AD_____

Award Number: DAMD17—94-J-4203

TITLE: Developing New Epidemiologic Tools for Investigating Breast Cancer Risk

PRINCIPAL INVESTIGATOR: Peter Gann, M.D., Sc.D.

CONTRACTING ORGANIZATION: Northwestern University
Evanston, Illinois 60208-1110

REPORT DATE: September 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1999	Annual (1 Sep 98 - 31 Aug 99)

4. TITLE AND SUBTITLE
Developing New Epidemiologic Tools for Investigating Breast Cancer Risk

5. FUNDING NUMBERS
DAMD17-94-J-4203

6. AUTHOR(S)
Peter Gann, M.D., Sc.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Northwestern University
Evanston, Illinois 60208-1110

E-Mail: pgann@northwestern.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

The overall aim of this research was to develop novel biomarkers for breast cancer prevention research using samples of biological fluids and breast tissue. Project 1 dealt with measurement of growth factors and steroids in nipple aspirate fluid, as well as estradiol and progesterone in saliva. Project 2 dealt with developing a histological index for estimating the degree of lobular differentiation in normal breast tissue. We successfully completed development of assays for EGF and TGF-a in breast fluid, and showed that differences in levels between women are significantly greater than those between breasts or over time in individual women. We also found weak correlations between breast fluid EGF and mammographic breast density. An assay for estradiol in breast fluid is sensitive and reliable, but validation studies have not yet been completed. We developed salivary assays for progesterone and estradiol; the latter is the first reported direct assay. Our results show that these salivary assays can be useful tools in studies involving cumulative ovarian hormone levels in pre-menopausal women. In Project 2, we completed development of a reliable method for measuring lobular differentiation by histologic branching, and found evidence for greater differentiation in parous and recently pregnant women. However, our data do not indicate strong differences in lobule type between breast cancer cases and controls.

14. SUBJECT TERMS
Humans, Anatomical Samples, Breast Cancer

15. NUMBER OF PAGES
63

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

PG Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

___In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

___For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

___In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

___In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

___In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

PI - Signature 3/17/00 Date

Final Report for Grant DAMD17-94-J-4203

TABLE OF CONTENTS

Note: The paper in this report marked Appendix E contains unpublished data and should be considered proprietary. It should not be distributed without the knowledge of Dr. Peter Gann.

Final Report: Grant DAMD17-94-J-4203

Principal Investigator: Peter Gann, M.D., Sc.D.

INTRODUCTION

The overall aim of this project is to develop new biological markers that can improve epidemiological investigations into the etiology of breast cancer. Our studies encompass three types of novel biomarkers: a) *breast fluid from nipple aspirates for measurement of growth factors and steroids,* b) *saliva for measurement of sex steroid concentrations,* and c) *normal breast tissue from biopsy samples for assessment of lobular differentiation.* These biomarkers would allow epidemiologists to study the development of breast cancer in greater biological detail than previously possible using conventional questionnaire-based research and will also allow clinical investigators to test the effects of dietary or chemopreventive interventions. We have conducted a series of studies whose purpose is to evaluate the feasibility and validity of deploying these markers in clinical or epidemiological research aimed at breast cancer prevention. **To allow consistency with our previous reports, we will refer to all activities related to hormone or growth factor levels as Project 1; activities related to lobular differentiation will be referred to as Project 2.**

BODY OF REPORT

Project 1: Hormones/Growth Factor Levels in Body Fluids and Their Determinants

Task 1: Characterizing growth factor levels in breast fluid and lab assay variation

Sub-tasks here included developing immunoassays for 4 growth factors in pooled breast fluid samples, evaluating intra-run variability and variability due to freeze-thaw cycles, and evaluating inter-run variability. All 3 sub-tasks were completed for epidermal growth factor (EGF), transforming growth factor-alpha (TGF-α), transforming growth factor-beta 1 (TGF-β1), and insulin-like growth factor 1 (IGF-1). We used expressed prostatic fluid (EPF) in some cases for assay development, because this fluid is similar to but more abundant than breast fluid, and because we had access to a large bank of frozen EPF samples.

Methods, Results and Discussion pertaining to this Task are fully described in the two publications appended to this report (**Appendices A and B**). Development of these assays proved successful in breast fluid for EGF and TGF-α, and additionally for TGF-β1 in EPF. As described in the last Annual Report, we encountered difficulty with the IGF-1 procedure in both breast fluid and EPF. We were able to detect predicted amounts of IGF-1 in serum, and concluded that there must be interference with the assay in breast fluid and EPF due to matrix effects (physical properties of the fluid) or presence of unidentified chemical inhibitors. An additional hypothesis is that IGF-1 is destroyed by proteases normally present in these biological fluids. TGF-β1 was easily detected in EPF – however, in breast fluid, we established an adequate intra-assay variability (<10%) and a wide range between subjects, but have not yet

overcome an unacceptable (>15%) degree of inter-assay variability. Further work on this assay could increase reproducibility across assay runs.

Task 2: Measuring intra-individual variation in breast fluid growth factors

We accomplished this task by recruiting women for the Repeat Sample Study in years 1 and 2 of the project period. Thirty-seven premenopausal volunteers provided breast fluid samples on 2-4 visits over a single menstrual cycle. The methods and results for EGF and TGF-α are described in the publication by Gann, et al in **Appendix A**. Our results indicate that breast fluid concentrations of these growth factors are consistent over time in individual women, and are relatively different between women. We also demonstrated that, within individuals, EGF secretion was correlated with secretion of TGF-α and with serum levels of estradiol.

Task 3: Correlating breast fluid growth factor levels and biopsy results

We were not able, for logistical reasons, to recruit women awaiting breast biopsy. However, we did complete two studies that related growth factor levels to tissue characteristics. In the first study, we compared levels of EGF, TGF-α and TGF-β1 in EPF from patients with benign hyperplasia, prostate cancer and normal prostates. These results were recently published (see **Appendix C**).

The Mammographic Density Study (MDS) was begun in project year 3 with the following aims: 1) to evaluate the association between breast fluid levels of EGF and TGF-α and breast parenchymal density as reflected in screening mammograms, and 2) to evaluate the association between these breast fluid growth factor levels and reproductive risk factors for breast cancer. Aim 2 also relates to SOW Task 4 in our original USAMRDC proposal. Mammographic density is increasingly recognized as a risk factor for breast cancer development and appears to be controlled, at least partially, by ovarian hormonal influences.[1] We hypothesized that women with characteristically high levels of estradiol and mitogenic growth factors in breast fluid will have increased mammographic density.

Following IRB approval, we began to collect breast fluid from women receiving mammograms at the Lynn Sage Breast Center at Northwestern. Women were eligible if they were between 35-60 years old, had no history of breast cancer, were at least 2 years post-lactation, were scheduled for a screening mammogram as opposed to a mammogram for follow-up of an abnormal finding, and had no nipple soreness or lesions that would preclude breast fluid sampling. After the films were taken, radiology technicians gave each eligible woman a cover letter and consent form to review and sign while waiting (still in examination gown) for the film quality to be checked. Thus, all nipple aspirations were performed soon after breast compression. Following nipple aspiration, an interviewer administered a brief questionnaire to obtain supplemental information on reproductive history, exogenous hormone use and family history of breast cancer. A staff member validated according to procedures described earlier in the project traced the whole breast area and areas of density on the cranio-caudal view mammogram. These tracings were then re-traced with a computerized planimeter to measure total breast area and total area occupied by radiographic densities. We completed breast fluid EGF measurements for 23 women in this study; and TGF-α for another 19. The correlation between EGF and breast density ($r=0.32$, $P = 0.08$) is shown in Figure 1. A similar analysis was done with TGF-α and breast density, however no correlation was shown ($r=0.12$, $P=0.58$)

Task 4: Association of growth factor levels with breast cancer risk factors

As mentioned above, the Mammographic Density Study is responsive to this Task, because breast density is a recognized independent risk factor for development of breast cancer. In addition, we have accumulated enough data on EGF and TGF-α levels in women from various studies to permit exploratory analyses examining possible determinants of these levels. We have combined data from the Repeat Sample Study (RSS – described in previous reports) and the MDS for these analyses. We have data on factors pertaining to lifestyle and reproductive history for each of these women, including smoking history, height, body mass index, use of oral contraceptives or estrogen replacement, alcohol consumption, menarche, menopausal status, number and timing of pregnancies, and lactation. Average values for breast fluid EGF for each individual were log-transformed to normalize the distribution for parametric analyses. Using SAS-PC, we examined the association between log EGF and the above variables.

The results indicate interesting associations between breast fluid EGF concentration and age, BMI, age at menarche, oral contraceptive use and smoking history. Women under age 35 had significantly lower EGF levels than women aged 35 and older. BMI was positively associated with breast fluid EGF, with borderline statistical significance. Eight women had menarche at less than age 12, and their EGF levels were significantly lower than those for women with older ages at menarche. Women who had ever used oral contraceptives had EGF levels almost 2-fold higher than those never using them. However, the strongest association observed was for smoking history. Past or current smokers had a mean EGF level of 781 ng/ml, compared to 324 ng/ml among never smokers ($P = 0.009$). We are performing similar analyses for TGF-α and plan to prepare a manuscript reporting results for both growth factors.

Task 5: [Proposed Addition to Original SOW] Development of new assays for estradiol and progesterone in saliva

Our work with saliva samples focused on development of new assay methods for estradiol and progesterone. These assays are direct assays - they avoid an extraction step that requires a large volume of sample and can introduce error. The new assays will provide ultrasensitive non-invasive methods for serial measurement of steroid hormone concentrations in premenopausal women. These are potentially important biomarkers for epidemiological and prevention research because the fluctuation of ovarian steroid hormone levels during the menstrual cycle complicates interpretation of values from single blood samples.[2] Moreover, the steroid concentrations in saliva are more representative of the active fraction of bioavailable hormone in tissues.[3]

To measure progesterone in saliva directly without extraction, we use an ultrasensitive competitive-binding radioimmunoassay (RIA). The antibodies for the assay were produced by Dr. Robert Chatterton, whose laboratory conducts the assay.[4] In the assay, 200 ul samples of saliva are mixed with radiolabelled progesterone and progesterone antibody. The resulting bound progesterone is separated using dextran-coated charcoal, counted in a scintillation counter, and counts are compared to a standard curve. For salivary estradiol, we previously evaluated a time-resolved fluoroimmunoassay (DELFIA), but have now switched to an ultrasensitive RIA (DSL, Webster, TX) that we have adapted and optimized for this project. In this assay, I^{125}-labeled estradiol competes with salivary estradiol for binding to antibody in buffer solution. Addition of anti-IgG results in the formation of antigen-antibody complexes, which are then precipitated and separated from unbound hormone by centrifugation. This assay thus

avoids potential problems caused by interference between substances in saliva and binding to the solid-phase secondary antibody used in DELFIA.

Optimization and validation of the assay for salivary estradiol was an arduous task, given our goal of measuring the low concentrations present in saliva without an extraction step. The RIA kit is optimized for measurement of estradiol in serum. We experimented with numerous antigen:antibody ratios and sample volumes to construct a standard curve that is linear and steep within the expected range of sample results. We measured both intra- and interassay variability in a variety of quality control pool samples. We conducted serial dilution experiments comparing observed to expected estradiol values. We also added known amounts of estradiol to charcoal-stripped male saliva samples to compare observed to expected values. Our final procedure calls for 400 ul of saliva for each assay. Details regarding the methods, results and interpretation of our salivary assay development work are provided in our published paper (**Appendix D**). A second manuscript concerning the 'tracking' of salivary steroid levels *between* menstrual cycles in individual women has been approved by co-authors and is now ready for submission (**Appendix E**).

Task 6: [Proposed Addition to Original SOW] Development of assay for estradiol in breast fluid

In developing the assay for estradiol in breast fluid we investigated 1) different solvents for extraction to minimize blank values, 2) measured concentrations from a breast fluid pool after extraction with ethyl acetate and after solvent partition into aqueous 0.4 M NaOH from isooctane, 3) the parallelism of response with dilution, 4) the ability to measure known amounts of estradiol added to a breast fluid pool.

1) Extraction of estradiol with ethyl ether is subjected to extreme variability due to impurities in the solvent. Redistillation is helpful but requires special equipment. We tested ethyl acetate, a solvent with similar polarity for extraction efficiency and blank values in the assay. The blank value from 5 samples was 20 ± 8 pg/ml.

2) Samples from the same pool were extracted on 3 different occasions. The value after ethyl acetate extraction without correction for recovery was 752 ± 88 pg/ml. After partition into 0.4 M NaOH from isooctane, neutralization of the solution, and reextraction into ethyl acetate, the values without correction for recovery were 676 ± 72 pg/ml. Recovery of ^{3}H-estradiol in the solvent partition procedure was 94.8 ± 0.4%. The corrected value after solvent partition is therefore 713 pg/ml, not significantly different from the value before the solvent partition purification step.

3) Parallelism of the response was tested by assaying volumes from 0.0 to 12.5 μl in duplicate. The response was linear as shown in Figure 2.

4) Measurement of estradiol added to BF (Figure 3). Estradiol in the amounts, 0.0, 1.25, and 2.5 pg/tube was added to 10 μl and 20 μl volumes of a pool of BF. The value of the zero added dose, 0.42 pg and 0.78 pg, respectively, was subtracted for display of the data graphically. The average recovery of added estradiol was 115%.

Fractionation of steroids from BF. We have modified a previous procedure for separation of steroids into aqueous (steroid sulfates), neutral (androgens), and phenolic (estradiol and

estrone) fractions. The breast fluid is diluted with water and neutral and phenolic steroids are extracted into ethyl acetate. After evaporating the ethyl acetate, the residue is partitioned between isooctane and 0.4 M aq. NaOH. The isooctane is evaporated and the steroids are dissolved in assay buffer. The NaOH fraction is neutralized with HCl, and phenolic steroids are extracted into ethyl acetate. The ethyl acetate is evaporated from the phenolic fraction, and the steroids are dissolved in assay buffer. Thus, steroid sulfates, androgens, and estrogens can be measured in a single volume of BF instead of requiring separate aliquots of BF for each. This will maximize the utilization of BF and contribute to the purity of the steroids being assayed. The recovery of the ^{3}H-steroids is shown below.

Steroid	Fraction	Percent Recovery
Estrone	Phenolic	95.4 ± 0.5
Estradiol	Phenolic	94.8 ± 0.4
Androstenedione	Neutral	93.4 ± 0.4
DHA	Neutral	89.9 ± 1.6
Estrone sulfate	Aqueous	92.8 ± 1.5

We recently completed breast fluid estradiol measurements for 8 women who participated in the Repeat Sample Study. For each of these women, we assayed 3 samples: one from the right and left breast obtained on the same day, and the third from either breast, obtained 2 weeks apart in time. This allowed us to evaluate right-left breast correlation, as well as individual correlation over time within a single breast. These preliminary results do not indicate strong correlations either between breasts or over time. The right-left correlation is –0.24 (P = 0.57), and the same breast over time correlation is 0.48 (P = 0.22). We are currently testing the hypothesis that these discrepancies are due to degradation of the RSS samples, which may have become dessicated by long-term storage. Therefore, we are repeating these analyses with freshly obtained breast fluid samples.

Project 2: Lobular Differentiation in Normal Breast Tissue

Project 2 deals with measurement of lobular differentiation in normal breast tissue. We postulate that it is feasible to use normal breast tissue from the margins of breast biopsies to obtain a histological index of the differentiation status of a woman's breast. This idea follows from the work of Russo, et al., which demonstrates the feasibility and usefulness of such a marker in a rodent model.[5] Data on human lobular differentiation previously has been obtained from breast reduction or autopsy specimens, and therefore are not abundant. Development of a histolologic differentiation marker that can be used in readily available tissue would allow us to begin epidemiological studies aimed at identifying the major influences on human breast differentiation. It is hypothesized that extensive lobular differentiation will protect against breast cancer development.

Task 1: Establishing measurement criteria and inter-reader reliability

We completed this task using a set of slides from breast biopsies with adequate amounts of normal tissue. Working together with our reference expert, Dr. Jose Russo, we established criteria for categorizing lobules according to the extent of lobular branching. Two breast pathologists at Northwestern – Drs. Robert Goldschmidt and Elizabeth Wiley – participated in iterative discussions with Dr. Gann and Dr. Russo concerning problems involved in arriving at a

valid and reproducible set of criteria for reading lobular type. Dr. Russo conducted a training session in Chicago and Dr. Goldschmidt spent two days reviewing cases with Dr. Russo in Philadelphia. Dr. Gann and Allison Ellman visited Dr. Jose Russo at Fox Chase Cancer Center in Philadelphia in September,1996. The purpose was to revise the standard criteria for scoring lobule type in histologic specimens, and review discrepancies between Dr. Russo and Ms. Ellman. After review of many cases, we concluded that inter-reader agreement would be improved by: 1. adopting strict criteria for presence of a central duct in order to count a lobule, and 2) excluding lobules with diffuse hyperplasia defined as greater than two epithelial cell layers of thickness.

Final protocol for scoring lobules

The reader's goal is to provide an estimate of the distribution of lobule type (types 1, 2 and 3) in normal breast tissue obtained by excision biopsy or mastectomy. Type 1, 2 and 3 lobules are distinguished solely on the basis of the number of acini contained within a lobular unit, as follows:

Type 1	2 - 29 acini
Type 2	30-79
Type 3	≥ 80

Slides stained with H&E are examined first at 40X to establish the suitability of the slide for reading (presence of lobular structures, absence of diffuse pathology or artifact), and the approximate number of slides that will be needed to complete the reading. Multiple slides are used for each patient. The reader then proceeds to type lobules at 100X, moving through adjacent fields in an organized manner. 200X magnification is used to resolve questions about specific lobules. Results for each patient are recorded on tally sheets that allow us to reconstruct the order in which lobules were read, by slide.

The following criteria must be met in order to count a lobule:

- Central duct must be visible (step sections can be used to identify a central duct), unless there are 80 or more acini already visible.
- There must not be artifact or atrophy that could distort the number of acini counted.
- There must be no more than 2 cell layers of epithelium in the acini, which would indicate hyperplasia.
- The lobule must be free of other definable pathology (fibrocystic change, adenosis, LCIS, etc.)

Analysis of data from scoring an unlimited number of lobules led us to conclude that improvement in inter-reader agreement is negligible after 100 lobules have been counted using the above procedures. Therefore, no more than 100 lobules are counted per case.

We then trained two readers – Dr. Elizabeth Bauer-Marsh and Allison Ellman – using a set of cases that had been consensus-read by the experts and was divided equally into a training and validation subset. The correlation for % lobule type 1 and type 2 for the validated readers versus Dr. Russo are shown in Figure 4.

These plots indicate a high level of agreement between readers, and that our standardized protocol allows us to readily validate new readers. Note also that the slopes of the line for inter-reader agreement are close to 1.0.

Task 2: The relation of lobular differentiation status to distance from primary breast lesion

Whole mount studies had indicated no significant systematic variation in lobule type across breast regions. To determine if the same was true for the histologic index, and to determine if distance from a tumor would influence lobule type, we examined slides from 6 mastectomy patients in detail (4 from previously stained sections and 2 from new sections made from formalin-fixed mastectomy tissue). We had a pathology technician make new slides from all quadrants of 6 recently-obtained mastectomy specimens - however, 4 of these had too few lobules present on the slides to permit scoring. For each patient, we read 10 slides from each breast quadrant, and recorded the distance of each section from the breast cancer.

In our homogeneity analyses, lobule type was essentially the same in all breast quadrants (including the involved quadrant) in patients with breast cancer. Lobule type was also the same for regions located either 2-4 cm or > 4 cm from the primary tumor (the number of lobules < 2 cm from tumors was insufficient for analysis). We counted 152-454 lobules per case in these analyses.

Task 3: Measure association of lobular differentiation status to breast cancer risk factors

We completed a case-control study on the relation of lobular differentiation to the risk of breast cancer. This study encompasses, and extends beyond Task 3 in the original SOW. The specific aims of this study are to apply the index to conduct pilot studies exploring the relationship between lobular differentiation and a.) breast cancer risk, and b.) reproductive and hormonal variables hypothesized to be determinants of breast tissue maturation.

We randomly selected 32 patients under age 55 diagnosed with invasive or in-situ breast cancer since 1995 at Northwestern Memorial Hospital. We also randomly selected 30 patients who had recent biopsies that did not reveal malignant disease. Age was restricted to younger women to reduce distortion of lobular typing by age-related lobular regression. Slides from each patient were scored for lobule type by validated readers. Previous results indicate that the number of lobules read can be limited to 100. Dr. Wiley initially reviews all cases and all slides containing breast cancer are not sent to the reader to avoid bias. Analysis is based on the difference in lobule type distribution (types 1, 2 or 3) in cases versus controls. We hypothesize that cases will have a higher percentage of type 1 lobules and a lower percentage of type 3.

We identified 4 patients in our breast cancer series (2 infiltrating, 2 DCIS) who were diagnosed with breast cancer within 2 years postpartum. Because pregnancy is hypothesized to produce a temporary shift towards a more mature lobular pattern, we examined these cases separately. In fact, the mean % type 1 lobule among these cases was quite low compared to all controls - 61.4% -and the % type 2 was elevated – 26.5%. One woman who was diagnosed during pregnancy had the highest % type 3 we observed – 36%. Although these data are sparse, they lead us to conclude that recent childbirth (or at least progression to late pregnancy) is associated with a differentiated lobular pattern. These 4 cases are therefore excluded from further analysis to reduce bias. We also identified one case with a predominant lesion of fibroadenoma and a small focus of lobular neoplasia (also termed LCIS). We elected to exclude this case also because the distinction between lobular neoplasia and severe atypical hyperplasia – which is represented in the "benign" group – is very small.

Table 1 shows the comparison of lobule type distribution in breast cancer cases versus benign controls. Considering all subjects, breast cancer cases (including DCIS) had a slightly higher

percentage of type 1 lobules and slightly lower percentage of type 2, but the difference is easily consistent with chance (P = 0.36 for type 1). Six patients in the benign group were under age 30 - including 4 with a diagnosis of fibroadenoma, 1 with stromal fibrosis and 1 breast reduction. These subjects had markedly higher percentages of immature lobules. When we restrict analysis to subjects 30 years of age or older, the age of cases and controls becomes similar. In this subgroup, the mean type 1 % for cases is 7.3% higher than the type 1 % for controls, and the type 2 % is 5.7 % lower. These differences do not reach conventional statistical significance levels, however. We also stratified subjects by parity status. All parous subjects were older than 30. In this stratum, breast cancer cases had a type 1 % 10.6% higher and type 2 % 9.8% lower than controls. These results were also of borderline statistical significance.

Table 1. Distribution of lobule types in breast cancer cases versus controls with benign breast histology (all P values are two-sided)

	N	Type 1,%	Type 2,%	Type 3,%	P value, Type 1	P value, Type 2
All subjects						
breast cancer	27	82.8	12.6	4.6		
benign	30	78.9	16.0	5.5	0.36	0.24
Subjects >=age 30						
breast cancer	27	82.8	12.6	4.6		
benign	24	75.5	18.3	6.2	0.08	0.05
Parous, >=age 30						
breast cancer	16	83.2	12.4	4.4		
benign	12	72.6	21.2	6.0	0.09	0.05

Altogether, the results of this pilot study suggest a possible relation between the histologic lobule differentiation marker and breast cancer, but the data are far from conclusive. The greatest difference between cases and controls was observed among parous women – which is consistent with the hypothesis that some women have a sub-optimal differentiation response to pregnancy and thus acquire a persistently elevated risk of breast cancer. A larger study would provide more definitive answers. In addition, a larger study would allow for additional control over factors such as parity, age, lactation and time since last pregnancy. Our data do indicate that during a relatively short interval after childbirth, lobular differentiation shifts to a more mature pattern. Future studies should include larger numbers of older women – despite the tendency for lobule type to regress with aging – because a substantial proportion of breast cancers in younger women is related to inherited factors that might not impinge on differentiation.

We plan to complete a manuscript describing these data. We have held discussions with colleagues who have access to a large set of archived biopsy sections from women who subsequently did or did not develop breast cancer. These materials would be ideal for a nested case-control study of our differentiation marker. However, due to the labor involved (an average of 1-1.5 hours per subject) in screening and reading slides and the weakness of the results from our pilot study, we have decided to postpone a larger study. We plan to discuss our results as widely as possible in order to solicit suggestions as to how and whether the histologic marker can be improved. We continue to investigate automated image analysis, which would be an ideal solution, but current methodology is still not adequate for the pattern recognition required.

We are also investigating immunohistochemical markers (such as inhibin, casein and TGF-β) that could reflect differentiation status.

KEY RESEARCH ACCOMPLISHMENTS

- Publication of the first paper describing a comprehensive evaluation of assays for EGF and TGF-α in breast fluid. (This research also contributed to the publication of a second paper, dealing with characterization of EGF and TGF-α in expressed prostatic fluid)

- Detection of assayable quantities of TGF-β1 and estradiol in breast fluid. Demonstration of a breast fluid estradiol assay that is sensitive and reproducible.

- Demonstration of a possible association between breast fluid EGF levels (but not TGF-α) and mammographic breast density.

- Completion of development of reliable, valid assays for salivary progesterone and estradiol. The estradiol assay is the first direct (non-extraction) assay to be developed, and our reliability/validation data are extensive. These could become valuable tools in studies of cumulative hormone level and cancer risk in premenopausal women.

- Completion of data collection and analysis of studies using salivary estradiol and progesterone to estimate the cycle-to-cycle variability of integrated hormone levels in premenopausal women. These studies, because they avoid daily venipuncture, are among the largest and most complete to date on cycle-to-cycle variability in ovarian steroid level.

- Published manuscript on differences in TGF-α, EGF and TGF-β in expressed prostatic fluid from men with prostate cancer, benign hyperplasia and normal prostates. This work relates closely to the assay development work we are undertaking with breast fluid, and is currently being used to improve the TGF-β assay for breast fluid.

- Development of a reproducible method for classifying normal breast lobules by degree of differentiation (branching complexity) in tissue samples obtained at biopsy.

- Completion of a pilot case-control study on the relation of breast lobular differentiation pattern to breast cancer development. We found weak associations between nulliparity and less developed lobular patterns, and have some data corroborating the effect of recent completed pregnancy on lobular differentiation. Our data also provide suggestive evidence that an association between immature lobule type and breast cancer might exist among parous women.

REPORTABLE OUTCOMES

Manuscripts (see Appendices A-E)

Gann P, Chatterton R, Vogelsong K, Dupuis J, Ellman A: Mitogenic growth factors in breast fluid obtained from healthy women: evaluation of biological and extraneous sources of variability. *Cancer Epidemiol Biomarkers Prev* 1997;6:421-428.

