Published in final edited form as:

J Am Psychiatr Nurses Assoc. 2008 August ; 14(4): 273–284. doi:10.1177/1078390308322944.

Ecological salivary cortisol specimen collection (Part 1): Methodological consideration of yield, error, and effects of sampling decisions in a perinatal mental health study

Julia S. Seng, PhD, CNM¹, Anthony King, PhD¹, Cynthia Gabriel, PhD¹, Caroline D. Reed¹,
Michelle Sperlich, MA, CPM¹, Sara Dunbar, MPH, Emily Fraker, and David L. Ronis, PHD^{1,2}

¹ University of Michigan, Ann Arbor

² Ann Arbor Veterans Administration Health System

Abstract

Background—Current health research strives to integrate biological, psychological, and social factors consistent with ecological models. Home-based biomarker specimens are consistent with an ecological approach, but deviations from laboratory norms could affect validity of results.

Objective—This paper uses salivary cortisol specimens collected early in a perinatal mental health study to describe (1) return rate and returner characteristics, (2) adherence to procedures, (3) sources of laboratory error, (4) effects of deleting specimens with ‘nuisance’ factors, and (5) effects that selection bias could have on cortisol concentration distribution.

Study Design—Methodological analysis of collection, assay, and pre-analysis decision components.

Results—Rates of return do not differ by sociodemographic, perinatal, or psychiatric characteristics of participants. Exclusion of smokers affects representativeness. Selection bias in favor of more- or less-disadvantaged participants affects cortisol distribution.

Conclusions—The large yield of useable specimens permits multivariate modeling of cortisol level in association with health outcomes, enhancing ecological validity.

Keywords

ecological validity; salivary cortisol; methodology; posttraumatic stress; community-based research

Introduction

Large psychosocial studies on mental and physical health contribute to understanding of social and contextual factors that account for variance in risk, resilience, and recovery, especially when stress- or trauma-related etiologies are in question. Laboratory studies contribute valuable data about biological mechanisms, but design norms intended to increase internal validity limit external validity and result in sample sizes too small to permit complex modeling. Ecological or “ecosocial” approaches that include biomarkers in large-sample psychosocial studies are advocated to advance understanding of complex causal disorders and to attend to disparities across gender, race, ethnicity, and economic strata (e.g., Rich-Edwards et al., 2001; Rowley, 2001; Gillman et al., 2006). When large community

samples are involved, logistic and cost considerations often dictate departures from laboratory study norms. Instead of intensive protocols that apply inclusion and exclusion criteria to the sample, standardize procedures and timing, and have staff oversee specimen collection, biomarker specimens often are provided by participants from home according to written instructions and mailed to a laboratory. Cortisol is a stress-response hormone that reflects hypothalamic-pituitary-adrenal (HPA) axis functioning, making it a potentially useful biomarker for health outcome studies where stress-related disorders are of interest (e.g., Marin, Martin, Blackwell, Stetler, & Miller, 2007; Rosmond, Holm, & Bjorntorp, 2000). Cortisol can be validly measured in saliva, even during pregnancy (Vining, McGinley, Maksvytit, & Ho, 1983). Reliability of assays conducted when specimens are collected at home and mailed has been established (e.g., Clements & Parker, 1998; Kahn, Rubinow, Davis, Kling, & Post, 1988). There is less information available about the implications for internal and external validity when such biomarker data are obtained from large samples using home-based specimen collection. The purpose of this paper is to describe the impact of using an ecological approach on data collection, mailing and processing factors, and pre-analysis processes by describing

1. return rates and “returner” characteristics,
2. extent of error due to mailing and processing factors,
3. extent of variance in returners’ adherence to procedures,
4. effects of specimen exclusion decisions to limit ‘nuisance’ error on (a) the representativeness of the sample and on (b) the distribution of evening salivary cortisol concentrations as an example of one useful parameter, and
5. potential effects of socioeconomic selection bias on the distribution of evening cortisol concentrations.

A companion paper in this issue follows which focuses on analytic concerns, including relative effect sizes of potential nuisance factors, coping with potential lack of independence of some of these factors, and presenting an exemplar analysis in which several levels of ecological variables, including the biomarker, are considered in a multiple variable model of a sample health outcome.