Gann PH, Chatterton R, Vogelsong K, Grayhack JT, Lee C: Epidermal growth factor-related peptides in human prostatic fluid: sources of variability in assay results. *Prostate* 1997;32:234-240.

Gann PH, Klein KG, Chatterton RT, et al: Growth factors in expressed prostatic fluid from men with prostate cancer, BPH, and clinically normal prostates. *Prostate* 1999;40:248-255.

Lu Y, Bentley GR, Gann PH, Hodges KR, Chatterton RT: Salivary estradiol and progesterone levels in conception and nonconception cycles in women: evaluation of a new assay for salivary estradiol. *Fertil Steril* 1999;71:863-868.

Gann PH, Giovanazzi S, Van Horn L, Branning A, Chatterton RT: Saliva as a medium for investigating intra- and inter-individual differences in sex hormone levels in premenopausal women. *Manuscript completed.*

Abstracts/Invited Presentations

Gann PH, Chatterton R, Vogelsong K, Grayhack J, Lee C. Variation in prostatic fluid concentrations of EGF and TGF-α. The Coleman Foundation Symposium, Robert H. Lurie Cancer Center, Northwestern University, Evanston, IL, September, 1995

Gann PH, Chatterton R, Vogelsong KM, Dupuis J, Ellman A. Mitogenic growth factors in breast fluid obtained from healthy women: biological and extraneous sources of variability. Annual Meeting of the American Association for Cancer Research, San Diego, CA, March, 1997.

Gann PH, Ellman A, Wiley EL, Russo J. A histologic index of lobular differentiation in breast cancer cases versus controls with benign breast disease. 20th Annual San Antonio Breast Cancer Symposium, San Antonio, TX, December, 1997.

Gann PH. Developing new biomarker tools: growth factors in breast fluid and lobular differentiation. Plenary Session, Department of Defense Breast Cancer Research Meeting: an Era of Hope., Washington, DC, November, 1997.

Gann PH. Salivary hormone assays for measuring cumulative exposure to ovarian steroids. International Agency for Research on Cancer, Lyon, France, November, 1998.

Lu Y-c, Bentley GR, Gann PH, Hodges KR, Chatterton RT. Evaluation of a new assay for salivary estradiol: salivary estradiol and progesterone levels in conception and non-conception cycles. 5th European Symposium on the Application of Saliva in Clinical Practice and Research, Egmond Zee, Netherlands, May, 1999.

Gann PH, Giovanazzi S, Van Horn L, Branning A, Chatterton RT. Saliva as a medium for investigating inter-individual differences in ovarian steroid hormone levels. Annual Meeting of the American Society for Preventive Oncology, Bethesda, MD, March, 2000.

Gann PH, How are diet and hormonal patterns linked to breast cancer risk? New Approaches in Nutrition and Chronic Disease Prevention, University of Pittsburgh, November, 1997.

Funding applied for based on work supported by this award

R01 CA66695 (Gann) 8/5/95 - 5/31/00 NIH/NCI, $231,738, Hormonal Responses to a Low-Fat High Fiber and Soy Diet.
The goal of this project is to quantify the impact of a low-fat high-fiber (LFHF) diet and/or a soy supplement on hormonal patterns associated with breast cancer risk. A secondary purpose is to assess the convenience and precision of using daily saliva samples to measure ovarian steroids throughout the menstrual cycle.

RO1 CA66691 (Van Horn), 8/5/95 – 5/31/00, NIH/NCI, $282,943, Low-Fat High-Fiber Soy Rich Diet in Premenopausal Women
Adherence and biologic response to a diet low in fat and high in fruit, vegetables, whole grains and fiber (LFHF diet) will be measured in healthy, free-living premenopausal women. Dietary assessment and biochemical measures of total fatty acids, antioxidants and phytoestrogens will be used to quantify adherence and biologic responses within and between groups. The supplemental effect of soy on these parameters, in combination with both the LFHF or usual diet, will also be evaluated.

(Gann), 1/01/00-12/31/01, American Cancer Society Rubloff Trust, $405,462
The Effect of Soy Phytoestrogens on Breast Tissue Proliferation Markers: A Randomized, Placebo-controlled Trial
This project will test the effects of supplemental soy phytoestrogens on biomarkers of proliferation and cell growth regulation in women at high risk for breast cancer. Participants will consume 60 grams of soy protein per day; placebo will consist of soy from which isoflavones have been extracted. Fine needle aspiration samples from the breast will be obtained before and after 6 months of treatment, along with breast fluid and blood samples. Immunohistochemistry will provide measures of Ki67, EGF-R, ER-α, PR , p53 and HER-2/neu in normal breast tissue. Ancillary studies will measure effects of the soy on markers of DNA oxidation damage.

(Chatterton), 9/1/00 – 8/31/05, NIH/NCI, $150,000
Regulation of Estradiol Levels in the Breast
This project, part of a Breast Cancer SPORE proposal, has an overall objective of elucidating factors that explain variation in levels of estradiol within the breast, as measured by concentrations in nipple aspirate fluid. The project utilizes four separate study populations to determine the associations between breast fluid estradiol and the following: reproductive history and established breast cancer risk factors, body mass index, androgen levels in serum and breast fluid, salivary and serum estradiol, breast fluid IL-6 and IL-13, tamoxifen exposure and oral contraceptive use.

Personnel receiving salary support from this award

Amy Klein
Yu-cai Lu, MD
Amy Branning
Kelly Hodges
Hui-yan Cai

CONCLUSIONS

This award allowed us to complete development of new assays for growth factors in breast fluid and steroid hormones in saliva. These assays are now ready to be deployed by investigators examining the effects of chemopreventive or lifestyle factors on intermediate markers of breast cancer risk. The saliva assays for estradiol and progesterone are particularly important for measuring cumulative hormone exposure in women undergoing menstrual cycling. We did not complete development of the assay for breast fluid estradiol, although considerable progress was made. Future work must be done on fresh breast fluid samples to determine the correlation between breasts in individual women as well as the correlation in levels from a single breast over time. This assay will be extremely useful, because it may provide an estimate of estradiol levels within the breast tissue itself. The award also allowed us to develop a standardized procedure for measuring lobular differentiation by histologic features on breast biopsy slides. We have shown that this procedure is reproducible between observers and between regions of the same breast. However, our comparison of lobular differentiation in breast cancer cases versus controls did not reveal associations as strong as we had hoped. Future work should focus on refining and perhaps automating the lobule counting process. Successful development of a simple measure of lobular differentiation would be very significant, because it would provide researchers with a tool for opening a new path for prevention of breast cancer, by promoting differentiation of immature and therefore vulnerable breast tissue.

References

1. Byrne C, Schairer C, Wolfe J, et al: Mammographic features and breast cancer risk: effects with time, age, and menopause status. *J Natl Cancer Inst* 1995;87:1622-1629.

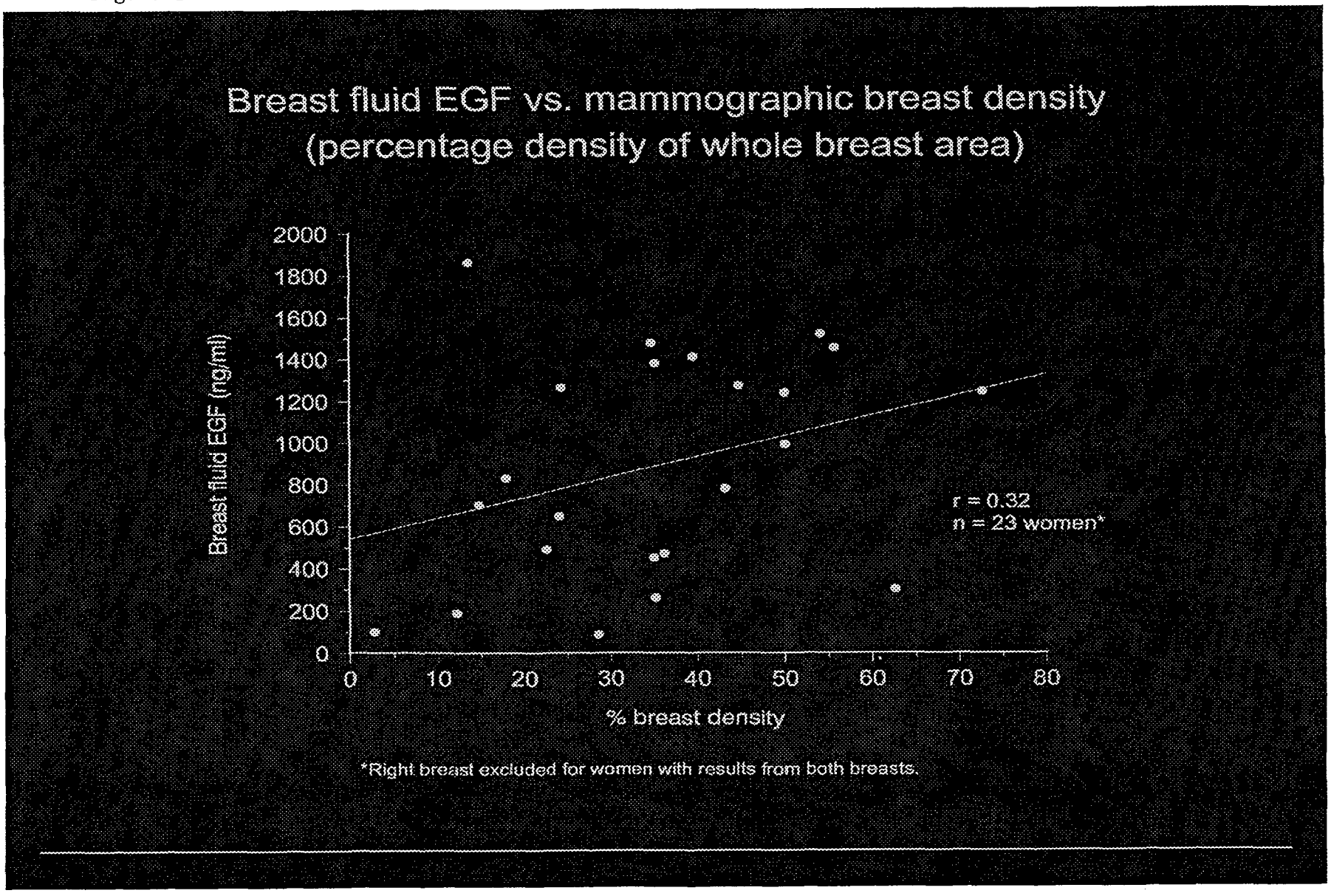
2. Michaud DS, Manson JE, Spiegelman D, et al: Reproducibility of plasma and urinary sex hormone levels in premenopausal women over a one-year period. *Cancer Epidemiol Biomarkers Prev* 1999;8:1059-1064.

3. Riad-Fahmy D, Read GF, Walker RF, Walker SM, Griffiths K: Determination of ovarian steroid hormone levels in saliva: an overview. *J Reprod Med* 1987;32(4):254-272.

4. Lu YC, Chatterton RTJ, Vogelsong KM, May LK: Direct radioimmunoassay of progesterone in saliva . *J Immunoassay* 1997;18:149-163.

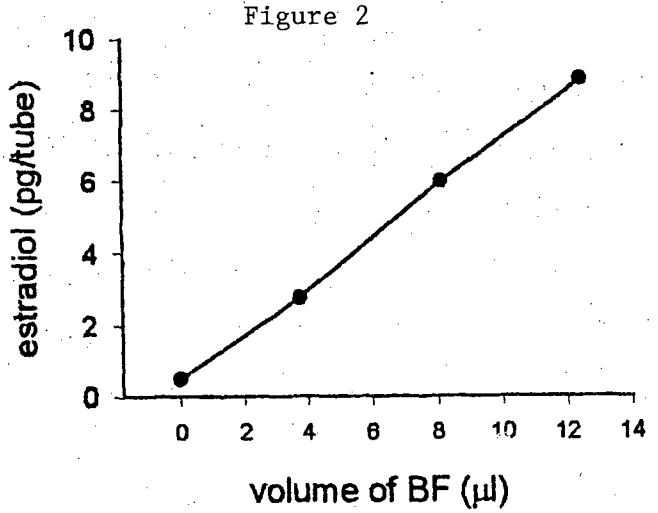
5. Russo J, Russo IH: Toward a physiological approach to breast cancer prevention. [Review]. *Cancer Epidemiol Biomarkers Prev* 1994;3:353-364.

Figure 1



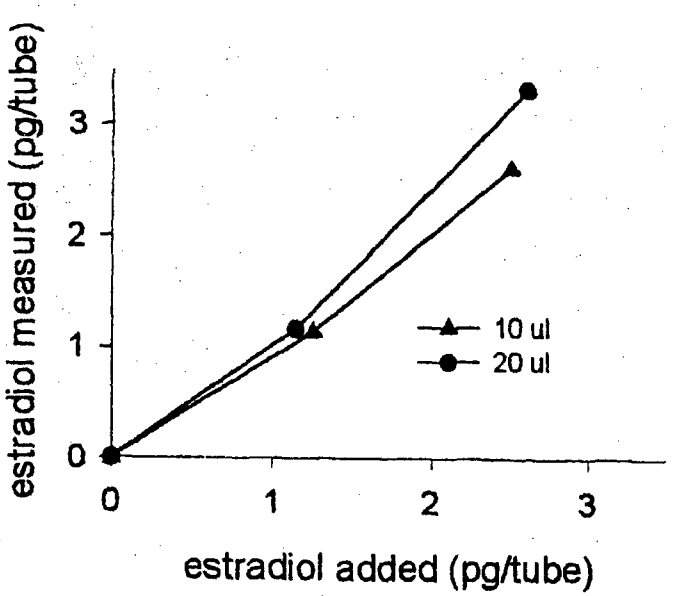
r=0.32, P=0.08

Figure 2



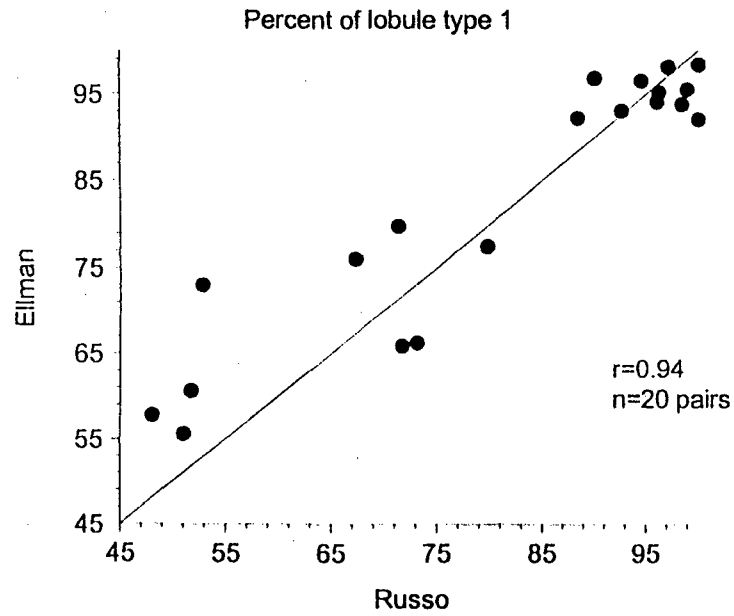
Breast fluid estradiol dilution experiment

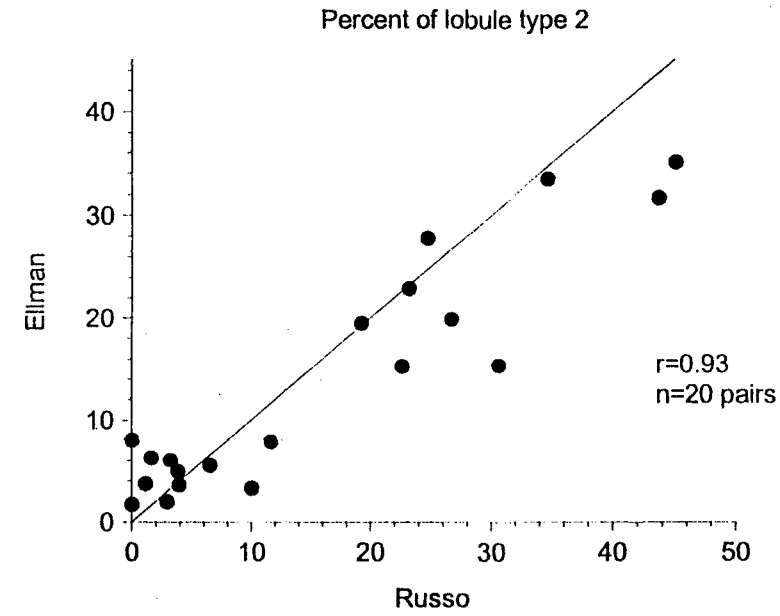
Figure 3

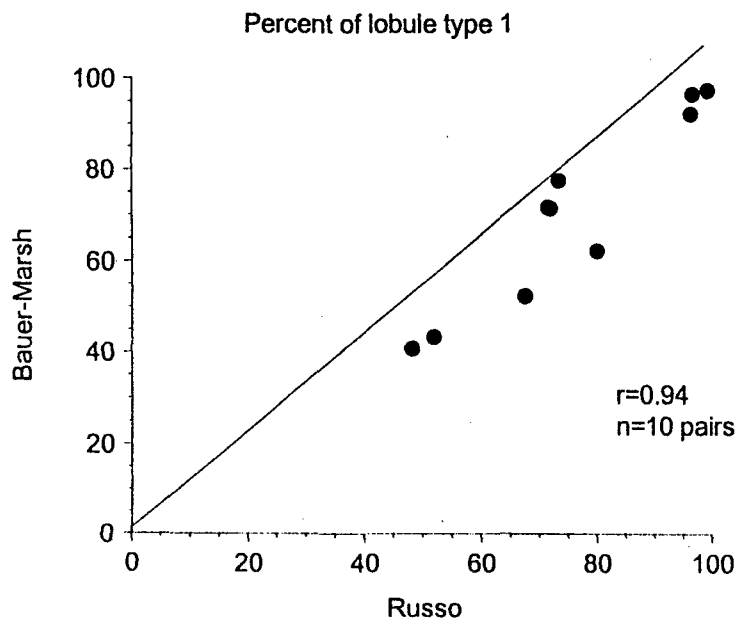


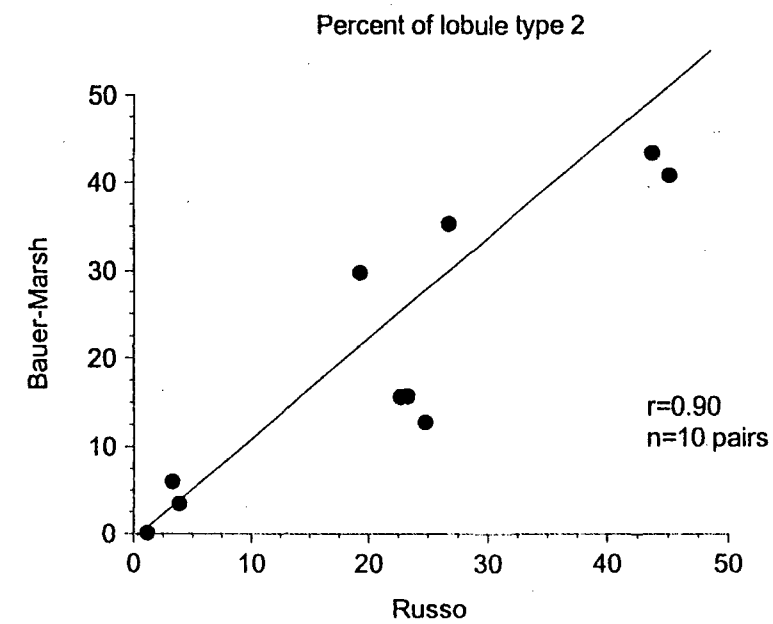
Breast fluid estradiol add-back experiment

Figure 4: Scatterplots of percent of lobule type: correlation between readings of non-professional (Ellman and Bauer-Marsh) readers vs. an expert reader (Russo)









Appendices A-E

Appendix A

Mitogenic Growth Factors in Breast Fluid Obtained from Healthy Women: Evaluation of Biological and Extraneous Sources of Variability[1]

Peter Gann,[2] Robert Chatterton, Kirsten Vogelsong, Josée Dupuis, and Allison Ellman

Departments of Preventive Medicine [P. G., J. D., A. E.], Obstetrics and Gynecology [R. C., K. V.], and Physiology [R. C.], Northwestern University Medical School, Chicago, Illinois 60611

Abstract

Peptide growth factors (GFs), including epidermal GF (EGF) and transforming GF-α (TGF-α), are presumed to play an important role in the local regulation of breast cell proliferation. Breast fluid collected by nipple aspiration provides a potential means to assess the concentration of these factors in contact with the ductal epithelium. Although identification of immunoreactive EGF-like GFs in breast fluid has been reported previously, we performed this study to evaluate the sensitivity and reliability of newer RIA methods and to characterize the sources and amounts of both intra- and intersubject variability. We also evaluated the relationship of breast fluid EGF and TGF-α levels to each other and to plasma levels of estradiol and progesterone. Breast fluid and plasma samples were obtained two to four times at weekly intervals from 18 healthy, premenopausal women. EGF and TGF-α were measured by competitive binding RIA. Both GFs were detected with good precision in all breast fluid samples analyzed, using dilutions as low as 1:100 for EGF (1 μl) and 1:25 for TGF-α (4 μl). The correlations between the right and left breasts, sampled concurrently, were $r = 0.78$ ($P = 0.003$) for EGF and $r = 0.89$ ($P = 0.0001$) for TGF-α. For both GFs, the variation between women was substantially greater than the variation between breasts or over time in an individual woman, particularly for EGF, for which there were 100-fold differences between women in mean levels. When samples from multiple women were analyzed together, we found no apparent relationships between EGF and TGF-α levels or between either GF level and menstrual cycle phase or plasma hormone concentrations. However, in random effects analyses, EGF levels within an individual were significantly associated overall with both TGF-α ($P = 0.02$) and plasma estradiol levels ($P = 0.01$). These data, which are the first comprehensive results on the feasibility of measuring mitogenic GFs in breast fluid, support the conclusion that women secrete consistent and individually distinct levels of EGF and TGF-α and that, in at least some women, EGF secretion *in vivo* covaries with both TGF-α in breast fluid and circulating estradiol.

Introduction

The prevailing model of breast cancer development assigns an important role to locally acting autocrine/paracrine GFs[3] (1). Peptides such as EGF and TGF-α, which have a 30–40% amino acid homology, demonstrate potent mitogenic effects on human breast cancer cells *in vitro* (2). In addition, the EGF receptor and the homologous product of the oncogene c-*erbB2*, which bind both EGF and TGF-α, are overexpressed in a significant proportion of breast cancers, particularly those with a poor prognosis (3, 4). Undoubtedly, these GFs, the structures of which are highly conserved across species, also play a role in control of normal breast cell proliferation. Current evidence indicates that estradiol and antiestrogens such as tamoxifen have direct and opposite effects on production of these GFs by epithelial or stromal cells in the breast (5). High levels of ovarian activity and of estrogen itself are related to increased breast cell proliferation and indeed provide the most cogent explanation for the increased breast cancer risk attributable to diverse factors such as age at menarche, age at menopause, and obesity. Overexpression of EGF-type GFs is strongly associated with mammary cancer in transgenic mice and the early stages of spontaneous mammary tumor development in normal mice (6, 7). It is plausible, therefore, to hypothesize that healthy women with excessive production of these mitogenic GFs have an elevated risk of developing breast cancer.

Because EGF and TGF-α are produced locally and act locally, their concentrations in serum or urine are not necessarily relevant. On the other hand, nipple aspiration provides a noninvasive method for sampling fluid that is in close contact with ductal epithelial cells. Several groups of investigators have demonstrated that a small volume of breast fluid can be obtained from 40–70% of nonlactating women by using a simple pump-like device (8). On the basis of numerous biochemical analyses, this fluid appears to provide reasonable insight into the hormonal and metabolic microenvironment of the breast. One group of investigators has reported immunoreactive EGF and TGF-α concentrations in breast fluid (9, 10). Other studies

Received 11/1/96; revised 2/20/97; accepted 2/24/97.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
[1] This work was supported by National Cancer Institute Grant KO7 CA66185, United States Army Medical Research and Materiel Command Grant DAMD17-94-J-4203, National Center for Research Resources (NIH) Grant RR-00048, the Lynn Sage Breast Cancer Research Fund, and the Shellie Harris Memorial Fund.
[2] To whom requests for reprints should be addressed, at Department of Preventive Medicine, Northwestern University Medical School, 680 North Lake Shore Drive, Suite 1102, Chicago, IL 60611. Tel: (312) 908-8432; Fax: (312) 908-9588; E-mail: pgann@nwu.edu.

[3] The abbreviations used are: GF, growth factor; EGF, epidermal GF; TGF-α, transforming GF-α; CV, coefficient of variation; QC, quality control; ICC, intraclass correlation coefficient.

Table 1 Intra-assay and interassay variability for replicate samples (including QC pools and individual subject samples) assayed at various dilutions: EGF and TGF-α in breast fluid

	Dilution	Mean intra-assay CV (no. of replicate pairs)	Mean interassay CV (no. of replicate pairs)
EGF	1:50	0.12 (26)	0.05 (6)
	1:75	0.04 (2)	
	1:100	0.13 (36)	
TGF-α	1:25	0.11 (59)	0.11 (4)
	1:50	0.25 (11)	

Table 2 Breast fluid GF levels: variation between and within Repeat Sample Study participants (left *vs.* right breast, same day)

	No. of subjects	No. of samples	Mean	CV between	CV within	ICC
EGF	12	24	604 ng/ml	0.82	0.47	0.48
TGF-α	13	26	2.26 ng/ml	0.46	0.11	0.88
EGF/protein	12	24	7.11 ng/mg	0.74	0.65	0.08
TGF-α/protein	13	26	39.1 pg/mg	1.49	1.16	0.21

have reported on these or similar GFs in human milk or breast cyst fluid, although the comparability of milk or cyst fluid to nipple aspirate fluid is questionable (11–13).

The studies described in this report were designed to address basic methodological questions concerning the assay of EGF and TGF-α in breast fluid. We evaluated the sensitivity and reproducibilty of these assays and the effect of specimen handling and storage. We then explored the variation in levels within women between breasts and within the same breast over time, comparing these within-woman variations to the amount of variation seen between women. For a biomarker to be useful in clinical or epidemiological research, it is critical that there be a substantial amount of variation between individuals relative to the variation within individuals (14). Finally, we determined whether breast fluid EGF and TGF-α levels were related to the menstrual cycle phase, to plasma estradiol or progesterone levels, or to each other. By repeat sampling of individual women, we were able to assess these relationships within individual women, as well as in the group as a whole.

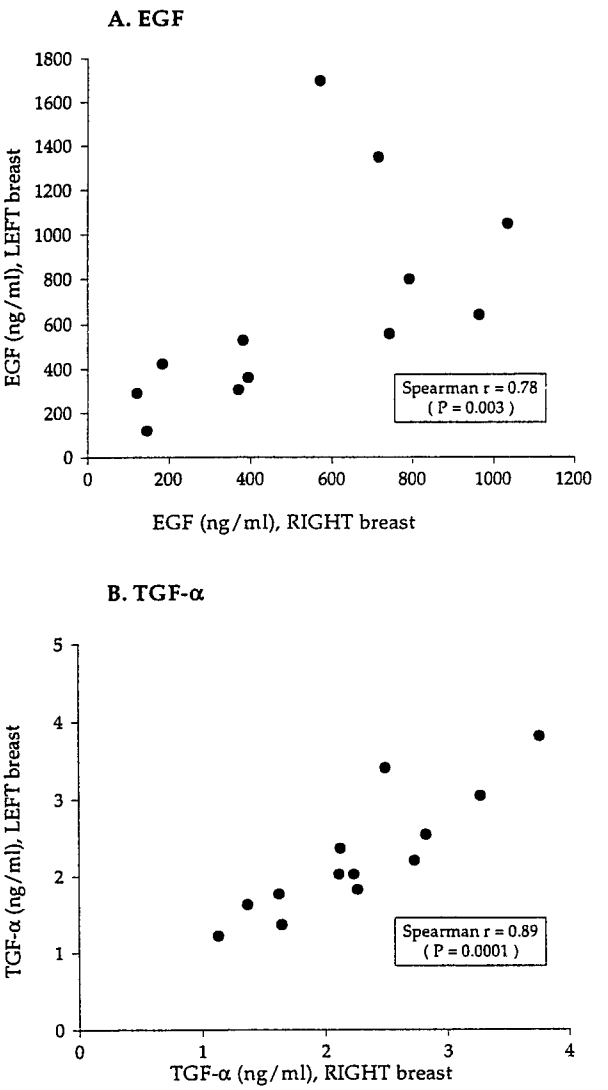


Fig. 1. Scatterplots showing concentrations of EGF (*A*) and TGF-α (*B*) in breast fluid obtained at the same time from the right *versus* left breast. EGF, $n = 12$; TGF-α, $n = 13$.

Subjects and Methods

Study Population and Sample Collection. Following approval of the protocol and informed consent procedures by the Institutional Review Board, we recruited women from the Chicago area to participate in the Repeat Sample Study. Criteria for eligibility included: ages 25–45 years, no history of breast cancer, regular menstrual periods, no lactation within 6 months, no use of oral contraceptives or other exogenous hormones within 6 months, and no major concurrent illnesses. Sixty-five eligible women were scheduled for four outpatient appointments each, 1 week apart, at the Clinical Research Center at Northwestern Memorial Hospital. Participants were allowed to start their visits during any day of the menstrual cycle, and they arrived at the Clinical Research Center in the morning after an overnight fast. The position of each visit day in the menstrual cycle was determined by recording the dates of onset of all menstrual bleeding immediately prior to and after the four visits. This allowed cycle position to be estimated by reverse dating, in which the midcycle day is defined as the first day of bleeding minus 14 days, the average length of the luteal phase. At each visit, we collected plasma and breast fluid. In addition, we measured body size and fat composition, physical activity, and dietary intake, for analyses not presented here. The mean age of the participants was 33.7 years; 76% were white, 11% were African-American, and 4% were Asian.

Breast Fluid Collection. At each visit, a trained nurse attempted to aspirate breast fluid from both breasts of each participant. After lightly scrubbing the nipple with a water-moistened gauze pad, the nurse asked the participant to compress the breast at its base with both hands. A sterile suction device made from a 20-cc plastic syringe body was then applied over the nipple, and vacuum pressure was gradually applied (15). Suction was discontinued if fluid failed to appear at the nipple surface after 10 s. Droplets of breast fluid appearing at the duct openings were collected in 75-mm plastic-coated capillary tubes that were then clay-sealed at both ends and kept on ice until storage at −20°C, no more than 1 h after collection. We obtained at least 2 μl of breast fluid from 39 (60%) of the 65 women who attended at least one visit. For the assays in this report, we selected samples from 18 women, 15 of whom gave samples at three or four visits and 13 of whom gave samples from both breasts on at least one visit.

Table 3 Breast fluid GF levels: variation between and within Repeat Sample Study participants (same breast, different days)

	No. of subjects	No. of samples	Mean	CV between	CV within	ICC
EGF	15	57[a]	494 ng/ml	1.20	0.26	0.83
TGF-α	15	58	2.68 ng/ml	0.78	0.42	0.37
EGF/protein	15	57	6.87 ng/mg	1.61	1.05	0.24
TGF-α/protein	15	58	73.2 pg/mg	3.81	2.58	0.21

[a] This sample was not included in analysis of EGF levels due to a laboratory error.

Assays for Plasma Estradiol, Bioavailable Estradiol, and Progesterone. Plasma estradiol was measured using a RIA kit obtained from Diagnostic Systems Laboratories (Webster, TX). This double antibody assay has a sensitivity of 29 pmol/liter and low cross-reactivity with forms of estrogen other than estradiol. The mean intra- and interassay CVs based on blinded QC samples were 0.08 and 0.04, respectively. Sex hormone binding globulin-bound estradiol was measured using the concanavalin A-Sepharose method described by Bonfrer *et al.* (16). This assay gave intra- and interassay CVs of 0.04 and 0.09, respectively, in our QC samples. Plasma progesterone concentrations were measured by RIA using antibodies prepared by Dr. R. Chatterton (17). The intra- and interassay CVs were 0.09 and 0.10, respectively, in our QC samples.

Assays for Breast Fluid EGF, TGF-α, and Protein. Breast fluid was removed from the capillary tube while it was in a semi-frozen state, and, except when undiluted sample was needed for sensitivity studies, it was diluted with a Tris-saline buffer (pH 8.0) prior to assay. The initial dilution was made without BSA in the buffer, an aliquot was taken for protein assay, and the final dilutions were made with BSA to a concentration of 0.15%. We used competitive binding RIA kits purchased from BioMedical Technologies (Stoughton, MA) for both EGF and TGF-α. The EGF assay has an estimated cross-reactivity of less than 0.1% with human TGF-α and undetectable binding with other human peptides tested. The TGF-α assay has cross-reactivities of less than 0.1% with human EGF and undetectable binding with other human peptides tested. To evaluate assay sensitivity, dilutions of breast fluid ranging from undiluted to 1:200 were prepared. The intra-assay CV at each dilution was evaluated by analyzing replicates in the same assay run to determine the dilution level at which assay reliability became unacceptable. Total protein in breast fluid was measured by the Bradford method.

Data Analysis. We calculated CVs and ICCs to assess intra- and interassay variability and the amount of variation within *versus* between individuals. Interassay CV was calculated from the variance between assays with the intra-assay variance removed. The ICC is defined as the between-person variance divided by the total variance (between plus within; Ref. 18). To determine which method minimized extraneous variation, we compared within-person CVs for GF concentrations expressed per unit breast fluid volume to those expressed per weight of total protein. To compare right *versus* left breast results and to compare EGF *versus* TGF-α or either GF *versus* plasma hormone levels, we computed the nonparametric Spearman correlation coefficients (r). These coefficients are unbiased, but because there were multiple measurements from the same person, the conventional variance estimate of the coefficient estimates was too low. To obtain correct variance estimates and compute Ps for r, we used a permutation method, which generates an approximate distribution of r under the null hypothesis (19).

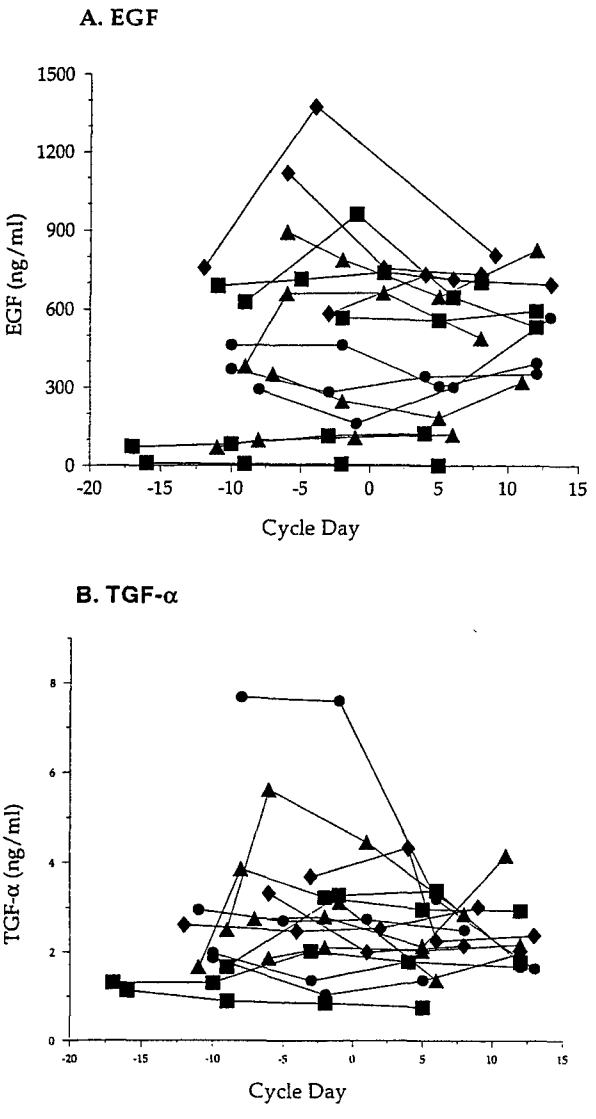


Fig. 2. Concentrations of EGF (*A*) and TGF-α (*B*) in breast fluid across the menstrual cycle: repeated measures from 15 women. Midcycle day (day 0) was determined by reverse dating from subsequent menses.

Growth factor and hormone concentrations were not normally distributed. To facilitate parametric analyses, we evaluated several data transformations and concluded that the square-root transformation provided the best normalization for the key variables as a group. Therefore, to compute mean GF concentration plus estimated 95% confidence intervals for each phase of the menstrual cycle, we obtained SEs and confidence intervals from the transformed data and then converted back to the original units for reporting purposes. The menstrual cycle was divided into the following six phases, with 0 designated as the midcycle day: early (days < -10), mid- (days -10 to -6), and late (days -5 to -1) follicular and early (days 0–4), mid- (days 5–9), and late (days 10–13) luteal. Ps for comparison of GF levels by cycle phase were obtained by random effects modeling using PROC GLM in SAS (SAS Institute, Inc., Cary, NC). We also used random effects models, with EGF as the dependent variable, to determine the degree of linear association between EGF and TGF-α and plasma estradiol (20). An inter-

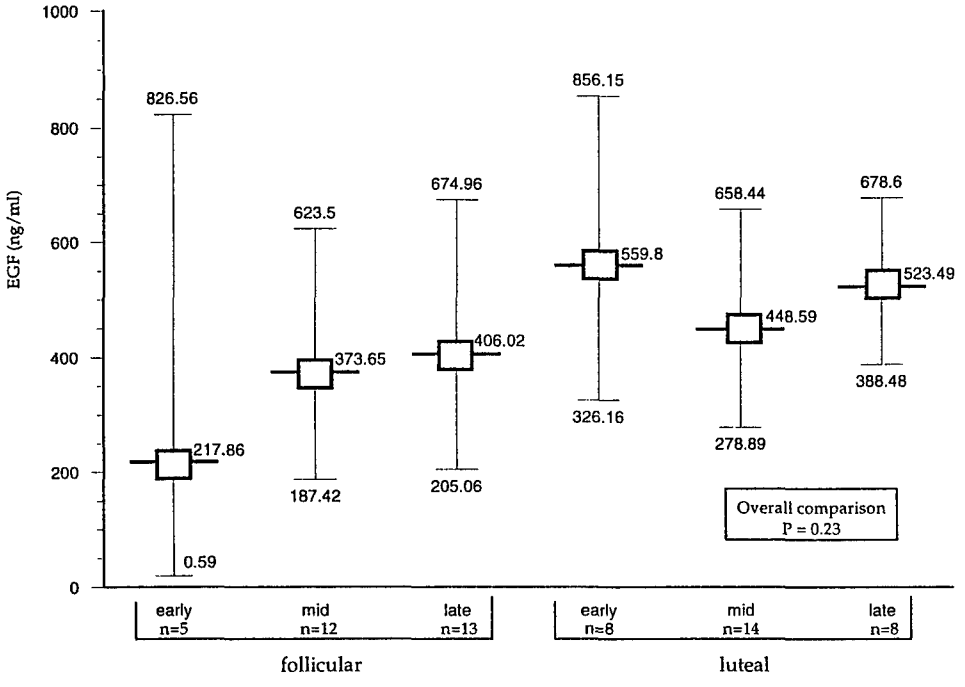


Fig. 3. Mean concentration (plus 95% confidence intervals) of EGF in breast fluid obtained during six phases of the menstrual cycle.

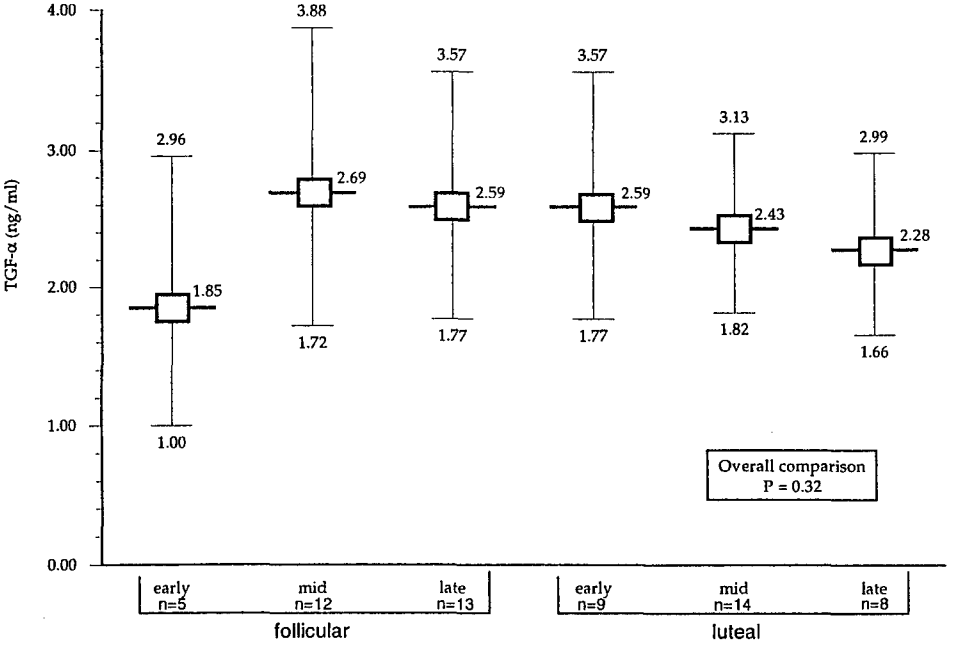


Fig. 4. Mean concentration (plus 95% confidence intervals) of TGF-α in breast fluid obtained during six phases of the menstrual cycle.

action term consisting of a binary dummy variable for each subject multiplied by the predictor level (TGF-α or plasma estradiol) allowed us to evaluate the biomarker associations within individuals. Similar models were developed with TGF-α as the dependent variable and EGF and plasma estradiol as the predictors. From these models, we obtained estimates of the total variance and within-woman variance explained by each model term. Although we had enough data to estimate the contribution of within-woman correlation to overall model fit, the small number of samples available per woman precluded testing hypotheses about correlation in specific individuals.

Results

Table 1 shows results indicating the sensitivity and reliability of the EGF and TGF-α measurements in breast fluid. For EGF, analysis of samples diluted 1:100 with assay buffer gave acceptable intra-assay reliability. Thus, we were able to obtain reliable results using only 1 μl of breast fluid, which contained concentrations in the range of 4–6 ng/ml in diluted samples from various QC pools. For TGF-α, both intra-assay and interassay reliability were acceptable at dilutions of 1:25, but not at 1:50. We were therefore able to reliably measure TGF-α in only 4 μl of breast fluid, with measured concentrations in pooled, diluted samples of about 0.1 ng/ml. EGF and TGF-α

concentrations in pools prepared from women with abundant *versus* scant volumes of breast fluid were indistinguishable. In one experiment, the number of freeze-thaw cycles (ranging from two to six) was not associated with any trends in measured GF concentrations.

Mean GF levels and results on the variation between the right and left breast are shown in Table 2. EGF levels were unobtainable from both breasts in one woman due to a laboratory error. For EGF, the within-woman variation (between breasts) was considerably less than the variation in EGF levels between women. The ICC implies that 48% of the total variance in EGF could be attributed to between-woman differences. For TGF-α, the within-woman variation between breasts was even lower, and the ratio of between-woman to within-woman variation was even higher. Eighty-eight % of the total variance in TGF-α was attributable to between-woman differences. Table 2 also shows that expressing GF levels per weight of total protein rather than per unit volume did not improve, and in fact substantially reduced, the level of agreement between breasts. Total breast fluid protein levels were correlated between breasts (data not shown).

Fig. 1 displays the agreement in EGF and TGF-α levels between breasts for individual women sampled in both breasts on the same date. Right and left breast levels were well correlated: EGF, $r = 0.78$ and $P = 0.003$; and TGF-α, $r = 0.89$ and $P = 0.001$.

Table 3 shows the variation in GF levels between and within women in the same breast over time. Once again, the variation between women was far greater than the variation within individual women over time for both GFs. The range of breast fluid EGF concentrations between women was extremely high (over 100-fold differences), and thus, the between-women variance for EGF was by far the dominant component of total variance (ICC = 0.83). Variances over time were not reduced by expressing results per weight of total protein rather than per volume. Fig. 2 shows the EGF (Fig. 2*A*) and TGF-α (Fig. 2*B*) results for each woman over time. This graph provides visual evidence that women tend to have highly distinct levels of EGF that are relatively consistent over time. TGF-α levels for individual women (Fig. 2*B*) also tend to remain stable over time, although the decreased variation between women, compared to EGF, is evident.

To more closely examine whether GF levels in breast fluid vary in conjunction with the menstrual cycle, we plotted the mean EGF and TGF-α concentrations for six cycle phases, as shown in Figs. 3 and 4. We found no significant differences for either GF across cycle phases ($P = 0.23$ and 0.32, respectively, based on a random effects model accounting for repeated measures). For EGF, there is the appearance of an increase during the luteal phase, but direct comparison of, for example, midluteal *versus* early or midfollicular EGF indicated that the differences were highly compatible with chance ($P = 0.61$).

In Fig. 5, EGF (Fig. 5*A*) and TGF-α (Fig. 5*B*) levels are plotted against concurrent total plasma estradiol levels. Neither GF was meaningfully correlated with plasma estradiol ($r = 0.15$ for EGF and $r = 0.02$ for TGF-α by the permutation method). We obtained similar results using plasma bioavailable (*i.e.*, non-sex hormone binding globulin bound) estradiol or progesterone instead of total estradiol.

The results shown in Fig. 6 indicate that, when all samples from all women were considered, EGF and TGF-α concentrations in the same sample were not well correlated ($r = 0.17$, $P = 0.50$). However, when we examined the EGF and TGF-α relationship for individual women, some striking covariation was apparent. Table 4 shows results from a random effects

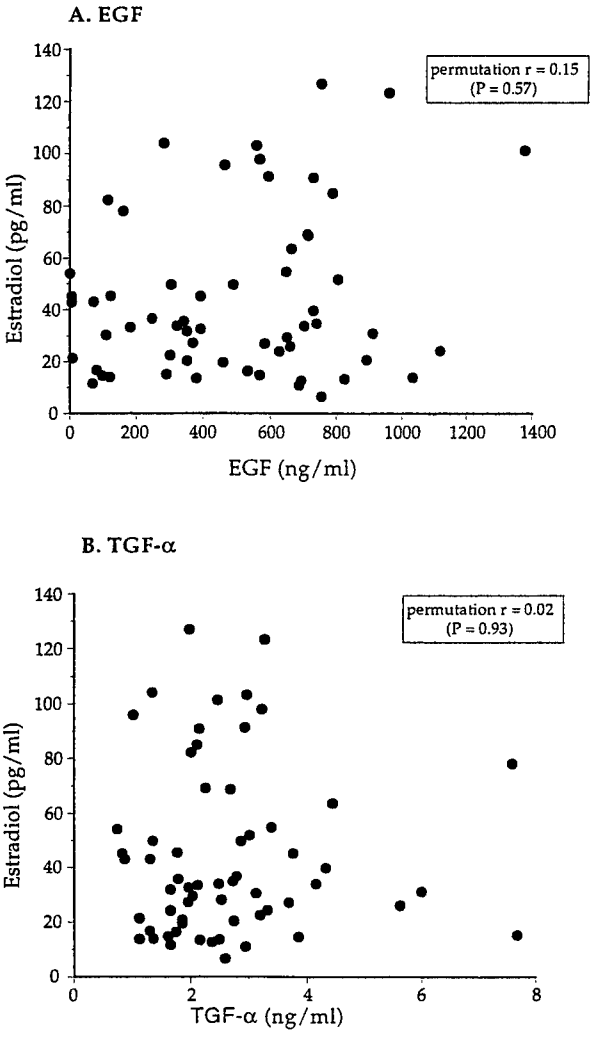


Fig. 5. Plasma estradiol *versus* breast fluid EGF (*A*) and TGF-α (*B*): contemporaneous samples from 18 women. Correlation coefficients and *P*s determined by a permutation test to account for multiple samples per woman. EGF, $n = 60$; TGF-α, $n = 61$.

model that evaluates the EGF-TGF-α association within women. This model includes a universal coefficient reflecting the common relationship of EGF to TGF-α, as well as a term reflecting the relationship for each individual woman. The universal coefficient was very small, consistent with the low r seen in Fig. 6. However, the association within individual women was statistically significant ($P = 0.02$) and explained nearly 56% of the variance in EGF within women. The high percentage (93.5%) of variance explained by simply specifying the individual subject corroborates the large amount of variation for EGF between women that was seen in the earlier analysis.

Table 5 shows a similar random effects analysis for the relationship of EGF to plasma estradiol. Again, although the overall correlation between EGF and estradiol was poor, the results indicate a significant correlation within individual women ($P = 0.01$). Fifty-eight % of the within-woman variance in EGF was explained by the plasma estradiol values. The within-woman association between breast fluid TGF-α and plasma estradiol was not substantial. Fig. 7 includes graphs of selected participants showing strong covariation between EGF

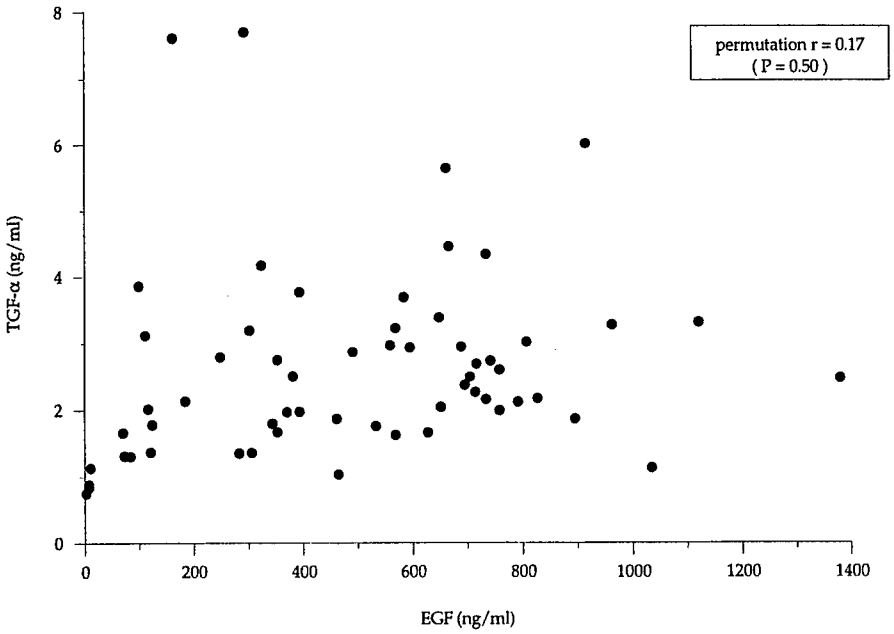


Fig. 6. Scatterplot of EGF and TGF-α concentrations measured in the same breast fluid sample. $n = 60$, obtained from 18 women.

Table 4 Random effects model evaluating the associations within individual women between breast fluid EGF and TGF-α[a]

Independent variable	Sum of squares	% variance explained	% within-woman variance explained	*P*
Individual subject	3808.17	93.49		
TGF-α	0.12	0.003	0.05	0.99
Subject × TGF-α	148.0	3.63	55.82	0.02
Error	116.87	2.87		

[a] Model: dependent variable, EGF; $n = 60$ samples from 18 women.

Table 5 Random effects model evaluating the associations within individual women between breast fluid EGF and total plasma estradiol[a]

Independent variable	Sum of squares	% variance explained	% within-woman variance explained	*P*
Individual subject	3808.17	93.49		
Plasma estradiol	2.39	0.06	0.92	0.65
Subject × plasma estradiol	154.66	3.80	58.37	0.01
Error	107.95	2.65		

[a] Model: dependent variable, EGF; $n = 61$ samples from 18 women.

and TGF-α levels (Fig. 7*A*) and covariation between EGF and plasma estradiol (Fig. 7*B*). These participants were not atypical, although the extent of covariation in some individuals appeared to be lower. The small number of replicates per woman, however, prevents us from drawing reliable conclusions about correlation in specific individuals.

Discussion

Peptide GFs such as EGF and TGF-α are potent signaling molecules for regulating the growth and perhaps differentiation of breast epithelial cells. Abnormal expression or activity of these factors could result from mutations of proto-oncogenes transcribing the GFs themselves or their receptors. Alternatively, because these GFs have a role in normal growth and therefore must be regulated by endogenous signals, abnormal expression could occur as a result of up- or down-regulation of gene transcription by compounds such as steroid hormones. The data in this report, although preliminary to the study of any relationships *in vivo* between breast cancer and GF expression, indicate that immunoreactive EGF and TGF-α can both be detected reliably in breast fluid from healthy premenopausal women and that individual women secrete distinctive amounts of these factors, amounts that are consistent both over time and between breasts. This study also provides evidence that levels of breast fluid EGF tend to covary over time with TGF-α and plasma estradiol, within individual women.