1. Background

Cortisol is a marker of hypothalamic-pituitary-adrenal (HPA) axis function of interest for mental health and health (including perinatal) outcomes research because it has long been associated with stress regulation (Selye, 1956). Cortisol is found in blood, urine, and saliva. Saliva specimens are easily obtained, feasible to mail, and reliably assayed if frozen within 14 days (Clements & Parker, 1998; Kahn et al., 1988). Stress and HPA axis dysregulation evidenced in plasma cortisol have been considered in relation to adverse pregnancy outcomes, including labor processes and prematurity (e.g., Lederman, Lederman, Work, & McCann, 1978; McCool, Dorn, & Susman, 1994; Green, Damus, Simpson, et. al., 2005), and validity of salivary cortisol measures have been affirmed in pregnant women (Vining et al., 1983). Cortisol is currently of interest for mental health research because alterations in the diurnal cortisol profile, including both lower and higher levels (reviewed in Yehuda, 2006), have been associated with stress- and trauma-related psychopathology, including major depressive disorder (MDD), and our team’s main area of interest, posttraumatic stress disorder (PTSD). Recent studies have most often found differences between stress and psychopathology case and control groups in evening cortisol levels (e.g., Bremner, Vermetten, & Kelley, 2007; Inslicht et al., 2006; Ranjit, Young, & Kaplan, 2005; Young & Breslau, 2004), thus the analyses below will focus on that sample time point.

1.1. Importance of sample representativeness for ecological validity

Cortisol sensitivity to the focal factor of psychopathology might not be adequately specific, however. Research done primarily with women indicates that alterations in cortisol levels also are associated with the early life stress, especially maltreatment, which predisposes to psychopathology (e.g., Bremner et al., 2007). Cortisol levels also can be altered in the presence of acute stressors, such as intimate partner violence, that are more likely to occur among women with psychopathology, including both PTSD and depression (Inslicht et al., 2006; Pico-Alfonso, Garcia-Linares, Celda-Navarro, Herbert, & Martinez, 2004). Cortisol levels also can vary in relation to chronic stressors and daily hassles that do not affect all people equally (van Eck, Berkhof, Nicolson, & Sulong, 1996), in particular sociodemographic disadvantages such as material hardship and racism (Ranjit et al., 2005; Tull, Sheu, Butler, & Cornelious, 2005). Thus, multiple types of stress factors—past, acute, and differentially experienced—that might be relevant in mental health and perinatal studies, and that would be particularly valued under an ecological or “ecosocial” approach (Krieger, 2001), all might affect cortisol levels. If rates of returning specimens differ by sociodemographic characteristics, pregnancy risk factors, or trauma history and psychiatric status, the relative impact of these various stressors may not be captured. The extent to which the use of mailed instructions and shipment procedures used in large-sample studies result in a diverse sample needs further demonstration.

1.2. Importance of mailing mishaps and laboratory processing sources of error

Factors affecting quality of traditional laboratory-based neuroendocrine studies at the assay phase all remain important to the validity of large-sample, home-based biomarker studies, including intra- and inter-assay variability. Mailing specimens can result in additional sources of laboratory error, which may or may not turn out to be random, including packages being delayed or damaged in the mail resulting in moldy, broken, or missing tubes. This source of error of reduced specimen quality also remains to be quantified.

1.3. Importance of standardization of procedural and individual “nuisance” factors

Procedural factors also affect cortisol levels. Eating and smoking prior to specimen collection could raise cortisol levels (Badrick, Kirschbaum, & Kumari, 2007; Rosmond et al., 2000). The timing of specimens in relation to diurnal patterns of light-dark and sleep-wake cycles could also affect cortisol levels, making either standardization of timing or analytic techniques to cope with variance (e.g., Ranjit, Young, Raghunathan, & Kaplan, 2005) an important feature of study design if deviations are an important source of error. Several other nuisance factors also potentially affect cortisol levels: endocrine diseases, regular medication and substance use, including nicotine, steroid hormone levels related to menstruation or gestation, and body mass index or weight distribution. Depending on the research question, these influences on cortisol levels may represent random error or important covariates or confounds. They potentially represent threats to internal validity that could be handled via analytic procedures or by exclusion of specimens or participants from analysis. If nuisance factors occur at high rates or if they are not independent of the main variables of interest (e.g., psychopathology, cortisol, health outcome), then exclusion of these specimens could be a substantial threat to external validity. Empirical data regarding these factors would be useful to investigators designing data collection and analysis of biomarker data in ecological outcomes research.