The presence of EGF in human milk has been demonstrated previously. In fact, Carpenter (11) reported that 90% of the mitogenic activity of human milk in cell culture systems was negated by a neutralizing antibody for human EGF. Connolly and Rose (9, 10) have reported previously the detection of EGF and TGF-α in breast fluid from healthy premenopausal women. They found concentrations of TGF-α in single samples from 21 women ranging from 0 to 50 ng/ml, with a median of 5.1 ng/ml, levels somewhat higher than we found. They also found EGF levels in samples from 17 women similar to ours; intriguingly, however, 9 of these women with unspecified biopsy-confirmed benign breast disease appeared to have higher EGF levels than the 8 women designated as controls (9). Although the results published previously probably required higher volumes of breast fluid for analysis, we found no difference in GF concentrations between women with abundant *versus* scanty breast fluid volume. We used highly sensitive RIAs that require only 1 and 4 μl of breast fluid for EGF and TGF-α, respectively, and therefore make it possible to obtain measurements even on women with scanty breast fluid samples. The median volume of breast fluid we obtain is approximately 25 μl, but the frequency distribution for sample volume is highly skewed, and many women have samples under 10 μl. Highly sensitive assay methods are therefore important. We found no evidence for an effect of thaw-refreeze cycles on GF concentrations, nor did we find any evidence for a decay in measured GF during 18 months of storage at temperatures of -15 to -20°C.

Several aspects of our findings require further elaboration. In looking at variation over time, we found that be-

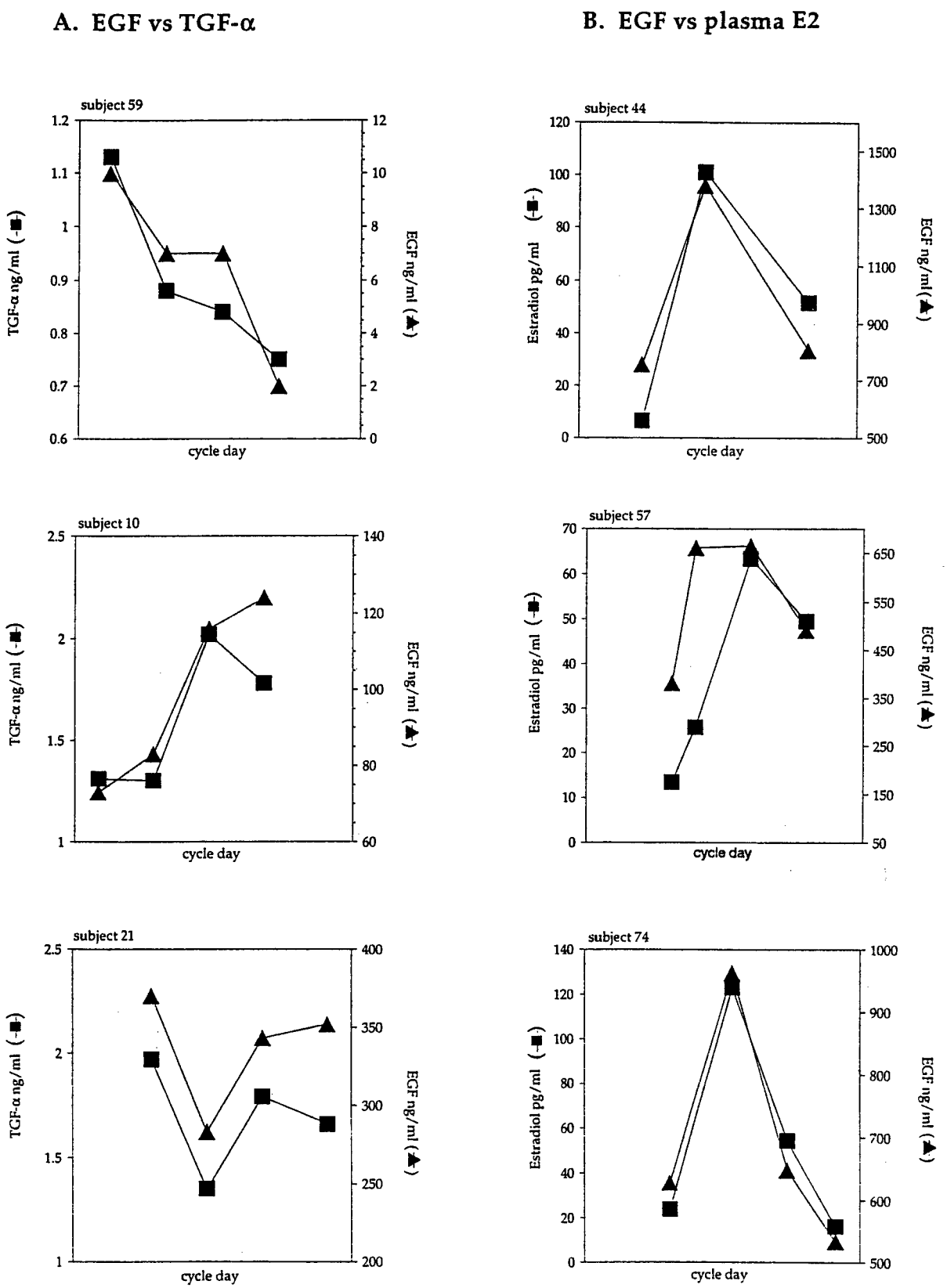


Fig. 7. Breast fluid EGF *versus* TGF-α (*A*) and breast fluid EGF *versus* plasma estradiol (*B*) across time for selected individuals.

tween- *versus* within-woman variation was greater for EGF than TGF-α, largely because of a much greater difference between women for EGF. However, TGF-α variation between the right and left breasts was very low (r = 0.89, CV = 0.11), so that the between- *versus* within-woman variation was greater for TGF-α than for EGF, although the differences between women were once again greater for EGF. Taken together, these findings imply that breast fluid levels of TGF-α, perhaps the more potent of the two GFs, are regulated within more narrow ranges than EGF in healthy women. More data are needed on this question.

The lack of a correlation between EGF and TGF-α or plasma estradiol concentrations when samples from many women are considered and the presence of significant correlations within individual women can be puzzling at first. However, this type of result arises if the quantitative relationship of EGF to TGF-α and estradiol varies from one woman to another, whereas the levels of breast fluid EGF within an individual woman tend to change over time in proportion to changes in TGF-α and plasma estradiol. Correlation analyses composed of samples from multiple women, such as those shown in Figs. 3–6, can fail to reveal these relationships. We conclude that these data provide preliminary evidence that breast fluid EGF and TGF-α are coregulated *in vivo* and that, in the case of EGF, regulation could involve circulating levels of estradiol. Other compounds, including other steroid hormones related to estradiol, could be involved in regulating TGF-α levels, or alternatively, estradiol and TGF-α levels could be related but less well synchronized than estradiol and EGF.

One limitation of these studies is that we were able to obtain breast fluid from only 60% of the women on whom nipple aspiration was attempted. Factors related to success in obtaining breast fluid have been studied extensively and appear to include age, parity, lactation history, Asian ethnicity, and cerumen type (21). In this study, such factors did not strongly differentiate those who provided fluid and those who did not; however, the study population was small and relatively homogeneous. Although we cannot rule out the possibility that our findings would not apply to the nonsecretors, had breast fluid been available from them, we find that argument to be implausible. The distinctions between secretors and nonsecretors are more likely to involve differences in the volume of breast fluid secretion and/or the physical consistency of material that normally plugs the nipple ducts, which is in part genetically determined (22). We found no evidence for a relationship between GF concentration and the amount of breast fluid obtained. Another limitation is that we have not yet confirmed the precise immunoreactive species in each RIA. The kit manufacturer's testing indicates only minimal cross-reactivity with other peptides for the antibodies used in the EGF and TGF-α assays; however, these cross-reactivities were determined for selected peptides that might or might not be present in breast fluid. We are currently conducting Western blot analyses to identify the immunoreactive species by molecular weight and ionic charge.

From these results, it appears that breast fluid EGF and TGF-α could eventually serve as useful biomarkers in studies of breast cancer etiology. Our next series of studies will examine whether GF concentrations are related to hyperproliferative states in the normal breast, to known or suspected breast cancer risk factors, and to the occurrence of breast cancer itself. If altered local GF secretion can be established as a link in the causal pathway of breast cancer development, breast fluid GF levels could provide novel intermediate end points for the evaluation of suspected risk factors or of interventions designed to reduce breast cancer risk.

Acknowledgments

The authors gratefully acknowledge Paulette O'Donnel, Yuppadee Sukosol, and the staff of the Clinical Research Center at Northwestern Memorial Hospital for assistance in collecting samples; Lynette May and Yu-Cai Lu for laboratory assistance; Susan Gapstur and Susan Colilla for study planning and coordination; Cheng-Fang Huang for statistical programming; and the women volunteers in the Repeat Sample Study for their patience and commitment.

References

1. Dickson, R. B., and Lippman, M. E. Growth factors in breast cancer. Endocr. Rev., *16:* 559–589, 1995.
2. Lippman, M. E., Dickson, R. B., Bates, S., Knabbe, C., Huff, K., Swain, S., McManaway, M., Bronzert, D., Kasid, A., and Gelmann, E. P. Autocrine and paracrine growth regulation of human breast cancer. Breast Cancer Res. Treat., *7:* 59–70, 1986.
3. Paik, S., Hazan, R., Fisher, E. R., Sass, R. E., Fisher, B., Redmond, C., Schlessinger, J., Lippman, M. E., and King, C. R. Pathologic findings from the National Surgical Adjuvant Breast and Bowel Project: prognostic significance of erbB-2 protein overexpression in primary breast cancer. J. Clin. Oncol., *8:* 103–112, 1990.
4. Fox, S. B., Smith, K., Hollyer, J., Greenal, M., Hastrich, D., and Harris, A. L. The epidermal growth factor receptor as a prognostic marker: results of 370 patients and a review of 3009 patients. Breast Cancer Res. Treat., *29:* 41–50, 1994.
5. Clarke, R., Dickson, R. B., and Lippman, M. E. Hormonal aspects of breast cancer: growth factors, drugs and stromal interactions. Crit. Rev. Oncol. Hematol., *12:* 1–23, 1992.
6. Dickson, R. B., Gottardis, M. M., and Merlino, G. T. Molecular insights into breast cancer from transgenic mouse models. Bioessays, *13:* 591–596, 1991.
7. Medina, D., Kittrell, F. S., Oborn, C. J., and Schwartz, M. Growth factor dependency and gene expression in preneoplastic mouse mammary epithelial cells. Cancer Res., *53:* 668–674, 1993.
8. Petrakis, N. L. Nipple aspirate fluid in epidemiologic studies of breast disease. Epidemiol. Rev., *15:* 188–195, 1993.
9. Rose, D. P. Hormones and growth factors in nipple aspirates from normal women and benign breast disease patients. Cancer Detect. Prev., *16:* 43–51, 1992.
10. Connolly, J. M., and Rose, D. P. Epidermal growth factor-like proteins in breast fluid and human milk. Life Sci., *42:* 1751–1756, 1988.
11. Carpenter, G. Epidermal growth factor is a major growth-promoting agent in human milk. Science (Washington DC), *210:* 198–199, 1980.
12. Ness, J. C., Sedghinasab, M., Moe, R. E., and Tapper, D. Identification of multiple proliferative growth factors in breast cyst fluid. Am. J. Surg., *166:* 237–243, 1993.
13. Valenti, G., Zanardi, S., Cerruti, G., Fassio, T., Bruzzi, P., DeFranchis, V., Barreca, A., Del Monte, P., and Minuto, F. Epidermal growth factor in breast cyst fluid: relationship with intracystic cation and androgen conjugate content. Cancer Res., *48:* 5860–5863, 1988.
14. Wilcosky, T. C. Criteria for selecting and evaluating markers. *In:* B. S. Hulka, T. C. Wilcosky, and J. D. Griffith (eds.), Biochemical Markers in Epidemiology, pp. 28–55. New York: Oxford University Press, 1990.
15. Kesaree, N., Banapurmath, C. R., Nanapurmath, S., and Shamanur, K. Treatment of inverted nipples using a disposable syringe. J. Hum. Lact., *9:* 27–29, 1993.
16. Bonfrer, J. M. G., Bruning, P. F., and Nooljen, W. J. A simple method for the measurement of the steroid fraction bound to sex hormone binding globulin. J. Steroid Biochem., *33:* 227–231, 1989.
17. Chatterton, R. T., Haan, J. N., Jenco, J. M., and Cheesman, K. L. Radioimmunoassay of pregnanediol concentrations in early morning urine specimens for assessment of luteal function in women. Fertil. Steril., *37:* 361–366, 1982.
18. Fleiss, J. L. Reliability of measurement. *In:* The Design and Analysis of Clinical Experiments, pp. 1–32. New York: John Wiley & Sons, 1986.
19. Good, P. Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses, pp. 96–97. New York: Springer-Verlag New York, Inc., 1994.
20. Diggle, P. J., Liang, K., and Zeger, S. L. *In:* Analysis of Longitudinal Data, p. 169. Oxford: Oxford University Press, 1994.
21. Wrensch, M. R., Petrakis, N. L., Gruenke, L. D., Ernster, V. L., Miike, R., King, E. B., and Hauck, W. W. Factors associated with obtaining nipple aspirate fluid: analysis of 1428 women and literature review. Breast Cancer Res. Treat., *15:* 39–51, 1990.
22. Petrakis, N. L. Cerumen phenotype and epithelial dysplasia in nipple aspirates of breast fluid. Am. J. Phys. Anthropol., *62:* 115–118, 1983.

The Prostate 32:234–240 (1997)

Epidermal Growth Factor-Related Peptides in Human Prostatic Fluid: Sources of Variability in Assay Results

Peter H. Gann,[1]* Robert Chatterton,[2] Kirsten Vogelsong,[2] John T. Grayhack,[3] and Chung Lee[3]

[1]*Department of Preventive Medicine, Robert H. Lurie Cancer Center, Northwestern University Medical School, Chicago, Illinois*

[2]*Department of Obstetrics and Gynecology, Robert H. Lurie Cancer Center, Northwestern University Medical School, Chicago, Illinois*

[3]*Department of Urology, Robert H. Lurie Cancer Center, Northwestern University Medical School, Chicago, Illinois*

BACKGROUND. Prostatic fluid (PF) provides a unique medium for noninvasive evaluation of critical growth and differentiation signals in the prostatic microenvironment. The purpose of this study was to establish the feasibility of measuring two prostatic mitogens, epidermal growth factor (EGF) and transforming growth factor-alpha (TGF-α) in PF, and specifically to quantify extraneous variability attributable to the assay itself, sample handling, or biological variation within an individual over time.

METHODS. PF was collected by transrectal massage from consecutive patients attending a urology clinic. Pooled PF and individual samples from 25 men with stable benign prostatic hyperplasia (BPH) were analyzed for EGF and TGF-α by radioimmunoassay and for total protein.

RESULTS. Reproducibility was adequate at dilutions as low as 1:50 (2-μl pooled sample) and 1:5 (20 μl) for EGF and TGF-α, respectively. Results were not affected by freeze-thaw cycles, time in storage, or protease inhibition in fresh PF. EGF and TGF-α were detectable in 100% and 92% of individual men, with respective means of 152 and 0.2 ng/ml. Correlations between two samples obtained from the same man within 12 months were highly significant (EGF $r = 0.89$, TGF-α $r = 0.71$). Protein concentrations were consistent over time; expression of either peptide per weight of protein rather than per volume did not improve within-man correlation. Between-man variability far exceeded within-man variability for both peptides, and was estimated to account for 84% and 61% of the total variability in EGF and TGF-α, respectively. There was no correlation between EGF and TGF-α in the same samples.

CONCLUSIONS. We conclude that men with BPH secrete consistent and distinct levels of EGF-related peptides in PF, and that these levels can be detected with acceptable sensitivity

Contract Grant sponsor: National Cancer Institute; Contract Grant number: KO7 CA66185; Contract Grant sponsor: U.S. Army Medical Research and Development Command; Contract Grant number: DAMD17-94-J-4203; Contract Grant sponsor: National Institute of Diabetes and Digestive and Kidney Diseases; Contract Grant number: DK 39250.

*Correspondence to: Peter Gann, M.D., Sc.D., Department of Preventive Medicine, Northwestern University Medical School, 680 N. Lake Shore Drive, Suite 1102, Chicago, IL 60611. E-mail: pgann@nwu.edu

Received 5 April 1996; Accepted 12 September 1996

and precision by radioimmunoassay (RIA). Measurement of TGF-α, which has not been reported previously, requires a relatively larger sample. *Prostate 32:234–240, 1997.*

KEY WORDS: **prostate; epidermal growth factor-urogastrone; transforming growth factor-alpha**

INTRODUCTION

A substantial body of evidence suggests that epidermal growth factor (EGF) and transforming growth factor-alpha (TGF-α) play a role in controlling the replication of prostatic epithelial cells. These peptides, which both interact with the EGF receptor, are potent mitogenic stimuli for human prostate cells in vitro [1], are overexpressed in cancerous compared to benign prostate [2], and are potentially important in the local mediation of androgen effects in the prostate [3]. Development of a noninvasive tool for assessing growth factor levels would be of considerable benefit to clinical or epidemiological research aimed at identifying exposures that enhance or inhibit prostate carcinogenesis. However, since these growth factors act primarily through autocrine or paracrine processes, systemic levels (as can be measured in serum or urine) are not likely to serve as useful biomarkers. Prostatic fluid (PF) produced by prostatic epithelium provides a reflection of the metabolic status of the prostate, and can be obtained repeatedly from most men by transrectal massage [4].

Although EGF-related peptides have previously been identified in PF, no systematic study of assay variability has been conducted. In this study, we used pooled and individual samples of PF to evaluate the assay sensitivity and degree of variability in immunoreactive EGF and TGF-α measurements attributable to sample handling, variation in total protein, biological variation over time, and the assay procedure itself. Assessment of these sources of variation is critical before we go on to examine the role played by alterations in levels of EGF-related peptides in the development and progression of prostate pathology.

MATERIALS AND METHODS

Prostatic Fluid Samples

Prostatic fluid samples collected from consecutive patients seen in our Urology Clinic were examined microscopically for cellular elements, sperm, and seminal vesicle globules. Uncontaminated PF samples were immediately placed in a refrigerator freezer and transported on ice to a −20°C freezer within 4 hr. The mean sample volume was approximately 75 μl. Samples were entered into a database that includes information on the principal diagnoses at time of collection. For analyses of assay sensitivity and precision, we made several pools of PF, each containing fluid from at least 10 men. For analyses relating to variation over time, we selected samples from 25 men in the sample bank who met the following criteria: 1) no history of prostate cancer or prostatic-specific antigen (PSA) greater than 6 ng/ml, 2) two PF samples of at least 30 μl each obtained on separate visits within 12 months, 3) stable benign prostatic hyperplasia (BPH) with no evidence of changes in symptoms, physical examination, PSA, or therapy during the interval between samples, and 4) no medications that would possibly influence androgen levels. For 2 men we selected a third sample obtained within the 12-month period, thus yielding 23 pairs of samples and two triplets. All but five sets of samples were obtained within 6 months of each other. At sampling, subjects' ages ranged from 54–78 years, with a mean of 68 years.

Assays for EGF and TGF-α

Immunoreactive EGF and TGF-α were measured in prostatic fluid by competitive binding radioimmunoassay (RIA) using commercially available reagents (Biomedical Technologies, Stoughton, MA). Hereafter, the terms "EGF" and "TGF-α" refer to their immunoreactive identities. Prostatic fluid was diluted in a saline-Tris-BSA buffer as provided in the kits for EGF and TGF-α assays. Growth factors were measured in a double-antibody RIA with ^{125}I-labeled ligands. For each assay run, we included a wide range of purified growth factor concentrations in order to construct a standard curve. The range of standard concentrations was selected to provide at least one standard level below the lowest measurable sample. Occasional samples with concentrations above the highest standard were diluted to bring them within the standard curve range. Total protein assays were done on 18 sets of samples using the standard Coomassie blue method of Bradford, which requires only 4 μl of PF.

To determine assay sensitivity, we made progressive dilutions of pooled PF (from 1:2–1:200), assayed identical aliquots at each dilution level, and calculated the within-assay coefficient of variation (CV) at each dilution. The lowest dilution yielding a CV of less than 15% was used as one indication of assay sensitivity, and was used as the standard dilution for assaying individual samples. Assay precision was determined

by calculating the mean intraassay CV for replicate samples from several assay batches. Interassay CV was determined by comparing results for identical samples of pooled PF inserted into each assay batch. These quality-control (QC) pools were also used to monitor for interassay drift, indicating possible degradation of samples in storage.

Effects of Freeze-Thaw Cycles and Protease Activity

To evaluate the effect of freeze-thaw cycles on measured growth factor levels in PF, we thawed and refroze pooled samples 1, 2, 4, and 6 times and then assayed them together. To assess the possible effect of proteases in PF on EGF and TGF-α, we collected fresh PF from 3 men, and immediately divided the PF into a regular tube and a tube containing 100 μl of glycine-HCl buffer (0.21% glycine in 0.13 M HCl) to obtain a sample pH of 2.0. This level of acidification is known to inhibit nearly all proteases [5]. Acidified samples were neutralized, and both these and nonacidified aliquots were then assayed together for EGF and TGF-α, as well as PSA.

Data Analysis

Intra- vs. interindividual variability was assessed primarily by calculation of intraclass correlation coefficients (ICC) based on pairs of samples from 25 men. For the 2 men with triplet samples, we used the two samples collected closest together in time, unless one of these samples had a missing protein value. The ICC is the proportion of total variance (including between- and within-subject components) contributed by between-subject variability. We calculated the exact lower bound of the 95% confidence interval for each ICC using the method described by Fleiss [6]. We also calculated CVs within and between men and F statistics from a one-way analysis of variance (ANOVA) comparing variance within and between men. We plotted EGF and TGF-α measurements at two timepoints and calculated both Pearson and Spearman correlation coefficients. The two types of coefficient were virtually identical; we chose to report the nonparametric Spearman coefficients. For each growth factor, calculations were performed with concentrations expressed per unit volume and per weight of total protein. We used a scatterplot and correlation analysis to compare EGF and TGF-α for the same sample (same date).

RESULTS

Data shown in Table I indicate that TGF-α in pooled prostatic fluid could be reliably measured with RIA at a dilution of 1:5, requiring 20 μl of fluid. This dilution corresponded to a concentration of approximately 0.04 ng/ml. Purified TGF-α standards were precisely measured with linear results across a range from 0.015–2.5 ng/ml. EGF could be reliably measured with a 1:50 dilution, which required 2 μl of sample and corresponded to a concentration of approximately 3 ng/ml. The assay provided precise and linear results for purified EGF standards at concentrations between 0.25–50 ng/ml. We did not observe a trend towards lower values for either growth factor with progressive freeze-thaw cycles, or a downward trend in values for QC pools assayed up to 11 months apart. Growth factor and PSA levels in fresh PF, acidified immediately after collection, were similar to those in the unacidified aliquots.

TABLE I. Intraassay and Interassay Coefficients of Variation (CVs) at Various Dilutions of Pooled Prostatic Fluid: TGF-α and EGF

	Dilution	Concentration (ng/ml)	Mean intraassay CV (%)	Mean interassay CV (%)
TGF-α				
	Undiluted	0.21		5.8%
	1:5	0.04	10.3%	16.1%
	1:10	0.02	36.9%	
EGF				
	1:25	5.82	3.6%	
	1:50	3.03	2.7%	7.3%
	1:100	1.54	22.6%	29.8%

Table II shows the comparison of variability within men over time vs. variability between men. TGF-α was not detectable in any sample from 2 men. Three other men had one sample that was considered nondetectable, yielding a total of 20 TGF-α pairs for analysis. Fourteen pairs of samples had both detectable levels of TGF-α and protein levels available. Results are shown both with and without adjustment for total protein concentration. Total protein averaged 15 mg/ml, and was highly correlated when measured at separate time points ($r = 0.83$). EGF was easily detectable in every sample assayed, and at a higher concentration than TGF-α in each sample. Between-man variability was far greater than within-man variability for both growth factors, by several measures. P values for the F statistics, which test the hypothesis that samples from the same man represent less variability than samples drawn from the entire data set, were all extremely low. The intraclass correlation coefficients indicated that most of the total variance (e.g., 84% for EGF and 61% for TGF-α) was contributed by variability between men. Expressing results relative to total protein

TABLE II. Variability Within Individuals Over Time Compared to Variability Between Individuals: Prostatic Fluid TGF-α and EGF

Reproducibility measure	TGF-α	TGF-α/protein	EGF	EGF/protein
No. of sample pairs	20	14	25	18
Mean level	0.23 ng/ml	0.017 ng/mg	152.02 ng/ml	10.56 ng/mg
SD	0.11	0.014	83.35	4.84
Standard error of measurement	0.07	0.005	33.88	2.27
$CV_{between}$	0.65	1.12	0.78	0.65
CV_{within}	0.31	0.30	0.22	0.22
CV_b/CV_w	2.10	3.73	3.55	2.95
F statistic	4.41	13.97	12.11	9.07
F statistic *P* value	0.0009	>0.0001	>0.0001	>0.0001
Intraclass correlation coefficient (ICC)	0.61[a]	0.86	0.84	0.80
ICC 95% confidence interval, lower bound	0.35	0.70	0.72	0.61

[a]ICC = 0.82 (95% CI lower bound = 0.63) when restricted to same pairs used in TGF-α/protein analysis.

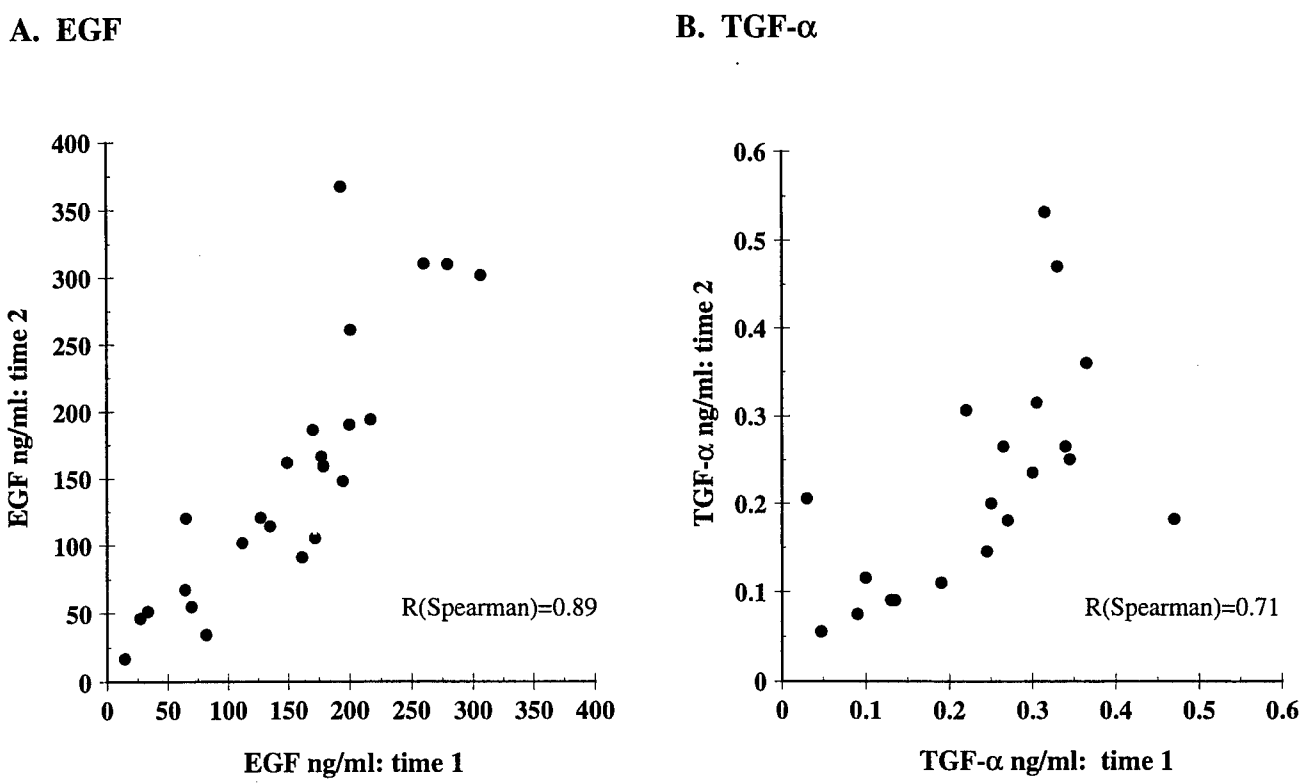


Fig. 1. Correlation of **(A)** EGF (n = 25) and **(B)** TGF-α (n = 20) concentrations in prostatic fluid obtained at two time-points from the same individual.

rather than volume had no discernible effect on the ICC for EGF. The increase in ICC for TGF-α expressed per weight protein occurred because, by chance, several sample pairs that had a missing protein value were less highly correlated. The ICC for TGF-α alone, using the same pairs as in the TGF-α/protein analysis, was 0.82.

Figure 1 shows scatterplots for EGF (Fig. 1A) and TGF-α (Fig. 1B) measured from the same individual at two time points. Figure 2 shows similar plots for EGF/

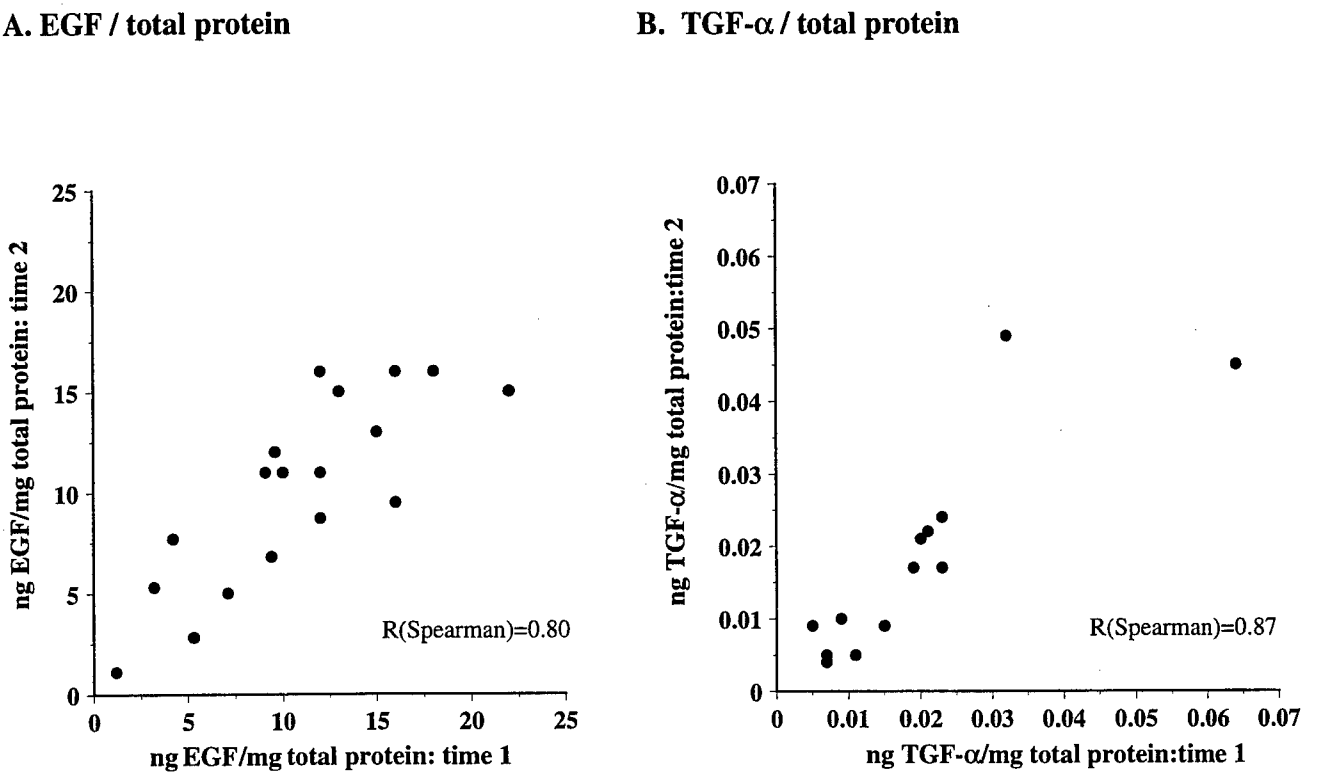


Fig. 2. Correlation of **(A)** EGF (n = 18) and **(B)** TGF-α (n = 14) in prostatic fluid obtained at two time-points from the same individual, expressed per weight of total protein.

total protein (Fig. 1A) and TGF-α/total protein (Fig. 1B). For EGF expressed per unit volume, the nonparametric correlation coefficient was 0.89, and 0.80 for EGF expressed per weight of protein. For TGF-α, the correlation coefficient was 0.71 per unit volume, and 0.87 per weight of protein. The TGF-α per unit volume correlation, with analysis restricted to sample pairs with protein levels available, was 0.86. The growth factor concentrations detected ranged widely, from 14.5–367 ng/ml for EGF and from 0.04–0.47 ng/ml for TGF-α.

Figure 3 shows the scatterplot for EGF vs. TGF-α measured in the same samples from 23 men. These results were not correlated ($r = -0.16$).

DISCUSSION

Peptide growth factors such as EGF and TGF-α appear to be potent signalling molecules for regulating the growth and differentiation of prostate cells, and, in all likelihood, their abnormal expression plays a role in the carcinogenic process [7]. Abnormal expression could result from mutation of protooncogenes transcribing the growth factors themselves or their receptors. However, research to date has not identified any strong associations between such protooncogenes and prostate cancer [8]. Alternatively, since these growth factors have a function in normal growth and development and therefore must be regulatable by endogenous signals, abnormal expression could occur as a result of up- or downregulation of the normal mechanisms for controlling growth factor gene transcription. This view allows that growth factor expression could be altered diffusely in prostatic tissue during the early stages of cancer development. An imbalance of stimulatory and inhibitory signals, for example, could create a "field effect" in which hyperproliferation leads to somatic mutation and clonal selection. Our motivation for studying growth factors in prostatic fluid therefore stems more from an interest in the influence of etiologic factors in the environment (including diet) on growth factor expression than it does from an interest in detecting prostate cancer earlier due to specific patterns of growth factor expression confined to nests of neoplastic cells.

Our results indicate that EGF and TGF-α can be reliably measured in prostatic fluid by radioimmunoassay. TGF-α- and EGF-like material were detectable in nearly all samples, with levels of EGF 700–800 times higher on average than TGF-α. Required sample volumes are small enough to permit analysis of both growth factors in most individual specimens. However, TGF-α will be difficult or impossible to measure in some samples with low concentrations and low sample volume. Men with low sample volumes were excluded from this study.

Levels of TGF-α and EGF in individual men re-

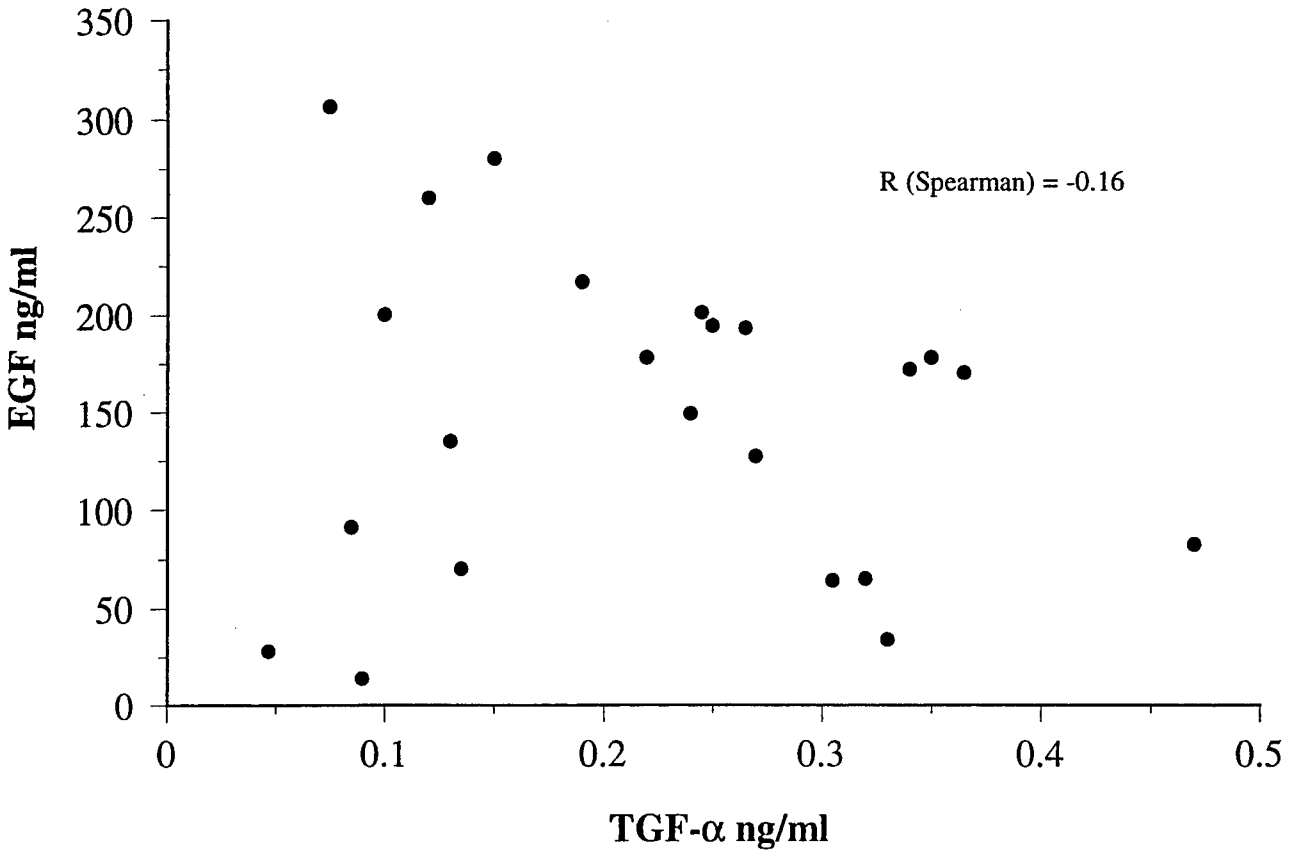


Fig. 3. Correlation of EGF and TGF-α concentrations in the same prostatic fluid sample (n = 23).

mained fairly constant when sampled twice within 12 months. Greater variability between vs. within men suggests that individuals can be correctly ranked within study populations, which is vital for epidemiologic analyses. The within-man variability for TGF-α appeared to be reduced by computing the ratio of growth factor to total protein rather than sample volume. However, most of the apparent improvement in correlation, e.g., in comparing Figures 1B and 2B, was due to chance dropout of poorly-correlated observations, because both protein and TGF-α measurements were not available for some samples. The correlation coefficient for TGF-α per unit volume for the same 14 pairs of samples plotted in Figure 2B was 0.86. Measurement of total protein is easily performed with as little as 4 μl of sample; however, it does not appear to be helpful.

Levels of TGF-α and EGF in prostatic fluid are not correlated. We previously found them to be highly correlated ($r = 0.88$) in breast fluid. The reasons for this difference need to be explored, because they suggest that in PF, expression of these growth factors might have different regulatory mechanisms. We are currently determining the relationship between growth factor and steroid hormone levels in PF.

Only two previous studies provide data on growth factors in PF. Tackett et al. [9] identified a 30-kD peptide in PF that was mitogenic to cultured fibroblasts as well as a smaller, unidentified inhibitory peptide. Gregory et al. [10] measured EGF-like material by RIA in PF from men with BPH and clinically normal prostate glands. Among the "normal" men with a mean age of 67.3, the mean EGF concentration was 272 ng/ml, whereas among similarly aged men with BPH the mean EGF was 155 ng/ml, a statistically significant difference. EGF levels in PF did not appear to vary by age among the normals. This study failed to detect EGF-like material in tissue sections, which led the authors to speculate that the prostate does not produce EGF itself, but instead just concentrates and secretes EGF obtained from the blood. BPH could thus involve an impairment in the ability to package and secrete EGF. These intriguing observations on EGF, to our knowledge, have not been subjected to further study. Both EGF- and a TGF-α-like material have been identified in seminal fluid [11]. We were unable to find any data published prior to this report on TGF-α in PF.

The available data on EGF and TGF-α in human prostatic tissue are sparse and difficult to interpret. In one study, EGF was detected by immunohistochemistry in both BPH and carcinoma; TGF-α was detected in carcinoma only [2]. An earlier study reported less frequent detection of EGF in BPH tissue compared to cancer (6% vs. 68%) [12]. Yang et al. [13] measured

EGF/TGF-α in homogenized tissue by RIA and found equivalent amounts of both GFs in BPH and cancer tissue. Since the distribution of these GFs within tissue sections is not homogeneous, assays of total GF per weight of tissue might not reflect the biologically relevant concentrations. Saturation analysis of EGF binding sites on resected human prostate tissue revealed lower levels of EGF binding in samples of containing BPH compared to cancer or histologically normal tissue [14]. Considered together, these results suggest a role for the EGF family in human prostatic disease; however, we do not have enough data yet to formulate detailed hypotheses regarding mechanisms.

This is the first study to report on sources of variability in growth factor measurements in prostatic fluid and the first to report detection of TGF-α. However, we note several limitations to our data. All samples assayed were obtained from men with BPH, because these were the most numerous in the sample bank, and it is possible that assay characteristics are different in men without clinical prostatic disease or those in other age groups. Furthermore, if growth factor levels in PF are associated with the condition of the gland, then interindividual variation in a more general, less highly-selected population would be even greater than our estimates indicate. We limited patients selected to those with stable BPH and a minimal interval between samples. Failure to identify changes in growth factor levels related to disease progression would have led to conservative overestimation of within-man variability. We refer to the measured growth factors as EGF and TGF-α, but do not presume that the immunoreactive species necessarily correspond to pure EGF and TGF-α. Although the antibodies display minimal crossreactivity in other biological media, crossreactivity in PF has not been specifically tested. Furthermore, it is possible that the detectable growth factors include higher molecular weight forms. Earlier investigators identified a 6-kD form consistent with pure EGF in addition to a higher molecular species in PF [10], and only a 6-kD form of TGF-α similar to pure TGF-α in seminal fluid [11]. Our preliminary analyses indicate that the immunoreactive EGF we are measuring includes some higher molecular weight forms. The biological activity of the forms in PF is also yet to be determined.

We are currently pursuing more detailed characterization of the EGF-related species present in PF and are investigating assays for additional growth factors. In the meantime, we conclude that these data support the feasibility of using assays for EGF-like peptides in prostatic fluid as biomarkers in clinical or epidemiological research. We plan to further explore the associations between EGF-like peptides in prostatic fluid and prostate cancer, as well as the identification of factors influencing growth factor levels, such as local steroid concentrations.

ACKNOWLEDGMENTS

The authors thank the following individuals for their contributions to this work: Eva Braun, Nicole Wilson, Barbara Stuppy, and Susan Collila for assistance in collecting and cataloging samples, and Allison Ellman for assistance with data analysis and preparation of the manuscript.

This work was supported by a Preventive Oncology Academic Award (KO7 CA66185) from the National Cancer Institute to P.G., a grant (DAMD17-94-J-4203) from the U.S. Army Medical Research and Development Command to P.G., and a grant from the National Institute of Diabetes and Digestive and Kidney Diseases (DK 39250) to J.T.G.

REFERENCES

1. Sherwood ER, Lee C: Epidermal growth factor-related peptides and the epidermal growth factor receptor in normal and malignant prostate. World J Urol 1995;13:290–296.
2. Myers RB, Kudlow JE, Grizzle WE: Expression of transforming growth factor-alpha, epidermal growth factor and the epidermal growth factor receptor in adenocarcinoma of the prostate and benign prostatic hyperplasia. Mod Pathol 1993;6:733–737.
3. Liu XH, Wiley HS, Meikle AW: Androgens regulate proliferation of human prostate cancer cells in culture by increasing transforming growth factor-alpha (TGF-alpha) and epidermal growth factor (EGF)/TGF-alpha receptor. J Clin Endocrinol Metab 1993;77:1472–1478.
4. Grayhack JT, Lee C, Oliver L, Schaeffer AJ, Wendel EF: Biochemical profiles of prostatic fluid from normal and diseased prostate glands. Prostate 1980;1:227–237.
5. Clement-Jones V, Lowry PJ, Rees LH, Besser GM: Development of a specific extracted radioimmunoassay for methionine enkephalin in human plasma and cerebrospinal fluid. J Endocrinol 1980;86:231–243.
6. Fleiss JL: "The Design and Analysis of Clinical Experiments," New York: Wiley, 1995:12.
7. Habib FK, Chisholm GD: The role of growth factors in the human prostate. Scand J Urol Nephrol [Suppl] 1991;138:53–58.
8. Peehl DM: Oncogenes in prostate cancer: An update. Cancer 1993;71:1159–1164.
9. Tackett RE, Heston WDW, Parrish RF, Pletscher LS, Fair WR: Mitogenic factors in prostatic tissue and expressed prostatic secretion. J Urol 1985;133:45–48.
10. Gregory H, Willshire IR, Kavanagh JP, Blacklock NJ, Chowdury S, Richards RC: Urogastrone—Epidermal growth factor concentrations in prostatic fluid of normal individuals and patients with benign prostatic hypertrophy. Clin Sci 1986;70:359–363.
11. Yie S, Lobb DK, Clark DA, Younglai EV: Identification of a transforming growth factor alpha-like molecule in human seminal plasma. Fertil Steril 1994;61:129–135.
12. Fowler JE Jr, Lau JLT, Ghosh L, Mills SE, Mounzer A: Epidermal growth factor and prostatic carcinoma: An immunohistochemical study. J Urol 1988;139:857–861.
13. Yang Y, Chisholm GD, Habib FK: Epidermal growth factor and transforming growth factor α concentrations in BPH and cancer of the prostate: Their relationships with tissue androgen levels. Br J Cancer 1993;67:152–155.
14. Davies P, Eaton CL: Binding of epidermal growth factor by human normal, hypertrophic, and carcinomatous prostate. Prostate 1989;14:123–132.

The Prostate 40:248–255 (1999)

Growth Factors in Expressed Prostatic Fluid From Men With Prostate Cancer, BPH, and Clinically Normal Prostates

Peter H. Gann,[1]* Karin G. Klein,[1] Robert T. Chatterton,[2] Allison E. Ellman,[1] John T. Grayhack,[3] Robert B. Nadler,[3] and Chung Lee[3]

[1]*Department of Preventive Medicine and Robert H. Lurie Cancer Center, Northwestern University Medical School, Chicago, Illinois*

[2]*Department of Obstetrics and Gynecology and Robert H. Lurie Cancer Center, Northwestern University Medical School, Chicago, Illinois*

[3]*Department of Urology and Robert H. Lurie Cancer Center, Northwestern University Medical School, Chicago, Illinois*

BACKGROUND. Although growth factors such as epidermal growth factor (EGF), transforming growth factor (TGF)-α, and TGF-β are important regulators of prostate cell growth in vitro and in animal models, evidence to support their role in human prostate cancer development remains sparse. We previously showed that men without prostate cancer have concentrations of EGF and TGF-α in expressed prostatic fluid (EPF) that are individually distinct and stable over time. This study addressed whether growth factor levels in EPF are associated with the presence or progression of prostate cancer.

METHODS. We measured levels of immunoreactive EGF, TGF-α, and TGF-β1 in stored EPF samples from three age-matched groups: 19 men with untreated, histologically diagnosed prostate cancer (CaP), 38 with benign prostate hyperplasia (BPH), and 19 with normal prostate glands (NPD).

RESULTS. Median TGF-α was lower in the BPH group (0.45 ng/ml) than in either CaP (0.63 ng/ml) or NPD (0.58 ng/ml) groups (P = 0.03 and 0.12, respectively). For EGF, the median was lowest in the CaP group and highest in the NPD group (92.5 ng/ml vs. 175.5 ng/ml, P = 0.006). For TGF-β1, the median level in CaP was 2.7 times higher than the median level among all controls (6.65 ng/ml vs. 2.46 ng/ml, P = 0.002). Growth factor levels were not associated with tumor stage or Gleason score. However, the single case with distant metastases had TGF-β1 levels 23-fold higher than the CaP median.

CONCLUSIONS. The results suggest that at the time of CaP diagnosis, EGF levels in EPF are significantly lower, and TGF-β1 levels significantly higher, than normal. Marked overexpression of TGF-β1 in advanced CaP might be reflected in extremely high EPF levels. *Prostate 40:248–255, 1999.*

KEY WORDS: **prostate; prostatic fluid; TGF-α; TGF-β; EGF**

INTRODUCTION

Proliferation and differentiation within the human prostate gland are controlled, in part, by peptide growth factors, including epidermal growth factor (EGF) and transforming growth factors alpha and beta (TGF-α, TGF-β) [1]. EGF and TGF-α, which are homologous and can both interact with the EGF receptor,

Grant sponsor: National Cancer Institute, National Institutes of Health; Grant numbers: KO7 CA66185, R21 CA69851; Grant sponsor: US Army Breast Cancer Research Program; Grant number: DAMD17-94-J-4203.

*Correspondence to: Dr. Peter Gann, Department of Preventive Medicine, Northwestern University Medical School, 680 N. Lake Shore Drive, Suite 1102, Chicago, IL 60611. E-mail: pgann@nwu.edu

Received 25 November 1998; Accepted 23 March 1999

have been associated with mitogenic effects and possible local mediation of androgen action [2]. On the other hand, TGF-β, and in particular the TGF-β1 isoform, could have a bifunctional role in prostate cancer development [3]. TGF-β1 generally inhibits epithelial cell proliferation. However, in more advanced tumors in vivo it could promote tumor growth and metastasis by stimulating angiogenesis, inhibiting local immune response, or enhancing invasiveness through changes in the stromal matrix [3]. Despite abundant evidence that these growth factors are critical mediators, we have only sparse translational data relating levels of their expression to prostate cancer development in humans. This is at least partly due to the lack of noninvasive methods for sampling and comparing affected and nonaffected populations.

Expressed prostate fluid (EPF), which can be readily obtained from most men by transrectal massage, provides a unique medium for noninvasively investigating the biochemical microenvironment of the prostate gland. It is produced by the apocrine and merocrine secretory activity of the prostate epithelium and, based on measurements of many biochemical constituents, it appears to reflect the metabolic status of the prostatic epithelium as a whole [4]. The small volume of EPF that is obtained (mean, approximately 75 μl) presents a challenge in developing assays sensitive enough to allow measurement of multiple growth factors and other compounds from an individual sample. We previously reported the development of two such assays to measure EGF and TGF-α in EPF, and found that EGF levels were 700–800 times higher on average than those of TGF-α [5]. We also demonstrated that the levels of EGF and TGF-α remain consistent over time for individual men, but vary greatly between men. Here we describe the development of an assay to measure TGF-β1 in EPF, and application of the assays for EGF, TGF-α and TGF-β1 to a set of EPF samples representing men with prostate cancer, benign prostatic hyperplasia, or no prostatic disease.

MATERIALS AND METHODS

Sample Collection and Storage

Samples of expressed prostatic fluid used in this study were identified from a sample bank maintained since 1992 by the Department of Urology at Northwestern University Medical School. The bank contains more than 2,200 specimens from approximately 700 patients collected during consecutive office visits by one of the authors (J.T.G.). The fluid was examined microscopically, immediately placed in the freezer compartment of a refrigerator, and transported on ice for permanent storage at −20°C.

Selection of Samples for Analysis

Preliminary selection of eligible samples was performed using a computerized sample inventory. Final selection was confirmed after a structured audit of each patient's medical records. We selected 19 for the prostate cancer (CaP) group based on the following criteria: 1) histologic confirmation of prostate cancer by needle biopsy, and 2) no evidence of surgical, hormonal, or radiation therapy prior to EPF sampling. For each CaP patient, we selected 2 patients for the benign prostatic hyperplasia (BPH) group who matched the CaP patient within 3 years on age and 6 months on sample storage length. Criteria for the BPH group included: 1) gland size ≥4.5 cm in width by palpation, 2) no suspicious focal findings based on digital rectal or ultrasound examinations, 3) all serum PSAs <10 ng/ml with no more than a 15% increase per year in PSA throughout the observation period available, 4) no drug or surgical therapy prior to sampling, 5) no history of CaP before or after sampling, and 6) no evidence of active prostatitis (including no more than rare white blood cells in EPF) at the time of sampling. From the sample bank, we selected 17 patients for the "no prostatic disease" (NPD) group who were frequency matched to the CaP cases on age and length of sample storage. Two additional fresh EPF samples were collected to bring the total number in the NPD group to 19. NPD patients met the following criteria: 1) gland size <4.5 cm by palpation, 2) serum PSA <4 ng/ml throughout the observation period, and 3) no history of symptoms or diagnosis of CaP, prostatitis, or BPH before or after sampling.