1.4. Effects of deleting cases affected by nuisance factors

Deleting specimens affected by the above sources of error could introduce bias and threaten external validity if the “nuisance” factor is not randomly distributed in the sample. For example, use of asthma inhalers would likely be more prevalent among poor or urban

women (Wright & Subramanian, 2007), and use of psychiatric medications would occur more in the women with mental disorders. Thus the effect of deletions on representativeness of a large ecological sample is as important to consider as exclusion criteria in laboratory studies.

Conversely, not deleting these specimens could threaten internal validity of conclusions drawn about cortisol levels. There is little empirical data to inform these decisions. In laboratory studies that do not exclude smokers, for example, is it customary for authors to state whether smokers' cortisol values did or did not differ statistically significantly from non-smokers' values, and to include or exclude the smokers accordingly. This is likely to be the case in small laboratory studies. But an effect on cortisol level that is not statistically significant if there are 10% smokers in a sample of 20 could indeed be statistically significant if there were 10% smokers in a sample of 200. Whether this amounts to significant "error" or important "data" likely depends on the health condition being studied. Quantification of the effect on cortisol with and without smokers is needed.

1.5. Potential effects of selection bias

When an ecological approach is used, the disadvantage of any increased error introduced by greater variance in procedures should be balanced or outweighed by what is, theoretically, its greatest advantage: a large and representative sample. Large and representative samples should increase the validity of conclusions in stress-related psychiatric and perinatal health outcome studies by including women with diverse levels of sociodemographic disadvantage (an ecological stressor) and also by modeling the impact of such environmental factors. The requirement to come to a research or clinical facility to provide specimens in laboratory-based studies may introduce a *de facto* selection bias in favor of demographic groups who are either able or motivated to put forth this effort. Since health research done from an ecological perspective is particularly interested in contextual stressors, selection bias occurring on sociodemographic, geographic, or age group lines could be particularly threatening to ecological validity. The effect of *de facto* limiting of a small sample's socioeconomic status to either disadvantaged or affluent people has not been quantified in relation to cortisol levels.

2. Methods

The goal of this methodological analysis is to present descriptive data from the early months of an on-going study to begin to fill the above empirical gaps related to return rates, laboratory error, protocol adherence, and specimen exclusion decisions in an ecological approach, and to explore the potential impact of selection bias. In a companion paper that focuses on analytic issues we also will present sample cortisol analyses to illustrate and quantify the impact of analytic decisions on results and conclusions. In the following sections we will provide a brief description of pertinent aspects of the large-sample psychobiological perinatal outcomes study for which these salivary cortisol specimens were collected (R01 NR008767). We then describe details of the procedures related to the cortisol specimen collection, including some early adjustments made in response to low return rates. We also describe the mailing and laboratory processes, and outline our statistical analysis.

Description of the parent study

The analysis for this paper and the companion paper was conducted from preliminary survey and cortisol data collected from June 2005 through October 2006 as part of an on-going study of the effects of pre-existing posttraumatic stress disorder (PTSD) on perinatal outcomes. The parent study uses a prospective, nested case-cohort design to test the hypotheses (1) that PTSD is associated with adverse perinatal outcomes and (2) that