Growth Factor Assays

TGF-α and EGF were measured using radioimmunoassay kits from Biomedical Technology, Inc. (Stoughton, MA) that are designed to measure growth factors in serum or plasma. The optimization, sensitivity, and accuracy of the kits for measuring EGF and TGF-α in prostate fluid were described previously [5]. Intraassay coefficients of variation (CV) for these growth factors in EPF range from 3–10%. EGF can be measured using only 2 μl, and TGF-α using 20 μl of EPF. TGF-β1 was measured using the Predicta™ human TGF-β1 kit from Genzyme, Inc. (Cambridge, MA), which is designed to measure TGF-β1 in serum or plasma. Its high sensitivity and small reaction volume presented the best opportunity to minimize the sample volume for accurately measuring EPF. Optimization of the kit protocol allowed detection of as

TABLE I. Selected Characteristics of Prostate Cancer, Benign Prostatic Hyperplasia, and No Prostatic Disease Patients With Prostatic Fluid Samples Evaluated*

	CaP	BPH	NPD
Number	19	38	19
Mean (SD) age, years	69.7 (6.95)	69.6 (7.45)	68.9 (6.43)
Mean (SD) time from sampling to assay, months	39 (2.9)	39 (2.8)	25 (3.6)
Mean (SD) serum PSA at sampling, ng/ml	na	3.3 (0.5)	1.6 (0.3)
Prostate cancer stage	A, 5 B, 9 C, 4 D, 1	na	na

*CaP, prostate cancer; BPH, benign prostatic hyperplasia; NPD, no prostatic disease; na, not available. Stage A, nonpalpable, localized; stage B, palpable, confined to prostate; stage C, extracapsular spread; stage D, metastases to pelvic nodes or distant sites.

little as 0.04 ng/ml TGF-β1 in the standard curve. Individual samples were assayed first, using a standard 5 μl volume. If TGF-β1 was below detection in that volume (n = 23), a second assay was performed with a larger sample volume, up to 25 μl. Intraassay coefficients of variation in a pool with high levels of TGF-β1 were 4% or below, at assay volumes ranging between 5–25 μl. In a pool with low TGF-β1, intraassay CVs ranged from 48% with an assay volume of 5 μl to 3% with an assay volume of 25 μl. All matched sets of samples were assayed in the same run to minimize the effect of interassay variability. Results were expressed in weight-per-volume units; previous results indicated no relationship between growth factor concentrations and total sample volume or protein concentration [5].

Data Analysis

We log-transformed growth factor concentrations to achieve more normal distributions. We present actual nontransformed values and medians. For calculating *P* values, we performed parametric *t*-tests, using the log-transformed values for all diagnostic groups. We calculated both Spearman and Pearson correlation coefficients on log-transformed values for determining the extent of correlation between growth factor concentrations. All *P* values reported are two-sided.

RESULTS

Selected characteristics of the study population are reported in Table I. Mean age in the three groups was not significantly different. Mean time since sampling was shorter in the NPD group due to the inclusion of two recently collected samples in the analysis. Samples in the CaP group included men with histologically identified tumors from a range of stages.

Figure 1 shows the EGF concentrations for each diagnostic group. The median value was lowest in the CaP group (92.5 ng/ml), and progressively higher in the BPH (131.0 ng/ml) and NPD (175.5 ng/ml) groups. The difference between the CaP group and the BPH or NPD groups was unlikely to have occurred by chance: $P = 0.02$ for CaP vs. BPH, and $P = 0.002$ for CaP vs. NPD. The difference between the BPH and NPD groups was not statistically significant ($P = 0.17$). We found little evidence for an association between EGF level and prostate cancer stage ($P = 0.21$ for stage A/B vs. C/D). Although the two highest EGF levels in the CaP group were found in men with extraprostatic disease, three other extraprostatic cases had levels near or below the median.

Due to the higher volume of EPF required to measure TGF-α accurately, only a subset of the samples with sufficient volume could be measured in the assay, including 11 CaP, 26 BPH, and 9 NPD samples. Figure 2 shows that all three groups contained some outlying values. Median TGF-α was lower in the BPH group (0.45 ng/ml) than in either the CaP (0.63 ng/ml) or NPD (0.58 ng/ml) groups ($P = 0.03$ and 0.12, respectively). The two extraprostatic CaP cases evaluated had TGF-α levels close to the median.

TGF-β1 was measured in a subset that included 17 CaP, 27 BPH, and 15 NPD samples (Fig. 3). The median level of TGF-β1 in the CaP group was 2.7 times higher than the median level among all other samples (6.65 ng/ml vs. 2.46 ng/ml, $P = 0.002$). The medians for the BPH and NPD groups were not significantly different. Only two of the CaP samples had levels equal to or less than the median for all controls; 59% of the CaP group had TGF-β1 levels above 4 ng/ml, compared to 33% and 13% in the BPH and NPD groups,

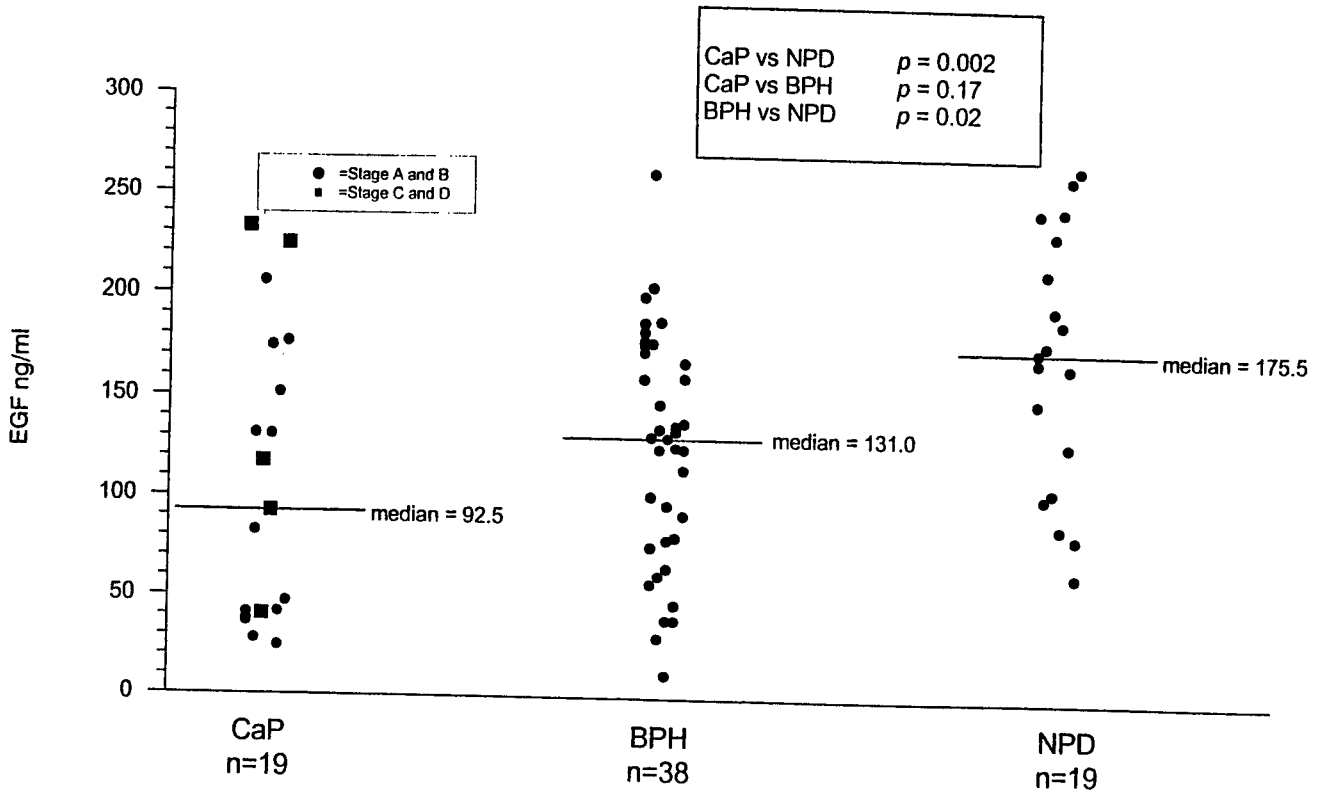


Fig. 1. EGF concentrations in prostatic fluid from prostate cancer (CaP), benign prostatic hyperplasia (BPH) and no prostatic disease (NPD) patients. The distributions of EGF concentrations for patients in each diagnostic group and the median are displayed, with the number of patients per group shown beneath the group designation. In the CaP group, results for patients with stage A or B tumors are represented by circles, and those for patients with stage C or D tumors are represented by squares. *P* values are two-sided and were calculated by *t*-tests on log-transformed values.

respectively. The sole patient with known metastatic prostate cancer at the time of EPF sampling had TGF-β1 levels over 20 times above the median. In the BPH and NPD groups, two samples had unusually elevated TGF-β1 levels. The clinical history of these patients indicates that they have not developed prostate cancer up to the present.

EGF and TGF-α levels were weakly correlated among all EPF samples ($r = 0.29$, $P = 0.05$), and more strongly correlated in the BPH group alone ($r = 0.52$, $P = 0.007$). EGF and TGF-β1 levels showed a weak inverse correlation ($r = -0.27$, $P = 0.04$), which was again most prominent in the BPH group ($r = -0.34$, $P = 0.08$). We observed no significant correlations between TGF-α and TGF-β1 levels overall or in any individual diagnostic group.

DISCUSSION

Our results, which are among the first data available on growth factor concentrations in prostatic fluid, suggest that there are important associations between these concentrations and the presence of neoplastic or hyperplastic disease in the prostate. Concentrations of EGF were 50% lower in EPF samples from men with prostate cancer than in samples from age-matched men with normal prostates. EGF levels were approximately 25% lower in men with BPH compared to the clinically normal group. Results for TGF-α were less clear; however, men with BPH had lower concentrations of EPF than those with either prostate cancer or normal prostates. TGF-β1 concentrations were more than 2.5 times higher among men with prostate cancer than among those with BPH or normal prostates.

EGF and TGF-α are strongly mitogenic for prostatic epithelial cells in vitro; therefore, we hypothesized that levels would be higher in EPF in conjunction with hyperproliferative states such as cancer and BPH [2]. Higher levels in these conditions could result from a field effect due to a generalized increase in cell proliferative activity, or to release of growth factors into the ductal system from focal lesions with hyperactive autocrine/paracrine loops. A substantial body of evidence, including findings with the TGF-α transgenic mouse, indicates that upregulation of EGF-related peptides is involved in the development of preneoplastic conditions in several types of tissue [6,7]. The possible role of overexpression of EGF-related pep-

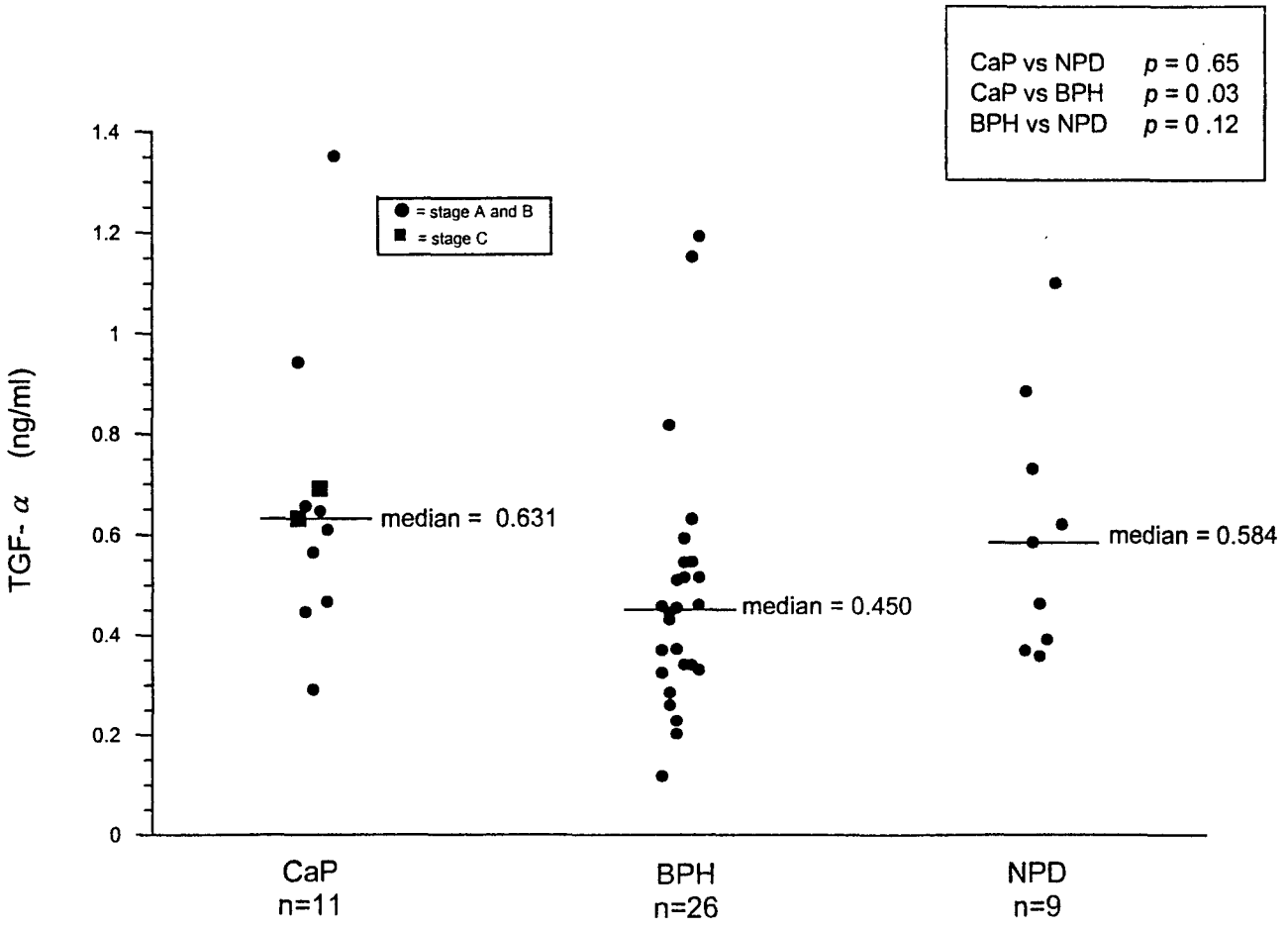


Fig. 2. TGF-α concentrations in prostatic fluid from prostate cancer (CaP), benign prostatic hyperplasia (BPH), and no prostatic disease (NPD) patients. The distributions of TGF-α concentrations for patients in each diagnostic group and the median are displayed, with the number of patients per group shown beneath the group designation. *P* values are two-sided and were calculated by *t*-tests on log-transformed values.

tides in the early stages of prostate carcinogenesis is suggested by the finding of increased TGF-α and EGF receptor in dysplastic foci in the dorsolateral prostates of steroid-treated rats [8]. In the normal prostate, EGF and TGF-α appear to be apocrine secretory products, with most immunostained material accumulating in ductal spaces [8,9]. With dysplasia and further progression to carcinoma, the normal cell polarity is lost and immunostaining patterns are altered so that intracytoplasmic localization of EGF and TGF-α is increased. Moreover, in steroid-treated rats, although TGF-α expression is increased in dysplasic foci and even more so in tumors, EGF is essentially absent in tumor foci. Taken together, these changes in growth factor expression and cellular localization might explain our results for EGF in expressed prostatic fluid, i.e., a marked depression in EGF among men with diagnosed prostate cancer and intermediate levels among men with BPH. The comparisons for TGF-α, which was expressed at far lower levels than EGF, were more ambiguous and could reflect opposing effects during tumorigenesis on expression, cellular localization, and secretory processes. Alternatively, reduced levels of EGF and TGF-α in prostatic fluid could be explained by binding of these ligands to excess EGF receptor released by rapidly dividing cells [10].

We found only one previous study that compared growth factor levels in EPF across diagnostic groups. In that study, which reported on EGF only, concentrations of EGF in prostatic fluid were lower among 20 men with BPH (mean, 155 ng/ml) than among 10 similar-aged men with clinically normal prostates (mean, 272 ng/ml) [11]. These results are quite similar to ours, and also suggest that BPH might involve disruption of the normal secretory process for EGF or loss of EGF required for maintenance of normal growth. Differences in growth factor expression patterns could exist between BPH characterized by patterns of either epithelial or stromal hyperplasia. We were unable to characterize BPH this way in our sample of patients.

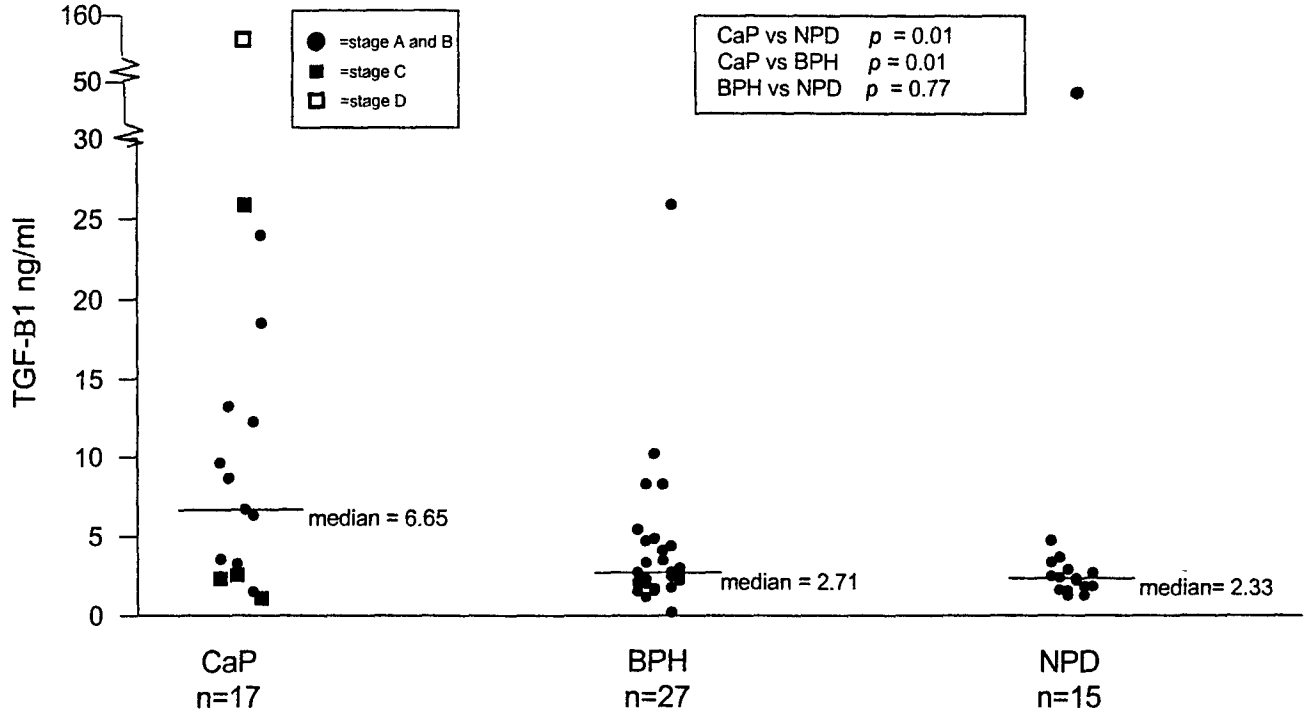


Fig. 3. TGF-β1 concentrations in prostatic fluid from prostate cancer (CaP), benign prostatic hyperplasia (BPH), and no prostatic disease (NPD) patients. The distributions of TGF-β1 concentrations for patients in each diagnostic group and the median are displayed, with the number of patients per group shown beneath the group designation. In the CaP group, results for patients with stage A or B tumors are represented by circles, those for patients with stage C by solid squares, and those for stage D tumors by an open square. *P* values are two-sided and were calculated by *t*-tests on log-transformed values.

In this sample of BPH patients we found a weak positive correlation between EGF and TGFα levels in prostatic fluid, in contrast to earlier results [5].

Members of the TGF-β family can have divergent effects in carcinogenesis, depending on the stage of tumor development [3,12]. TGF-β1, for example, generally inhibits the proliferation of epithelial cells, including many types of cultured prostatic epithelial cells. However, recent evidence suggests that prostate cancers lose responsiveness to the antiproliferative effects of TGF-β, thus allowing other effects to predominate. These effects include increases in proteolysis, cell motility, and angiogenesis, and suppression of local immune responses, all of which favor invasion and/or metastasis. Increased expression of TGF-β1 has been seen in prostate cancer tissue compared to BPH, and therefore we hypothesized that TGF-β1 would be elevated in EPF from men with established prostate cancer. Staining for TGF-β1 is particularly higher in tumors with metastatic disease [13]. One group reported that marked elevations in plasma levels of TGF-β1 can be detected in association with extraprostatic disease [14]. Two subsequent studies failed to confirm this, although one reported elevated TGF-β1 in urine from men with CaP, and a positive association with disease stage [15,16]. It is noteworthy that, in our study, the single patient with known metastatic prostate cancer at the time of EPF sampling had TGF-β1 levels over 20 times greater than the median for all prostate cancers. Results from a single patient must be interpreted cautiously, and furthermore, we did not observe elevated levels in stage C tumors. Nevertheless, the possibility that TGF-β1 levels (in either plasma or EPF) could serve as a clinically useful marker for occult metastatic disease deserves further attention.

This is the first study, to our knowledge, that reports on TGF-β concentrations in EPF. However, previous studies did report the presence of TGF-β or similar peptides in seminal fluid [17,18]. One study reported a mean concentration of TGF-β1 in seminal fluid of 238 ng/ml, considerably higher than the concentration we found in EPF [18]. We are continuing to refine the TGF-β1 assay and are characterizing intra- vs. interindividual variation as we have for EGF and TGF-α. The roles of TGF-β isoforms 2 and 3 in prostate cancer are less clear; however, these isoforms have different immunostaining patterns in prostatic tissue, and their concentrations in EPF could differ from those of TGF-β1 [19].

The strengths of this study include detailed criteria for selection of patients within diagnostic groups and close matching between groups on patient age and

length of sample storage. The acquisition of EPF samples from patients with CaP prior to treatment is an unusual and particularly valuable aspect of the sample bank that we used. This study was also strengthened by the experience gained by our group in optimizing immunoassays for measurement of growth factors in small volumes of biological fluids. However, the present study also has several limitations. The size of the study was restricted because samples from untreated men with prostate cancer and age-comparable controls with no prostatic disease are scarce, and some samples must be reserved for future studies. Unavoidably, staging of the CaP patients was imprecise because few of these patients were surgically staged. Therefore, it is conceivable that some CaP cases were more advanced at the time of EPF sampling. Comparison of cases by Gleason score yielded the same null results as comparisons by stage; however, it is possible that Gleason scores from biopsies were also underestimated.

Although the criteria for inclusion in the BPH group were quite strict, it was not possible to obtain biopsy evidence at the time of EPF sampling to rule out the existence of occult cancer. However, inclusion of such misclassified patients would have biased our results towards null. The TGF-β1 assay is relatively new, and we have not yet completed plans for fully characterizing both assay and biological sources of variation. However, we felt that the striking differences already observed between diagnostic groups should be reported. The most important limitation of the present study, however, lies in the analysis of EPF samples that were collected *concurrent* with determination of the pathological condition of the prostate. It is therefore impossible to determine the time relation of differences in growth factor levels to development of either BPH or cancer. Alteration of growth factor levels in EPF could be a late phenomenon, reflecting the effects of clinically detectable disease itself, or an earlier phenomenon, reflecting conditions that favor early progression of the disease process.

We are extending our efforts to evaluate additional growth factors in EPF. Because of our main interest in using EPF analyses to explore strategies for prevention, we plan to focus attention on the relation of changes in growth factor levels to early phases of carcinogenesis, and to the effects of potential dietary changes or chemopreventive agents on growth factor expression. In conclusion, these results indicate that growth factor levels in prostatic fluid, which are easily obtained noninvasively, are associated with hyperplastic or malignant disease in the prostate. Further studies will be required to determine if these markers convey important information about etiologic processes, staging, or prognosis.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Ramine Yazhari, Nicole Wilson, and Eva Braun for their contributions to data collection, and Geoffrey Bonn for his help with the laboratory assays.

REFERENCES

1. Habib FK, Chisholm GD. The role of growth factors in the human prostate. Scand J Urol Nephrol [Suppl] 1991;138: 53–58.
2. Sherwood ER, Lee C. Epidermal growth factor-related peptides and the epidermal growth factor receptor in normal and malignant prostate. World J Urol 1995;13:290–296.
3. Barrack ER. TGF-β in prostate cancer: a growth inhibitor that can enhance tumorigenicity. Prostate 1997;31:61–70.
4. Grayhack JT, Lee C, Oliver L, Schaeffer AJ, Wendel EF. Biochemical profiles of prostatic fluid from normal and diseased prostate glands. Prostate 1980;1:227–237.
5. Gann PH, Chatterton RT, Vogelsong KM, Grayhack JT, Lee C. Epidermal growth factor-related peptides in human prostatic fluid: sources of variability in assay results. Prostate 1997;32: 234–240.
6. Davies BR, Warren JR, Schmidt G, Rudland PS. Induction of a variety of preneoplasias and tumours in the mammary glands of transgenic rats. Biochem Soc Symp 1998;63:167–184.
7. Everitt JI, Walker CL, Goldsworthy TW, Wolf DC. Altered expression of transforming growth factor-alpha: an early event in renal cell carcinoma development. Mol Carcinog 1997;19:213–219.
8. Kaplan PJ, Leav I, Greenwood J, Kwan PW, Ho SM. Involvement of TGF-α and epidermal growth factor receptor in sex hormone-induced prostatic dysplasia and the growth of an androgen-independent transplantable carcinoma of the prostate. Carcinogenesis 1996;17:2571–2579.
9. Wu HH-H, Kawamata H, Kawai K, Lee C, Oyasu R. Immunohistochemical localization of epidermal growth factor and transforming growth factor α in the male rat accessory sex organs. J Urol 1993;150:990–993.
10. Glynne-Jones E, Goddard L, Harper ME. Comparative analysis of mRNA and protein expression for epidermal growth factor receptor and ligands relative to the proliferative index in human prostatic tissue. Hum Pathol 1996;27:688–694.
11. Gregory H, Willshire IR, Kavanagh JP, Blacklock NJ, Chowdury S, Richards RC. Urogastrone-epidermal growth factor concentrations in prostatic fluid of normal individuals and patients with benign prostatic hypertrophy. Clin Sci 1986;70:359–363.
12. Sehgal I, Baley PA, Thompson TC. Transforming growth factor beta1 stimulates contrasting responses in metastatic versus primary mouse prostate cancer-derived cell lines in vitro. Cancer Res 1996;56:3359—3365.
13. Eastham JA, Truong LD, Rogers E, Kattan M, Flanders KC, Scardino PT, Thompson TC. Transforming growth factor-beta1: comparative immunohistochemical localization in human primary and metastatic prostate cancer. Lab Invest 1995;73:628–635.
14. Ivanovic V, Melman A, Davis-Joseph B, Valcic M, Geliebter J. Elevated plasma levels of the TGF-β1 in patients with invasive prostate cancer. Nat Med 1995;1:282–284.
15. Wolff JM, Fandel T, Borchers H, Brehmer B, Jakse G. Transform-

ing growth factor-β1 serum concentration in patients with prostatic cancer and benign prostatic hyperplasia. Br J Urol 1998;81: 403–405.
16. Perry KT, Anthony CT, Case T, Steiner MS. Transforming growth factor beta as a clinical biomarker for prostate cancer. Urology 1997;49:151–155.
17. Lokeshwar BL, Block NL. Isolation of a prostate carcinoma cell proliferation-inhibiting factor from human seminal plasma and its similarity to transforming growth factor β1. Cancer Res 1992; 52:5821–5825.
18. Nocera M, Chu TM. Characterization of latent transforming growth factor-beta from human seminal plasma. Am J Reprod Immunol 1995;33:282–291.
19. Perry KT, Anthony CT, Steiner MS. Immunohistochemical localization of TGFβ1, TGFβ2, and TGFβ3 in normal and malignant human prostate. Prostate 1997;33:133–140.

Appendix D

FERTILITY AND STERILITY®
VOL. 71, NO. 5, MAY 1999

Published by Elsevier Science Inc.
Printed on acid-free paper in U.S.A.

Salivary estradiol and progesterone levels in conception and nonconception cycles in women: evaluation of a new assay for salivary estradiol

Yu-cai Lu, M.D., Gillian R. Bentley, Ph.D.,† Peter H. Gann, M.D., Sc.D.,‡ Kelly R. Hodges, B.S.,* and Robert T. Chatterton, Ph.D.*§*

Northwestern University Medical School, Chicago, Illinois, and Cambridge University, Cambridge, United Kingdom

Objective: To determine the usefulness of salivary E_2 and progesterone for noninvasive assessment of ovarian function.

Design: Prospective study of salivary hormone levels in women planning a pregnancy.

Setting: Department of Obstetrics and Gynecology at Northwestern University Medical School in Chicago, Illinois.

Patient(s): Fourteen women aged 23–39 years with regular menstrual cycles who were planning a pregnancy.

Intervention(s): None.

Main Outcome Measure(s): Salivary estradiol and progesterone concentrations.

Result(s): The sensitivity of the E_2 assay is 2.0 pmol/L; the interassay coefficient of variation was 5.2% (mean value 17 pmol/L). Recovery of E_2 added to saliva was 106%. The correlation with simultaneous serum samples was 0.71. Menstrual cycle patterns contained a preovulatory depression and a midcycle surge. By comparison with nonconception cycles, the luteal phases of conception cycles had significantly elevated salivary E_2 within the first 5 days after ovulation. Salivary progesterone was significantly elevated but not until 10 days after ovulation.

Conclusion(s): Salivary measurements of E_2 and progesterone can be used as noninvasive methods for assessment of ovarian function. Salivary specimens can be collected at home and brought to the laboratory for analysis, obviating the need for frequent phlebotomy. The sensitivity and precision of the salivary E_2 assay make it comparable with assays of serum E_2 for assessing changes in hormone levels. (Fertil Steril® 1999;71: 863–8. ©1999 by American Society for Reproductive Medicine.)

Key Words: Conception, saliva, E_2, progesterone, assay, immunoassay

Received August 3, 1998; revised and accepted December 15, 1998.

Supported by the National Science Foundation (SBR 9507320), the National Institutes of Health (R01 CA66695, K07 CA66185), the U.S. Army Medical Research and Development Command (DAMD17-94-J-4203), and the Carol Gollob Foundation for Breast Cancer Research.

Reprint requests: Robert T. Chatterton, Jr., Department of Obstetrics and Gynecology, Northwestern University Medical School, Chicago, Illinois 60611 (FAX: 312-908-1372; E-mail: chat@nwu.edu).

* Department of Obstetrics and Gynecology, Northwestern University Medical School.

† Department of Biological Anthropology, Cambridge University.

‡ Department of Preventive Medicine, Northwestern University Medical School.

§ Department of Physiology, Northwestern University Medical School.

0015-0282/99/$20.00
PII S0015-0282(99)00093-X

Deviations from normal levels of serum E_2 and progesterone are of interest in evaluating ovarian deficiencies or failure, for monitoring ovarian stimulation and ovulation, and for assessing individual or population differences in sex hormone exposure. However, the day-to-day variation in these levels necessitates serial blood sampling or elaborate efforts to time blood draws within specific phases of the menstrual cycle. A noninvasive method for assessment of the levels of sex steroids would be an asset to the clinician as well as the researcher. Previous attempts to develop a valid, reliable assay for salivary E_2 have not been wholly satisfactory (1).

Because of their lipid solubility, steroids that are not bound strongly to serum proteins such as sex hormone-binding globulin or corticosteroid-binding globulin are available to diffuse through the lipid bilayer of the cell membrane and enter cells where, in target organs, they bind to intracellular receptors (2, 3). Thus, salivary steroids are thought to reflect the concentration of unbound serum steroids. In this respect the salivary concentrations may be a better measure of the exposure of target organs to the steroids than the serum concentrations are. Actual measurements of the concentrations of E_2 in saliva relative to those in serum range from 0.2% (4) to between 2.2% and 5.3% (5) and a high of 7.9% (6). These are

closer to the free concentrations rather than to the concentrations of non–sex hormone-binding globulin-bound concentrations in serum (7).

Proportions of progesterone levels in saliva are similar, between 0.82% and 2.1% of total serum concentrations (8–11). Here the agreement among investigators is better. The major difference between estimates is due to a higher percentage in the follicular phase (11). Previously published correlations between saliva and serum levels range from 0.76 to 0.89 (4, 6, 12) for E_2 and between 0.75 and 0.93 (8, 10, 13) for progesterone.

The low concentrations of E_2 in saliva have made this steroid particularly difficult to measure accurately. A number of investigators have developed satisfactory methods for salivary E_2 determination, but they tend to be labor-intensive and without the desired sensitivity. Therefore, we have attempted to develop a simple, direct method with adequate sensitivity, precision, and accuracy for assessment of normal levels of E_2 in saliva. We have validated the method and have compared levels in women who conceived a normal pregnancy with their levels in the preceding nonconception cycles. Progesterone was measured by a previously validated method to characterize further the conception and nonconception cycles of these women.

MATERIALS AND METHODS

Subjects

Fourteen women aged 23–39 years of age with regular menstrual cycles of from 25 to 35 days in length who were planning a pregnancy were recruited from the local Chicago area by advertisements in the newspaper. Exclusion criteria were as follows: outside the normal body mass index limits according to the Metropolitan Life Insurance tables, medication for chronic illnesses, oral contraceptives within the last 3 months, lactating within the last 6 months, evidence of infertility, or presence of sexually transmitted diseases. All women who volunteered for the study signed a consent form that was approved by the Institutional Review Board of Northwestern University. Two of the women did not conceive and dropped out, and one had an early miscarriage. Eleven subjects completed at least one nonconception cycle and a conception cycle. The last nonconception cycle was compared with the conception cycle.

At the time of entry into the study, the women were given a package containing 30 vials (14 × 50-mm copolymer vials) each containing 0.5 mg of sodium azide dried on the bottom as a preservative, sugar-free gum, a calendar, and urine specimen cups. They were asked to collect approximately 5 mL of saliva every morning while chewing the gum. The vials were sealed with a cap and stored at room temperature. The subjects recorded vaginal bleeding and sexual intercourse on the calendar and collected urine specimens on day 25 of the cycle (day 1 was the first day of menstrual bleeding) and on day 30 if menstruation had not occurred by that time. At the end of the 30-day period, the subjects brought the package containing the diary, saliva, and urine specimens to the laboratory.

Another group of seven women not planning a pregnancy were recruited for comparisons of salivary and serum levels of E_2 during a single menstrual cycle. During a single menstrual cycle, serum was collected two to three times per week in the morning after an overnight fast. Saliva was collected in the evening as above, but specimens were preserved by freezing rather than by use of the sodium azide preservative.

Assays

Urine specimens were analyzed for hCG by use of the QUPID pregnancy test kit (Stanbio Laboratory, Inc., San Antonio, TX) as soon as the specimens were brought to the laboratory. Saliva samples were stored at −20°C until the day of the assay.

Salivary E_2

Salivary E_2 concentrations were measured with use of a double-antibody RIA with [^{125}I]E_2. Antiserum and tracer were obtained from Diagnostic Services Laboratories (Webster, TX). The antiserum cross-reacts 2.4% with estrone, 0.01% with estrone sulfate, 0.21% with 16-ketoestradiol, 2.56% with estradiol 3-glucuronide, 0.64% with estriol, and <0.1% with nonphenolic steroids tested. Pure E_2 for standards was obtained from Sigma Chemical Co. (St. Louis, MO). A solution of 1.0 μg E_2/mL in methanol was diluted with buffer A (0.1 M phosphate-buffered saline, pH 7.0, containing 0.015 M NaN_3 and 0.1% gelatin) to 50 pg/mL for the highest standard.

Additional standards were prepared by doubling dilutions to 1.56 pg/mL in buffer A. A precipitating solution was prepared by titrating the amount of sheep antirabbit gammaglobulin required for precipitation of 0.1 ml of 9% normal rabbit serum. This amount of sheep antirabbit gamma globulin was added to a solution of propylene glycol (4.8 g/dL) to form the precipitation solution.

Saliva was collected with sodium azide as preservative and was stored for an initial period at room temperature and, when received at the laboratory, was frozen at −20°C. Before analysis the saliva was thawed and centrifuged at 1,500 × *g* for 1 hour, and the supernatant was used for analysis. Samples of 0.4 mL were assayed in duplicate without extraction.

The procedure is as follows: The volume of standards added to assay tubes was 0.4 mL. A standard with no added E_2 was included. Tubes in which buffer was substituted for the antiserum were prepared to estimate nonspecific binding. The total [^{125}I]E_2 per tube was approximately 15,000 cpm. Antiserum was diluted to give approximately 40% binding. The total reaction volume was 0.6 mL/tube. Standards and

TABLE 1

Recovery of added E_2 to pools of saliva with different endogenous concentrations of E_2.

Sample	Endogenous (pmol/L)	Added (pmol/L)	Expected (pmol/L)	Observed (pmol/L)	Recovery (%)
I	7.0*	5.7	12.7	12.0	94%
		11.5	18.5	19.2	104
		23.0	30.0	36.4	121
		45.9	52.9	53.1	100
II	13.4	11.5	24.9	24.5	98
		23.0	36.4	38.0	104
		45.9	59.3	62.2	105
III	26.7	11.5	38.2	40.6	106
		23.0	49.7	51.8	104
		45.9	72.6	83.9	116
IV	41.4	11.5	52.9	61.4	116
		23.0	64.3	63.0	98
		45.9	87.3	97.6	112

* Male saliva pool. Other pools are from female subjects.

TABLE 2

Analysis of male saliva for E_2.

Subject	First assay (pmol/L)	Second assay (pmol/L)
1	5.80	3.19
	5.95	4.66
2	4.74	4.00
	5.32	5.03
3	11.60	12.70
	14.54	13.81
4	10.90	8.22
	13.55	7.86
Mean	9.10	7.44

samples were incubated for 3 hours at room temperature. Then 0.1 mL of 9% normal rabbit serum and 1.0 mL of the precipitation solution were added, mixed, incubated at room temperature for 15 minutes, and then centrifuged for 20 minutes at 1,500 *g*. The supernatant was decanted, and the pellet was counted in a gamma scintillation spectrometer. The standards were plotted on a logit-log scale, and the concentrations of the samples were calculated from the linear standard curve.

Multiple vials of low and high level quality controls (QC) were prepared and frozen for use in each assay. Quality control pools were prepared from saliva collected from women volunteers during the early to midfollicular phases of their cycles and from a woman in midpregnancy.

The mean ED50 was 4.3 pg/mL. The intraassay coefficient of variation (CV) calculated from the differences of duplicate determinations of samples was 5.1%. Interassay CVs from 3 QC pools of 17, 24, and 46 pmol/L were 8.4%, 6.0%, and 7.8%, respectively.

The recovery of E_2 added to four different pools of saliva is shown in Table 1. The average recovery was 106%. There was no trend toward higher or lower recovery with the amount of E_2 added or with the endogenous estradiol concentration in the pools.

Similarly, doubling dilutions to 64-fold of three samples with original values of 338, 152, and 109 pmol/L gave an average of 106% ± 3% of the expected values, demonstrating the parallelism of the response (data not shown).

In addition, saliva specimens from four young men were assayed on two occasions (Table 2). The mean concentration was 8.3 pmol/L. Statistically significant differences among subjects were detected by analysis of variance (ANOVA) with repeated measures ($P=.002$). The standard deviation calculated from the error mean square within subjects was 1.00 pmol/L. Therefore, the assay is capable of detecting differences as little as 2.0 pmol/L.

Salivary Progesterone

Salivary progesterone concentrations were measured with a modification of the competitive immunoassay described previously (8). Bound steroid was separated from unbound with dextran-coated charcoal (0.05% and 0.5%, respectively, in buffer A). [1,2,6,7-^{3}H]Progesterone was obtained from New England Nuclear Division of DuPont (Boston, MA). Pure progesterone for standards was obtained from Sigma Chemical Co. Sample preparation was as described for E_2. The standards for the assay were prepared in buffer A rather than in charcoal-stripped saliva. This buffer was found to give values identical to standards prepared in the stripped saliva. The volume of saliva used for assay was 0.2 mL. The sensitivity of the assay is 48 pmol/L. The intraassay and interassay CVs were 9.9% and 12.0%, respectively.

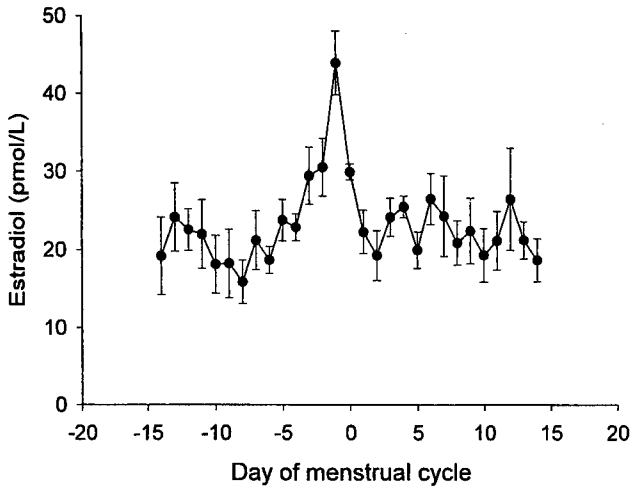
The antiserum, which was produced in this laboratory, cross-reacts 0.5% with 17-hydroxy progesterone, 2.9% with pregnenolone, 0.95% with corticosterone, 11.4% with 5α-pregnanedione, 1.9% with 5β-pregnanolone, and <0.1% with 20α-dihydroprogesterone and six other steroids tested. The correlation coefficient between this assay and a method using ether extraction of progesterone from saliva was 0.92 ($n=65$) as described previously (8).

RESULTS

The patterns of mean (± SD) daily salivary E_2 concentrations in nonconception and conception cycles for 11 subjects are shown in Figures 1 and 2, respectively. Individual cycles were normalized by setting the estimated day of ovulation as zero. Sexual intercourse occurred within 2 days of the estimated day of ovulation in all but one nonconception cycle (median of twice within this period of time). The day of ovulation was set according to Lipson and Ellison

FIGURE 1

Nonconception cycles. Levels of E_2 (means ± SE) in daily saliva samples from 11 subjects who were planning a pregnancy. Day zero is the day after the midcycle peak of salivary E_2 concentrations.



(12) as the first day after the midcycle peak of salivary E_2. This was a discernible value in each of the 11 cycles. Mean E_2 levels in both conception and nonconception cycles increased from approximately 17 pmol/L on day −10 to approximately 45 pmol/L on day −1.

According to a paired *t*-test of differences, there was no statistically significant difference between E_2 levels in conception and nonconception cycles in the preovulatory period. The mean difference across these days was 0.52 pmol/L (t = 1.55, P=.16). However, by comparison of differences in nonconception and conception cycles within subjects, E_2 levels were significantly higher in the conception cycles in the periods of days 1–5 (t = 6.06, P=.004), days 1–10 (t = 4.16, P<.002, and days 1–15 (t = 5.88, P<.001).

FIGURE 2

Conception cycles. Levels of E_2 (means ± SE) in daily saliva samples from 11 subjects who were planning a pregnancy. Day zero is the day of the first decline after the midcycle peak of salivary E_2 concentrations.

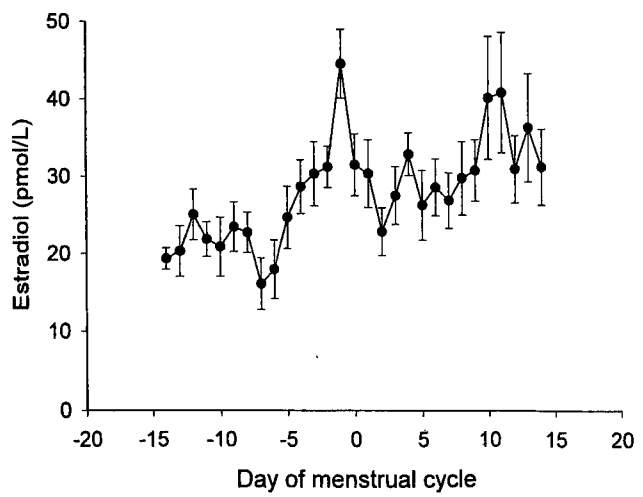


TABLE 3

Correlation of saliva and serum E_2 based on repeated sampling from the same women within a single menstrual cycle.

Subject	No. of saliva/serum pairs	Pearson's *r*	*P* value
1	7	−0.42	0.34
2	12	0.71	0.01
3	10	0.85	0.002
4	8	0.84	0.01
5	10	0.40	0.25
6	12	0.63	0.03
7	12	0.71	0.01
Total (all subjects)	71	−0.08	0.51
Excluding subject 1	64	0.21	0.09

The correlations between serum and saliva collected on the same day for each of the seven subjects recruited for this part of the study are shown in Table 3. Correlations for six of the subjects ranged between 0.40 and 0.85 with a median correlation coefficient of 0.71. Saliva-serum correlations were statistically significant for five subjects, even though the numbers of data pairs were small. The one subject who had a negative correlation had the fewest pairs and had two unexplained outlier saliva values that were >3 SDs from the mean.

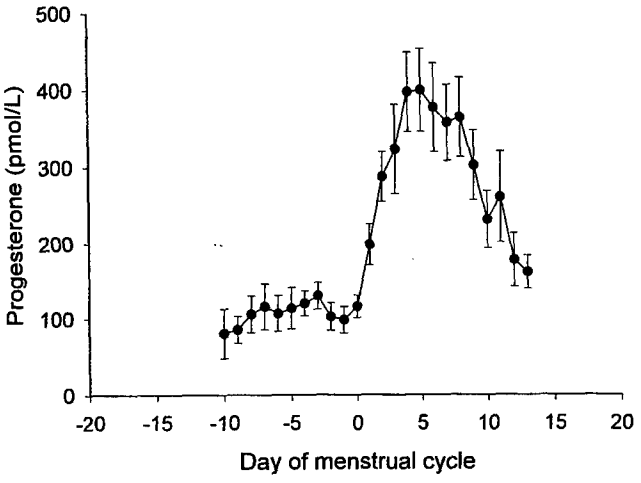
The salivary progesterone levels for the 11 subjects in the conception study are shown in Figure 3 and 4. The progesterone data confirm the precision of the method for determining the day of ovulation from the E_2 data as described above. There is a sharp rise in salivary progesterone immediately after the day of ovulation. The midcycle levels of progesterone, i.e., on day 7, are similar in conception and nonconception cycles with a sharp rise in progesterone thereafter in the conception cycles. The difference between nonconception and conception cycles between days 5 and 10 was 79.4 pmol/L (P=.27); the difference between days 10 and 15 was 429 pmol/L (P<.001).

DISCUSSION

This study validates a nonextraction method for measurement of salivary E_2. The method is adequately sensitive for detection of differences in levels of E_2 between nonconception and conception cycles soon after ovulation. The use of the ^{125}I-labeled E_2 as the tracer in the RIA greatly increases

FIGURE 3

Nonconception cycles. Levels of progesterone (means ± SE) in daily saliva samples from 11 subjects who were planning a pregnancy. Day zero is the day of the first decline after the midcycle peak of salivary E_2 concentrations.



the sensitivity over methods using the much lower specific activity tritium label. The elimination of the extraction step in the assay also leads to a considerably better precision than assays in which solvent residues can interfere with results, and sample recovery may vary from sample to sample.

By comparison with assays of serum E_2, results of the salivary assay suffer from a lack of knowledge about relationships to physiological function. Nevertheless, because the correlations between salivary and serum concentrations are generally quite high (14), changes in salivary steroids will have a close relationship to changes in serum steroids. The salivary methods have the advantage that multiple samples can be collected at home by the patient or research subject, providing more information than can be obtained from a single serum sample and without the need for phlebotomy.

In addition, it may be argued that salivary levels are a better indicator of the exposure of target organs to the hormone. Salivary steroid concentrations are more closely related to unbound, and therefore available, concentrations of hormones in serum. The contention is that steroid hormones that are not bound to serum-binding proteins such as sex hormone-binding globulin and corticosteroid-binding globulin are available to enter cells by diffusion (2, 3) including salivary gland cells. Unfortunately, this is difficult to prove because levels of several steroids including E_2 and progesterone are not present in salivary secretions in proportion to that fraction unbound to high affinity-binding globulins (4–6, 8–11). Undoubtedly, the steroid that enters the salivary gland is diluted by the process of active secretion of fluid (15).

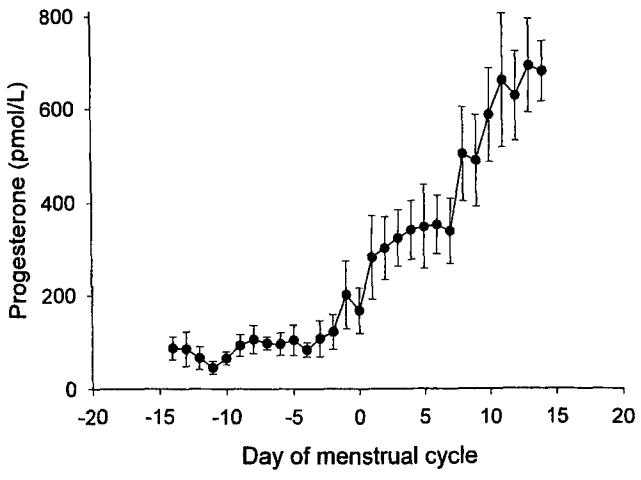
The patterns of E_2 found in this study are largely as expected, although, based on the well-established patterns of serum E_2, we would have expected to see a relatively higher concentration of salivary E_2 during the luteal phase of nonconception cycles than those in the early follicular phase of the same cycles. There is no indication that serum-binding globulins change during the course of the menstrual cycle (3), so there is no explanation for this apparent discrepancy at this time.

The correlation between salivary and serum E_2 was generally good within individuals. However, when the data for all seven subjects were pooled, there was no significant correlation between saliva and serum E_2. Thus, it is clear that the relationship differs greatly among individuals and that serum levels in one individual cannot be predicted from salivary concentrations in others. The reason for this probably derives from the interindividual differences in metabolism of E_2 within the salivary gland, the relative dilution by processes that regulate saliva flow, and possibly the level of binding proteins within the saliva.

We failed to confirm the greater preovulatory rise in salivary E_2 in conception cycles relative to nonconception cycles that was reported by Lipson and Ellison (12). The reason for this is not immediately obvious. Perhaps we lacked the statistical power to detect the difference, or there may be a difference between the populations studied. The finding of a highly significant increase in salivary E_2 immediately after ovulation in conception cycles relative to the levels in nonconception cycles was surprising but may confirm in part the data of Lipson and Ellison (12) because the patterns from their study also show a continued elevation in the postovulatory period. Salivary E_2 levels in conception

FIGURE 4

Conception cycles. Levels of progesterone (means ± SE) in daily saliva samples from 11 subjects who were planning a pregnancy. Day zero is the day of the first decline after the midcycle peak of salivary E_2 concentrations.



cycles also rose more rapidly in subjects of the present study than in subjects of the previous study (12).

In comparison with the recent study of urinary E_2 and progesterone metabolites in conception cycles by Baird et al. (16), we did not find an elevated progesterone level at the midluteal phase relative to nonconception cycles. Salivary progesterone was not significantly greater in conception cycles until after 10 days from ovulation.

Acknowledgments: The authors thank Lynette K. May and Kimberley G. Crone for excellent technical assistance.