dysregulations in the diurnal pattern of salivary cortisol are associated with both PTSD and adverse perinatal outcomes. Participants are pregnant women 18 and older, able to speak English without an interpreter, expecting a first infant, and initiating prenatal care at less than 27 completed weeks of gestation. They are recruited via three large health systems in a metropolitan area of the Midwestern United States so as to over-sample African Americans living in poverty, a group at increased risk of adverse perinatal outcomes. Participation begins with a standardized computer-assisted telephone psychiatric diagnostic interview. Measures include demographics, trauma history (Wolfe & Kimerling, 1997) the National Women's Study PTSD module (Resnick, Kilpatrick, Dansky, Saunders, & Best, 1993; Kilpatrick et al., 1994), Composite International Diagnostic Interview depression and generalized anxiety disorder modules (Wittchen, 1994), and the CDC's Pregnancy Risk Assessment Monitoring Survey items related to pregnancy substance use (PRAMS; Gilbert, Johnson, Marrow, Geffield, & Ahluwalia, 1999). Respondents are assigned to one of three cohorts for follow-up, or they are dismissed. The cohorts are women never exposed to a traumatic event, exposed women who did not develop PTSD, and exposed women who meet lifetime PTSD diagnostic criteria. Women provide informed consent verbally at the start of the telephone interview. Those whose responses conform to a cohort definition are invited to continue in the study and are asked to provide saliva specimens by mail.

Original and modified procedures for cortisol specimen collection

Procedures related to the saliva sample collection process were pilot tested via two academic health systems within faculty practice prenatal clinics (Seng, Kane Low, Ben Ami & Liberzon, 2005). In the first months of the subsequent study, return rates were lower than in the pilot, so procedures were modified. Originally our procedures were as follows. Telephone interviewers followed a script describing the specimen collection kit and stating how and when to do the specimens. The kit included Salivette tubes (Sarstedt, Newton, North Carolina) labeled with the participant identification number and the time for each specimen. There were three forms included. First, an illustrated instruction sheet written at the 7th grade level explained the procedure and provided helpful hints (such as putting the tube next to the bed for the morning sample). It also displayed our toll-free telephone number, where answers to "frequently asked questions" are pre-recorded, and where participants could request replacement kits. Second, a sample checklist asked about eating, smoking, and timing of each specimen. Third, a health checklist asked if the woman had any major metabolic disorder and if she used any medications, including asthma inhalers, since these could affect cortisol levels. A check for \$10 was included as thanks in advance for returning the sample. After the first three months of data collection, our cortisol specimen return rate was near 40%, lower than the 60–80% rates in other community sample studies (e.g., Young & Breslau, 2004), and lower than in our pilot (Seng et al., 2005).

An ethnographer (author CG; unpublished data) conducted a qualitative inquiry into possible reasons for this low return rate and found that some women had concerns about other tests being run on the saliva (e.g., drug testing), but that, otherwise, reasons they gave for not returning the kit could be characterized simply as it not being a priority task.

Given that this rate is lower than in other studies, a research assistant (author CDR; unpublished data) conducted an anonymous survey study of pregnant women in one of the prenatal clinics from which we recruit. The goal was to learn what barriers there might be to intending to return the saliva specimens that we could address. She explained the saliva specimen collection purpose and procedure, and participants then reported their level of intention to provide specimens by mail. She asked about specific daily time points (upon awakening, 45 minutes later, and prior to lunch, supper, and bedtime) and gestational time points (e.g., during pregnancy, around delivery, and across the postpartum period). Analysis of results showed that intention to return the samples was similar across the day, except for

much lower levels of intention to complete the 45-minutes post-awakening sample. They were similar across gestation but markedly lower postpartum. Thus, with our limited resources, the choice to use bedtime, awakening, and pre-supper samples seemed affirmed. This small survey also found there were no sociodemographic or obstetric differences among those whose level of intention to return the specimens was low. Women of low socioeconomic status were, however, more likely to be influenced by the amount of reimbursement and the extent to which mailing was convenient for them.

In response to these small investigations, we implemented four modifications. We added to our telephone script an explicit statement that we test only for cortisol. We started reimbursing only those women who actually mailed the kit to the laboratory. We increased the reimbursement to \$20 for returned kits. We implemented a reminder phone call followed by a reminder letter if a kit had not been received within 2–3 weeks. The analysis for this paper includes specimens collected at the beginning and through the months when the return rate was just beginning to improve, as reflected in a return rate that had increased to a cumulative rate of 47% by the time we conducted the analysis presented here.

Laboratory processes

Specimen collection kits include a postage paid return mailer addressed to our co-investigators' laboratory. Laboratory staff members log the kits into a tracking system, noting the date the tubes are put into the -80°C freezer and any evident problems, including missing, broken, or moldy tubes. Data on the checklist forms are entered as well. Assays are conducted periodically on the thawed samples, distributing specimens from all three cohorts within each plate. Each specimen is assayed in triplicate using the commercially available Coat-a-Count radioimmune assay kit.

Analysis plan

2.1. Return rate and returner characteristics—Overall return rate and rates of return for participants with differing demographic, perinatal risk, and psychiatric profiles are presented using bivariate odds ratios via chi-squared test and adjusted odds ratios after multiple variable logistic regression. Effects of additive risk within these profile categories on being a “returner” are examined using the tau-b correlation coefficient.

2.2. Mailing and laboratory processing sources of error—Rates of problems with tubes and delay in freezing are reported, along with summary information about the variability within and among the assays.

2.3. Extent of adherence to procedures—Rates of missing specimens and deviations from the requested collection procedures are presented for each of the three (bedtime, awakening, and pre-supper) time points.

2.4. Effects of deleting specimens with nuisance factors—Representativeness of effects of deletions from the sample will be examined in a series of chi-squared tests across all 15 participant characteristics examined. Detailed descriptive statistics portraying the effect of the deletions on distribution of the evening cortisol levels also will be provided, along with t-tests determining when deletions result in significant changes in the mean.

2.5. Effects of potential selection bias—Because laboratory studies often do not provide detailed analysis or discussion of sociodemographic characteristics (factors which are known to affect stress and health outcomes) in relation to cortisol level results, little is known about the relative importance of these contextual factors compared with focal factors. We use a cumulative disadvantage index, a sum of five characteristics associated with

increased stress and/or adverse health outcomes: poverty, low education, being African American, being pregnant as an adolescent, and living in an inner city. A bimodal distribution guided stratification into two groups, those with fewer (0 or 1) versus more (2 to 5) disadvantages. We compared the perinatal and psychiatric risk profile characteristics of each group using chi-squared tests and compared the evening cortisol distributions using t-tests.

3. Results

The sample for this analysis includes 470 women who completed the first survey and were enrolled for follow-up as of September 15, 2006. Of these 470 women, 32.6% were enrolled for long-term follow-up because they met diagnostic criteria for past or current PTSD, 36% were in the trauma-exposed control group, and 31.5% were in the non-exposed control group.

3.1. Return rate and returner characteristics

A participant was considered a “cortisol returner” if her kit was received in the laboratory by October 10, 2006. Although kits were received from 225 returners, two had to be excluded from the analysis due to not submitting specimens until after their infant was born. Thus the effective return rate is 223 of 470 participants, 47.4%. Descriptions of the overall sample and the returner subset are compared in Table 1, showing that those who are African American, living in poverty, with high school education or less, obtaining care in a central city clinic, smoking, and currently meeting full diagnostic criteria for PTSD are less likely to be among the returners (at $p < .05$ level). Once the interrelationships among these factors are considered in a multiple logistic regression model, however, of these 15 demographic, pregnancy, or psychiatric factors only being a smoker during pregnancy is an independent predictor of not returning a kit ($p = .032$). Within each category, we summed the five characteristics and used tau-b to assess the correlation of an increasing number of sociodemographic stressors, pregnancy risk factors, and psychiatric status factors with likelihood of returning the cortisol specimen kit. Only increasing sociodemographic stress was correlated with decreased likelihood of returning specimens.

3.2. Mailing and laboratory processing sources of error

Statistics summarizing mailing and laboratory processes are presented in Table 2. Problems with the mailing process (e.g., torn envelopes with lost tubes, broken tubes, delays getting tubes into the freezer) and concentrations below the detectable limit of the assay are rare events. In our collaborator's laboratory inter- and intra-assay coefficients indicate very satisfactory assay quality and stability over time, representing a source of error that is likely to be random and is minor enough that it need not be considered statistically in outcomes models.

3.3. Extent of adherence to procedures

Rates of entirely forgetting to do one of the specimens were low at all time points, but highest prior to supper. Analysis of the checklist where returners report adherence to the procedures is presented in Table 3. Rates of providing the specimen more than one hour outside the suggested time frame were lowest at bedtime and highest at the pre-supper time. Rates of smoking or eating ahead of giving the specimen were lowest at bedtime and highest at the pre-supper time point when 1.8% of those who returned specimens smoked in the preceding half-hour. Seventeen percent reported eating in the 30 minutes prior to the evening, pre-supper specimen.