References

1. Read GF. Status report on measurement of salivary estrogens and androgens. Ann NY Acad Sci 1993;694:146–60.
2. Rosner W. The functions of corticosteroid-binding globulin and sex hormone-binding globulin: recent advances. Endocr Rev 1990;11:80–91.
3. Petra PH. The plasma sex steroid binding protein (SBP or SHBG): a critical review of recent developments on the structure, molecular biology and function. J Steroid Biochem Mol Biol 1991;40:735–53.
4. Worthman CM, Stallings JF, Hofman LF. Sensitive salivary estradiol assay for monitoring ovarian function. Clin Chem 1990;36:1769–73.
5. O'Rourke MT, Ellison PT. Salivary estradiol levels decrease with age in healthy, regularly-cycling women. Endocr J 1993;1:487–94.
6. Belkien LD, Bordt J, Moller P, Hano R, Nieschlag E. Estradiol in saliva for monitoring follicular stimulation in an in vitro fertilization program. Fertil Steril 1985;44:322–7.
7. Toniolo P, Koenig KL, Pasternack BS, Banerjee S, Rosenberg C, Shore RE, et al. Reliability of measurements of total, protein-bound, and unbound estradiol in serum. Cancer Epidemiol Biomark Prev 1994;3:47–50.
8. Lu Y-C, Chatterton RT, Jr, Vogelsong KM, May LK. Direct radioimmunoassay of progesterone in saliva. J Immunoassay 1997;18:149–63.
9. Delfs TM, Klein S, Fottrell P, Naether OG, Leidenberger FA, Zimmermann RC. 24-hour profiles of salivary progesterone. Fertil Steril 1994;62:960–6.
10. Tallon DF, Gosling JP, Buckley PM, Dooley MM, Cleere WF, O'Dwyer EM, et al. Direct solid-phase enzyme immunoassay of progesterone in saliva. Clin Chem 1984;30:1507–1.
11. Evans JJ. Progesterone in saliva does not parallel unbound progesterone in plasma. Clin Chem 1986;32:542–4.
12. Lipson SF, Ellison PT. Comparison of salivary steroid profiles in naturally occurring conception and non-conception cycles. Hum Reprod 1996;11:2090–6.
13. Vuorento T, Lahti A, Hovatta O, Huhtaniemi I. Daily measurements of salivary progesterone reveal a high rate of anovulation in healthy students. Scand J Clin Lab Invest 1989;49:395–401.
14. Ellison PT. Measurements of salivary progesterone. Ann NY Acad Sci 1993;694:161–76.
15. Gallacher DV, Petersen OH. Stimulus-secretion coupling in mammalian salivary glands. In: Young JA, editor. Gastrointestinal physiology IV. Baltimore: University Park Press, 1983:1–52.
16. Baird DD, Wilcox AJ, Weinberg CR, Kamel F, McConnaughey DR, Musey PI, et al. Preimplantation hormonal differences between the conception and non-conception menstrual cycles of 32 normal women. Hum Reprod 1997;12:2607–13.

DRAFT – DO NOT CIRCULATE

Saliva as a Medium for Investigating Intra- and Inter-individual Differences in Sex Hormone Levels in Premenopausal Women

Peter H. Gann[1,2]
Susan Giovanazzi[1]
Linda Van Horn[1,2]
Amy Branning[1]
Robert T. Chatterton[1,2,3]

[1] Department of Preventive Medicine, [2] Robert H. Lurie Comprehensive Cancer Center, and [3] Department of Obstetrics and Gynecology, Northwestern University Medical School, Chicago, IL 60611

Corresponding author: Peter Gann, MD, ScD
Department of Preventive Medicine
Northwestern University Medical School
680 N. Lake Shore Drive, Suite 1102
Chicago, IL 60611
tel: 312.908.8432
fax: 312.908.9588
pgann@nwu.edu

Running title: Reproducibility of Sex Hormones in Saliva

ABSTRACT

Repeated measurement of ovarian steroids in saliva could provide an advantage in studies estimating long-term sex steroid exposure in premenopausal women, by reducing the measurement error associated with collection of serum or urine samples. We previously reported on characteristics of ultrasensitive radioimmunosassays adapted for extraction-free measurement of estradiol (E2) and progesterone (PG) in saliva. The purpose of the present study was to evaluate the consistency of E2 and PG levels in saliva in the same women across menstrual cycles, and to compare this to the variation observed between women. We also evaluated the effect of altering the number of consecutive daily samples considered and the method for locating a particular cycle day in relation to ovulation (day 0). Study participants included 12 healthy women who provided daily saliva samples for two consecutive, ovulatory menstrual cycles. A single mid-luteal serum sample was collected 7-8 days after detection of an LH peak in urine. We plotted individual cycle profiles and computed intraclass correlation coefficients (ICC) for various definitions of peak and cumulative daily hormone level. For peak PG, determined as the maximal running 3-day mean, ICC = 0.68. For cumulative PG, based on 8 consecutive cycle days (+2-+9), ICCs were 0.72 to 0.76 when reverse dating, LH peak or rise in salivary PG determined day 0. For E2, ICCs ranged from 0.74 to 0.79 by various dating methods for the 5 pre-ovulatory days (-4 –0), and from 0.85 to 0.92 for the 15 days from the center of the cycle (-6 -+8). With exclusion of just the first 5 days of the cycle, the ICC for E2 was 0.91. For both E2 and PG, selection of 5 or 7 days for estimation of the mid-luteal mean level provided separation of within and between subject variance that was comparable to an LH-timed serum sample. These results indicate that daily saliva samples can be combined to clarify the inter-individual differences in E2 and PG levels in premenopausal women, and that these inter-individual differences might be greater than previously imagined.

INTRODUCTION

According to the dominant paradigm, a woman's risk of developing breast (and perhaps endometrial) cancer is affected by long-term exposure to ovarian sex steroid hormones, primarily estradiol and progesterone (ref). Inter-individual differences in cumulative exposure to these steroids are therefore presumably important. It has proved difficult, however, to test aspects of this model in either etiologic or intervention studies involving premenopausal women, because of the fluctuation in these hormone levels across the menstrual cycle. Previous studies have reported a low correlation between two luteal-phase serum estradiol samples obtained from the same woman over a one year interval, even though sampling occurred approximately the same number of days after the onset of menses (refs). Standardized timing of blood samples can be improved by linking sampling to detection of ovulation or LH peaks, but such testing can be elaborate and/or time-consuming. We have adapted ultrasensitive radioimmunoassays for the direct measurement of estradiol and progesterone in saliva, and have previously reported on the sensitivity, reliability and serum-saliva correlations for these assays (ref). Saliva has several advantages over blood as a sampling medium: it can be easily collected by subjects themselves at repeated intervals, it requires no special collection or storage equipment, and the steroid concentrations measured exclude the fraction tightly bound to serum proteins and thus unavailable for biological action. Most importantly, consecutive daily samples can be grouped for analysis after the length of the menstrual cycle is known.

The aim of the present study was to determine how salivary levels of estradiol and progesterone track within women from cycle to cycle, and to evaluate how cumulative levels of salivary steroid based on daily sampling can be used to discriminate inter-individual from intra-individual differences. The ultimate goal is to develop a protocol for salivary measurement that reflects sustained exposure of target tissues to ovarian steroids, and thus provides a tool either for etiologic studies relating this exposure to disease risk, or for experimental studies of dietary or pharmacological interventions designed to reduce exposure.

METHODS

Study population

We recruited female participants from the Chicago area with the following eligibility criteria: age 20-40, no use of exogenous hormones within 6 months, regular menstrual cycles, and at least 12 months since previous childbirth or lactation. Potentially eligible participants were invited to the General Clinical Research Center (GCRC) at Northwestern University Medical School, where specially trained staff reviewed eligibility, obtained informed consent and collected baseline data on body size and past medical history. At the initial visit participants also received supplies and instructions regarding collection of the biological samples during the ensuing two consecutive menstrual cycles. A total of 20 women were recruited and 19 completed the two-cycle protocol.

Sample collection

During the two menstrual cycles, each participant provided samples of saliva, urine, breast fluid and blood. Beginning on the first day after the onset of menstrual bleeding, participants deposited their own saliva, while at home, into sequentially numbered empty plastic vials that were stored in cardboard boxes in the home freezer. The women were instructed to collect the

saliva daily between 8 and 11PM. Toothbrushing prior to sample collection was forbidden, to reduce the risk of contaminating the saliva sample with blood. Following thorough rinsing of the mouth with water, participants chewed sugarless gum to increase salivation while depositing 7-10 ml of saliva into the appropriate vial. The gum was demonstrated previously to have no effect on salivary steroid measurements (ref). Saliva collection was continued through the onset of bleeding marking the end of the second menstrual cycle. Completed boxes of saliva samples were returned to the GCRC where they were immediately catalogued and stored at -20°C.

For each participant, study staff calculated the expected day of ovulation based on the usual menstrual cycle length. Five days prior to this predicted day, participants began testing their first-void morning urine specimens using the Ovu-Quick® cassette for urinary LH (Quidel, San Diego, CA). Previous studies have validated this method for detection of LH in urine (ref). On the day the LH peak was detected, participants noted the day on a calendar provided and called the GCRC to schedule a mid-luteal appointment for 7 days following the LH peak. Mid-luteal appointments falling on Sunday were scheduled for the following day. In the event of a failure to detect a clear LH peak, participants were instructed to come to the GCRC 7 or 8 days after the predicted ovulatory day based on the usual cycle length. Venous blood samples were drawn in the morning after an overnight fast. Serum was separated and stored at -70°C.

Laboratory methods

All samples from a given individual were assayed in duplicate in the same run. Unidentifiable quality control samples were inserted in each batch to allow measurement of blinded intra- and interassay coefficients of variation (CV). Before analysis, frozen saliva was thawed and centrifuged at 1,500 x *g* for one hour; with the supernatant saved for assay. Salivary progesterone was measured with a modification of the competitive immunoassay described previously (Lu J Immunoassay 1997). Briefly, this assay employs tritium-labeled progesterone and an antiserum prepared by one of the authors (RC) which has known cross-reactivities of 0.5% with 17-hydroxy progesterone, 2.9% with pregnenelone, 0.95% with corticosterone, 11.4% with 5α-pregnanedione, 1.9% with 5β-pregnanolone, and <0.1% with seven other steroids tested. Standards were prepared in a special gelatinized buffer (0.1 M phosphate-buffered saline, pH 7.0, containing 0.015 M NaN_3 and 0.1% gelatin). The volume of sample used was 0.2 ml. Intraassay and interassay CVs were 13.5% and 18%, respectively.

Salivary E2 was measured with a double antibody radioimmunoassay also described previously (ref). Antiserum and [^{125}I]E2 tracer were obtained from Diagnostic Services Laboratories (Webster, TX). The antiserum has cross-reactivities of 2.4% with estrone, 0.01% with estrone sulfate, 0.21% with 16-ketoestradiol, 2.6% with estradiol 3-glucuronide, 0.64% with estriol and <0.1% with nonphenolic steroids tested. In the assay, the antiserum was diluted to give approximately 40% binding. Standards were prepared by diluting purified E2 in methanol with the same gelatin buffer used for progesterone. A precipitating antibody solution was prepared by titrating the amount of sheep antirabbit gamma globulin required for precipitation of 0.1 ml of rabbit serum, and adding this to propylene glycol (4.8 g/dL). The total volume of sample required was 0.4 ml. Intra- and interassay CVs were 9.9% and 11.6%, respectively.

Data analysis

We assayed salivary PG in all available samples, except those collected in the first 10 days of the cycle, from all 19 women who completed both cycles. For each cycle, midcycle day 0 was initially defined in two ways – as the day the urine LH peak was detected or by reverse dating, as the day located 14 days prior to onset of the subsequent menses. After daily salivary PG levels were measured, we identified the day 0 by onset of the rise in PG, defined as the middle day of the first 3-day running mean with an increase of at least 20% over the previous 3-day set,

and average daily level of at least 10 pg/ml. For PG, we defined two consecutive daily segments of interest a priori. The first was the 8-day segment from day +2 to day +9, and the second was the 1, 3, 5 or 7-day segment surrounding the mid-luteal center. Mid-luteal center was defined as the midpoint between day 0 and the end of the cycle. These definitions were arbitrary, but were chosen to include as many samples as possible while accommodating women with relatively short menstrual cycles. We computed 3-day running means for PG and defined peak PG as the highest running mean.

If a single daily saliva sample was missing we imputed a value by averaging levels from the preceding and following days. If 2 or more consecutive daily samples were missing, or if a given segment contained more than 2 missing samples, we excluded the segment from the analysis. Of the 19 women whose salivary PG levels were assayed, two had anovulatory cycles (one each) identified by absence of LH peak and serum PG < 3 ng/ml. Five other women had at least one cycle with segments disqualified by missing samples. Two women had one cycle apiece in which LH was not detected in urine, although serum and salivary PG suggested that ovulation had occurred.

For E2, we defined four segments of interest: the 5-day segment from day –4 to day 0 just prior to ovulation, the 15-day segment from day –6 to day +8, the mean daily E2 from all samples excluding the first 5 days, and the mid-luteal segment encompassing 1, 3, 5 or 7 days. We assayed all available saliva samples from 10 women, excluding those with anovulatory cycles or excessive numbers of missing samples. Peak E2 was not determined, because a single peak was not expected.

We used repeated measures ANOVA (SAS-PC®, SAS Institute, Inc., Cary, NC) to compute between-person and within-person variances. Intraclass correlation coefficients (ICCs) were computed as the between-person variance divided by the sum of between- and within-person variances (ref Fleiss). We calculated confidence intervals (95%) for the ICCs based on the expected distribution of the F statistic (ref Fleiss).

RESULTS

Among the 12 participants whose cycles were analyzed, mean age was 35 years (range 26-39), 6 were nulliparous, mean body mass index was 23.7, and mean age at menarche was 12.6 years. Mean daily saliva concentrations of E2 and PG across the menstrual cycle, aligned with the urine LH peak as day 0, are shown in Figure 1. Note that the E2 and PG curves correspond to the daily profiles normally observed in blood, but at far lower concentrations and with peaks of smaller amplitude. Figure 2 shows daily profiles of E2 and PG for consecutive cycles from selected individual participants, aligned by LH peak. Note that these representative curves show greater unexplained day-to-day variation in E2, leading to profiles that contain a discernible pre-ovulatory peak, but few other features typical of a classical serum profile. For PG however, salivary profiles always contained a distinct rise in PG at the onset of the luteal phase, followed by a relatively smooth, broad luteal peak. Individual and group cycle profiles were similar when reverse dating or rise in PG were used for alignment. Inter-individual differences in cumulative levels of both E2 and PG from cycle to cycle are readily apparent.

Cycle to cycle variability in peak and cumulative salivary PG is summarized in Table 1. For 24 eligible cycles, the intraclass correlation coefficient (ICC) for the peak level of PG was 0.68, meaning that only 32% of the total variance in peak PG was attributable to within-woman variation. For cumulative PG measured over 8 consecutive luteal days (+2 to +9), between-

woman differences accounted for 3 times as much total variance as within-woman differences. These ICCs are all significantly different from the null hypothesis of equal within and between-woman variance ($P < 0.01$).

Table 2 shows the comparison of within and between cycle variance for various cumulative measures of salivary E2. For a segment of 5 consecutive pre-ovulatory days (-4 to 0), ICCs ranged from 0.74 to 0.79, depending on the method of locating the midcycle day. For a 15-day segment from the center of the cycle, the ICCs were higher, ranging from 0.85 to 0.92 by method of alignment. For the mean daily E2, computed across the entire cycle excluding only the first 5 days, the ICC was 0.91. Again, all of the ICCs in this table have a very low likelihood under the null hypothesis of equal within and between-woman variance.

The effects on the ICC of varying the number of salivary samples used to estimate the mean mid-luteal steroid levels are shown in Table 3. The center of the mid-luteal period was defined as the midpoint between the midcycle day 0 and the end of the cycle. As expected, ICCs generally rise for both PG and E2 as one progresses from considering one to seven days in estimating the mean. The ICCs for a single mid-luteal serum sample, obtained 7-8 days after appearance of the urine LH peak, were 0.77 and 0.81 for PG and E2, respectively.

DISCUSSION

We have found that measurement of consecutive daily salivary estradiol and progesterone provides estimates of cumulative and peak concentrations that are consistent within women from one cycle to the next, and are relatively distinct between women. The ICC for peak progesterone was 0.68; and ranged from 0.72–0.76 for an 8-day luteal segment, depending on the method used for locating ovulation. A single mid-luteal saliva sample generally gave low ICCs for progesterone, but increasing the number of consecutive daily samples considered to 3, 5 or 7 gave ICCs comparable to that observed with a single serum carefully timed in relation to the urine LH peak. For estradiol, we observed substantial unexplained day-to-day within person variation in salivary concentration, but despite this, combination of consecutive daily saliva samples provided ICCs comparable to or greater than that provided by an LH-timed serum sample. When all but the first 5 cycle days were considered, only 9% of the total variance in mean daily salivary estradiol was attributable to within woman variance. This represents much higher within person correlation than observed for long-established measures used in population studies, such as serum total cholesterol measured one year apart (r = 0.65) (ref Shekelle, Stamler Liu).

It appears that the variation in daily salivary estradiol coexists with an even greater variation in time-integrated levels between women. Our data indicate that repeated sampling of saliva can be used to reveal potentially important differences in area-under-the-curve or time-integrated exposure of individual premenopausal women to endogenous sex steroids. The importance of even small differences in usual daily exposure, relative to cancer risk, are amplified because such differences are repeated month after month during a woman's reproductive years (ref Pike). Cross-sectional studies have reported substantially higher ovarian steroid hormone profiles in affluent Western women compared to women from non-industrialized cultures, providing evidence that the former have a high cumulative exposure to ovarian steroids due to higher exposure during a typical cycle, as well as a greater total number of menstrual cycles (ref Lancet). Within some agrarian cultures such as the Lese of Africa, seasonal variation in ovarian hormone levels suggests that large changes in caloric intake could be an important factor (ref Gillian). In the U.S., one study involving measurement of estrogen

and progesterone in daily urine samples from a single cycle in 175 women found lower progesterone levels in women with early menarche and higher body weight, and lower estrogen levels in cigarette smokers, but no other notable associations with reproductive variables associated with breast cancer risk (Westhoff). Regular exercise, whether vigorous or moderate, has also been associated with lower salivary progesterone profiles (Ellison). Polymorphisms in genes encoding key enzymes involved in sex steroid synthesis or metabolism are also under investigation as possible determinants of inter-individual differences in long term exposure (ref Ross Henderson). Thus far, however, it appears that most of the inter-woman variation in ovarian steroid levels remains unexplained.

We are aware of no previous studies comparing within and between woman variability in salivary estradiol. A single study of salivary progesterone in six cycles from each of eight women reported that inter-individual variance was approximately three times intra-individual variance – a result similar to ours (Sukalich abstract 1994). Several studies, however, have examined reproducibility of sex steroid hormone levels in serum or urine samples. Studies involving repeated measures in blood are few and tend to be small, due to the obvious practical constraints on performing many serial venipunctures. The largest such study measured daily plasma hormone levels during two cycles from 17 women (ref Lenton 1983). These investigators calculated the within-person correlation between cycles - roughly equivalent to the ICC - for plasma estradiol and progesterone during the interval +2 to +8 days after the LH peak. The correlation for progesterone was 0.80, very close to the estimate of 0.76 we obtained in saliva for nearly the identical interval. The between-cycle correlation for plasma estradiol was 0.49; although we did not assess this interval in saliva, the plasma result is very comparable to the ICCs we obtained for a 5-day mid-luteal segment (0.48-0.60) and considerably lower than we obtained for longer cycle segments.

In a study involving a single luteal phase serum sample, repeated one year apart, Muti, et al reported an ICC for estradiol of only 0.06 (ref). More recently, Michaud and co-workers, using a similar design, obtained an ICC of only 0.19 for a single luteal serum sample (ref). However, these investigators timed each blood sample collection to coincide with the *anticipated* mid-luteal day based on the individual's usual cycle length, and also had subjects report back on the onset of the subsequent menstrual cycle. They found that with exclusion of women with probable anovulatory cycles or samples obtained outside a 4-10 day window before the subsequent period, the estradiol ICC increased to 0.62. We note that our results, which also excluded anovulatory cycles, indicate a somewhat higher ICC (0.81) for a single serum sample timed more precisely, 7-8 days after detection of the urine LH peak. The method for timing a single blood sample used by Michaud, et al would have obtained a usable blood sample in 63% of all cycles. Use of saliva might exclude fewer cycles and provide higher ICCs over long segments, but it is not clear that these advantages outweigh the potentially higher costs of collecting daily saliva samples in all conceivable study applications.

Measurement of estradiol and progesterone metabolites in daily first-void urine samples provides some of the same benefits as salivary measurements in field study settings. Baird et al demonstrated that monitoring the urinary ratio of estrone-3-glucuronide and pregnanediol-3-glucuronide can be used to estimate the day of ovulation (Baird, 1995). Other investigators have reported menstrual cycle profiles for urinary estradiol and pregnanediol glucuronides – aligned by basal body temperature readings - that resemble those obtained in serum (Westhoff). Finally, the feasibility of daily urine collection was demonstrated in a study of 403 California women, which used creatinine-adjusted estrogen and progesterone metabolite levels to examine variation in follicular and luteal phase lengths. Urine samples require measuring

conjugated metabolites one or more steps removed from the active hormone, and thus could introduce possible additional error due to extraneous variations in metabolic activity.

Previous research has documented several advantages of saliva over blood or urine as a medium for frequent measurement of hormone levels (ref Read). Salivary samples can be collected over a matter of minutes by subjects themselves, and can be stored conveniently in home freezers, although studies have shown that most steroids remain stable for days in saliva, even at room temperature (Ellison field paper). Because it is easy to collect samples over an entire cycle, it is possible to select samples for assay afterward, once the length of the cycle is known. Estradiol and progesterone enter saliva by diffusion and represent the fraction of steroid not bound to carrier proteins. In principle, this means that salivary levels could provide a better reflection of the diffusable fraction available to target tissues such as the breast, however, this remains to be fully investigated. The salivary progesterone assay has been established for some time; our method is precise at levels as low as 15 ng/ml and provides a good correlation between synchronous saliva and serum samples ($r = 0.80$). Development of the estradiol assay, however, proved frustrating for many investigators. The evolution of immunoassay techniques permitted us to develop an assay for estradiol with good sensitivity and precision, and no required extraction step (ref). We found that serum-saliva correlations for estradiol (measured in serum as total or non-SHBG bound fraction) were not very high ($r = 0.21$) when serum-saliva pairs from many women were compared. However, median correlation within women was substantially higher ($r = 0.71$), suggesting that the ability of estradiol to move from serum to saliva varies between women.

Apart from the use of sensitive and precise direct assays for salivary hormones, the strengths of this study include estimation of ovulation by urine LH peak and rise in salivary progesterone as well as by reverse dating. Locating ovulation by detecting the rise in daily salivary progesterone is an established technique (Walker, 1985), as is detection of the mid-cycle decrease in salivary estradiol (Lipson and Ellison). However, measurement of individual daily samples is relatively expensive. In some field settings, it might be more practical to use self-detection of the urine LH peak, and in fact, our data show that a single serum sample collected at a fixed interval after LH detection is highly reproducible across cycles. However, some women with ovulatory cycles are not able to detect an LH peak with this method, and the daily urine testing around mid-cycle places a definite burden on study subjects and staff. The ICCs we observed for location of saliva samples by reverse dating were comparable to those obtained with LH or rise in progesterone dating. Therefore, we believe that it is most practical to collect samples from an entire cycle and then select samples from a broad consecutive segment located by reverse dating. The required sample volumes for a single assay for estradiol and progesterone are 400μl and 200μl respectively. It is possible to pool equal aliquots from consecutive days, and thus perform one assay to measure the cumulative or mean daily level over the segment.

We recognize that our data also have several limitations. The study size was not large, in part because it was necessary to assay many samples per subject in this methodological effort. In addition, we excluded women with anovulatory cycles or excessive missing samples. The criteria for missing samples were quite strict; it is possible that they could be relaxed in order to include more women and still provide adequate reproducibility, particularly for long cycle segments. Our definitions of cycle segments were somewhat arbitrary. In general, we would expect ICCs to increase as the number of samples considered increases. On the other hand, segments that are too long could exclude some women with unusually short cycles. Our study sampled saliva from two consecutive menstrual cycles. It is certainly possible that cumulative levels of sex steroids are more highly correlated in consecutive cycles than those further apart

in time. Individual women do have variations in estradiol from cycle to cycle, and this can even play a role in fecundity during specific cycles (ref Lipson, Hum Repro 1993). However, Lenton, et al reported that the separation between women in plasma progesterone during the interval +2 to +9 days after the LH peak was easily apparent even when they examined 5 cycles per woman spanning 3-6 years. Studies ongoing within our group will analyze reproducibility of consecutive daily segments from four cycles occurring over a one-two year period.

In conclusion, these results support the use of salivary measurements for exploring changes in cumulative exposure to endogenous estradiol and progesterone in premenopausal women. As a recent study has shown, a single mid-luteal serum sample could be adequate in a large prospective cohort setting where reverse dating can be used to exclude out-of-phase samples, and where the cost per subject is not great. On the other hand in trials or small observational studies with a high cost investment per subject, salivary measures could be more efficient by excluding fewer subjects and providing higher ICCs. More methodological research regarding the application of salivary methods is needed. In the context of breast cancer, the relevance of salivary steroid levels to the exposure of breast target tissue itself is a subject of great interest.

Table 1. Variation within versus between women in peak and cumulative salivary progesterone, measured during two menstrual cycles

	No. cycles	ICC	95% CI
Peak progesterone (3-day running mean)	24	0.68	(0.22-0.89)
Cumulative progesterone (days +2 to +9); midcycle (day 0) set by:			
Reverse dating	24	0.72	(0.30-0.91)
Urine LH testing	20	0.76	(0.32-0.93)
Rise in salivary progesterone	22	0.76	(0.37-0.93)

Table 2. Variation within versus between women in cumulative and mean daily estradiol, measured during two menstrual cycles

	No. cycles	ICC	95% CI
Cumulative estradiol			
Pre-ovulatory (days –4 to 0)			
Reverse dating	18	0.79	(0.35-0.95)
Urine LH testing	16	0.77	(0.27-0.95)
Rise in salivary progesterone	20	0.74	(0.27-0.93)
Center cycle (days –6 to +8)			
Reverse dating	20	0.85	(0.54-0.96)
Urine LH testing	16	0.92	(0.68-0.98)
Rise in salivary progesterone	20	0.86	(0.56-0.96)
Mean daily estradiol (entire cycle minus first 5 days)	20	0.91	(0.68-0.98)

Table 3. Intraclass correlation coefficients for midluteal[a] progesterone (PG) and estradiol, measured by a single, LH-timed serum sample versus varying numbers of consecutive saliva samples

	Number of consecutive daily samples			
	1	3	5	7
Saliva progesterone				
Reverse dating	0.21	0.63	0.65	0.66
Urine LH	0.73	0.83	0.68	0.73
Rise in salivary PG	0.16	0.58	0.75	0.81
Serum progesterone				
Urine LH	0.77	-	-	-
Saliva estradiol				
Reverse dating	0.27	0.35	0.53	0.67
Urine LH	0.26	0.41	0.48	0.62
Rise in salivary PG	0.23	0.32	0.60	0.69
Serum estradiol				
Urine LH	0.81	-	-	-