3.4. Effects of deleting specimens with nuisance factors

We now examine the effects of specimen deletion on representativeness (Table 4, part A) and cortisol concentration distribution (Table 4, part B) using the pre-supper, evening cortisol as the example. Nine of the 223 women who returned cortisol kits had the evening specimen missing, so the number of women's specimens considered is reduced to 214. The rest of the series of deletions are proxies for customary exclusion criteria that likely would be implemented because they are thought to introduce important levels of error (e.g., Scholtz, Schultz, Hellhammer, Stone, & Hellhammer, 2006). First, nine women's specimens (4%) were deleted because they had concentrations of .00 pg/mL. This could be due to concentration below the detection limit of the assay, due to contamination of the specimen, or due to wetting the cotton with water. Second, data provided on the health checklist indicate that a small proportion of returners have endocrine disorders (0.5%) or use medications that could be threats to the validity of cortisol concentration as a measure of stress or psychopathology (3%). Several participants were deleted in this next step, including those who reported they had diabetes (n=1) or who were using oral or inhaler steroids (n=5) and antidepressants (n=1), reducing the sample size to 198. Neither of these two sets of deletions resulted in statistically significant differences in the demographic, pregnancy risk, or psychiatric profiles compared with the women not deleted. Third, variations in the steroid hormone milieu can affect cortisol levels in women. In this study gestational age must be considered because cortisol concentrations have been found to rise across the last half of pregnancy (Challis & Patrick, 1983; Challis, Sprague, & Patrick, 1983). We thus deleted 30 participants whose interviews were completed after 20 weeks, resulting in a sample size of 168. Deleting the thirty women entering the study in the second half of pregnancy resulted in changes in the resulting sociodemographic profile, significant at the $p < .001$ level for all five characteristics. The rate of initiating prenatal care late (after 14 weeks) also decreases predictably and significantly. No other pregnancy or psychiatric characteristic rates changed with this deletion. Conversely, when the 19 women smoking during pregnancy were compared with the 195 not smoking, no differences in demographic profile result. However, rates of three pregnancy risk factors and three psychiatric diagnoses, all of which are known to be related to tobacco use, decrease, with levels of significance from $p = .033$ (late prenatal care) to $p < .001$ (PTSD and MDD).

In t-tests comparing the mean evening cortisol concentrations, none of the means of the progressively smaller samples differed significantly from the original sample of all 214 who returned evening cortisol or from each other, although the decrease in mean resulting from deletion of the remaining 16 smokers (sample E, $n = 152$) approached statistical significance ($t = 1.8752$, $df = 364$, $p = .062$). Skew and kurtosis were decreased when the thirty women in late gestation were removed. For subsequent analysis we use the sample that results when specimens from late gestation are deleted but smokers are retained ($n = 168$).

3.5. Effects of potential selection bias

Table 5, parts A and B respectively, depict the changes in sample characteristics and evening cortisol concentrations when samples are restricted to women with lower versus higher levels of sociodemographic disadvantage. Women with one or none of these were grouped as having less disadvantage ($n = 117$). Those with two or more were grouped as having more disadvantage ($n = 51$). By definition, these groups now have very different demographic characteristics ($p < .001$ for all five components of the disadvantage sum score). Their perinatal and psychiatric status profiles now also differ at a trend level ($p < .10$) for 8 of the ten individual factors. Only the characteristics of alcohol use and generalized anxiety disorder diagnosis do not appear to differ by sociodemographic status.

Sociodemographic disadvantage affects mean cortisol levels (Table 5, part B). The less disadvantaged group only differs from the whole sample of 168 women at trend level (versus .1048 pg/mL versus .1259 pg/mL $t=1.775$, $df=283$, $p=.077$). The more disadvantaged group, compared with the whole sample, has a significantly higher mean level (.1742 pg/mL versus .1259 pg/mL, $t=2.431$, $df=217$, $p=.016$). Contrasting the less- and more-disadvantaged groups with each other results in the largest difference (.1048 pg/mL versus .1742 pg/mL, $t=-2.981$, $df=60.2$, $p=.004$).

4. Discussion

The descriptive data from the early months of data collection which are presented here suggest that an ecological approach to collection of a salivary biomarker within a large-sample community-based study is feasible.

4.1. Return rate and returner characteristics

In this analysis conducted in the initial months of specimen collection, the return rate is lower in this study of diverse pregnant women than in other studies where an ecological approach was used but where specimen collection requests and instructions were made in person, and where the additional physical and psychological demands of pregnancy were not an issue (e.g., Young & Breslau, 2004). Returners did not differ from non-returners on any demographic, perinatal, or psychiatric factor when the interdependence of these characteristics was considered.

4.2. Laboratory sources of error

Mailing and laboratory sources of error were minimal. Only seven participants' specimens were delayed beyond 14 days in getting to the freezer. However, six of these seven women were in the more disadvantaged group, and were from the inner city, suggesting that processes to support women who do not have good access to mailing facilities would decrease bias in this potential source of error.

4.3. Extent of adherence to procedures

Adherence to procedures as reported on the specimen checklist forms varied across types of deviations and time of day. Adherence was generally worst for the pre-supper time. If the exact timing of specimen collection had been noted, additional analytic techniques could be used which turn this additional source of variance to advantage (Ranjit et al., 2005).

4.4. Effects of deleting specimens with nuisance factors

Deleting from analysis specimens affected by nuisance factors that are very customary to exclude from HPA axis studies (those with endocrine disorders, corticosteroid use, psychotropic medication use) did not result in loss of representativeness, nor did it change the mean evening cortisol values. Deletion of specimens from late gestation resulted in loss of African American, teen, poor, less educated, and urban women from the sample. Thus, the lower mean cortisol resulting with this deletion may be reducing variance associated with both higher steroid hormone levels and some extent of higher sociodemographic stress.

Although only three women reported smoking in the 30 minutes prior to giving the evening saliva specimen, those who were smokers had levels of cortisol that differed to an extent that stood out at trend level against all the error contained within the pre-deletion sample of 214 returners. Smoking co-occurred with other perinatal and psychiatric risk factors. Thus decisions about including specimens from smokers in ecological studies warrant careful consideration in light of the research question and smoking's relevance to a wide range of independent and dependent variables likely to be considered in health outcomes research.

4.5. Effects of potential selection bias

Analysis of the effect of limiting the sample to less- or more-disadvantaged groups showed that sociodemographic factors associated with increased stress have a significant association with evening cortisol levels. This finding is important as a plausible explanation for the mixed results in the posttraumatic stress literature if the current mixed results with regard to cortisol findings could be explained in part by high-stress versus low-stress ecological contexts affecting the samples in unquantified and often undiscussed ways. This is an area where using an ecological approach may improve internal validity as well as external validity. The relative effect sizes of these nuisance, contextual, and focal factors in relation to cortisol levels, their interdependence, and their relative independent contributions to analysis of an exemplar health outcome will be considered in detail in the companion paper.

There are several limitations to this methodological analysis. It is a *post hoc* examination; so many potentially relevant questions were not asked, including asking non-returners about barriers. It is limited to describing psychosocial characteristics and biological values of a community sample of pregnant women. The observed rates of endocrine diseases, smoking, and using medication may be lower than in samples of non-pregnant women, as might be their rate of returning the specimen kits. Recent literature suggests that both body mass index and fat distribution pattern affect cortisol profile, especially among women (Therrien et al., 2007), but we do not have these data to analyze. We also do not have adequate data to assess the impact of women whose wake-sleep cycle is significantly different than our procedures assumed. In future studies, asking women to note their usual times to go to bed, wake up, and eat supper as well as the times of the samples may provide more useful information and permit optimal analyses. This analysis may contribute to design of data collection for both future perinatal research and future psychiatric research; however, findings likely are most generalizable to studies where participants are both pregnant and selected for traumatic life events and/or psychopathology.

Strengths of this analysis include focus on a diverse population of young women where variance attributable to ecological factors is maximized. The sample was large enough to consider the potential interdependence of numerous sociodemographic, perinatal, and psychiatric factors that likely would be important to consider in studies focused on either pregnancy or psychiatric topics.

The companion paper focuses on the specimen analysis phase of the project and provides effect size data for a wide range of these variables, as well as examples of hypothesis testing and modeling to extend the descriptive data provided here about specimen collection processes.

From this analysis, investigators designing future studies from an ecological perspective have more extensive descriptive data than has been available to date. These data can inform specimen collection planning to optimize return rate and make use of increasingly sophisticated procedures for analyzing response to awakening and diurnal cortisol patterns (e.g., Ranjit et al., 2005). These data also will provide a quantitative basis from which to estimate the number of deletions, covariates, and thus the sample size needed to answer the research questions with the least error and the most generalizable results. These findings support the value, in terms of enhancing ecological validity, of going outside the laboratory setting to collect biomarker specimens.

References

- Badrick E, Kirschbaum C, Kumari M. The relationship between smoking status and cortisol secretion. *The Journal of Clinical Endocrinology and Metabolism*. 2007; 92(3):819–824. [PubMed: 17179195]

- Bremner D, Vermetten E, Kelley ME. Cortisol, dehydroepiandrosterone, and estradiol measured over 24 hours in women with childhood sexual abuse-related posttraumatic stress disorder. *The Journal of Nervous and Mental Disease*. 2007; 195(11):919–927. [PubMed: 18000454]
- Challis JRG, Patrick JE. Changes in the diurnal rhythms of plasma cortisol in women during the third trimester or pregnancy. *Gynecologic and Obstetric Investigation*. 1983; 16:27–32. [PubMed: 6684078]
- Challis JRG, Patrick JE. Relation between diurnal changes in peripheral plasma progesterone, cortisol, and estriol in normal women at 30–31, 34–35, and 38–39 weeks of gestation. *Gynecologic and Obstetric Investigation*. 1983; 16:33–44. [PubMed: 6684079]
- Clements AD, Parker CR. The relationship between salivary cortisol concentrations in frozen versus mailed samples. *Psychoneuroendocrinology*. 1998; 23:613–616. [PubMed: 9802131]
- DeSantis AS, Adam EK, Doane LD, Mineka S, Zinbarg RE, Craske MG. Racial/ethnic differences in cortisol diurnal rhythms in a community sample of adolescents. *Journal of Adolescent Health*. 2007; 41(1):1–2. [PubMed: 17577527]
- Gilbert BJC, Johnson CH, Morrow B, Gaffield ME, Ahluwalia I. Prevalence of selected maternal and infant characteristics, Pregnancy Risk Assessment Monitoring (PRAMS), 1997. *Morbidity and Mortality Weekly Report*. 1999; 48(SS-5):1–37.
- Gillman MW, Rich-Edwards JW, Huh S, Majzoub JA, Oken E, Taveras EM, Rifas-Shiman SL. Maternal corticotropin-releasing hormone levels during pregnancy and offspring adiposity. *Obesity*. 2006; 14(9):1647–1653. [PubMed: 17030976]
- Green NS, Damus K, Simpson JL, et al. Research agenda for preterm birth: recommendations from the March of Dimes. *American Journal of Obstetrics and Gynecology*. 2005; 193:626–35.
- Inslicht SS, Marmar CR, Neylan TC, Metzler TJ, Hart SL, Otte C, et al. Increased cortisol in women with intimate partner violence-related posttraumatic stress disorder. *Psychoneuroendocrinology*. 2006; 31:825–838. [PubMed: 16716530]
- Kahn JP, Rubinow DR, Davis CL, Kling M, Post RM. Salivary cortisol: a practical method for evaluation of adrenal function. *Biological Psychiatry*. 1988; 23(4):335–349. [PubMed: 3257706]
- Kilpatrick, DG.; Resnick, HS.; Freedy, JR.; Pelcovitz, D.; Resick, P.; Roth, S.; van der Kolk, B. *DSM-IV Sourcebook*. Vol. 4. Washington, DC: American Psychiatric Press; 1994. The posttraumatic stress disorder field trial: Emphasis on Criterion A and overall PTSD diagnosis.
- Krieger N. Theories for social epidemiology in the 21st century: an ecosocial perspective. *International Journal of Epidemiology*. 2001; 30(4):668–677. [PubMed: 11511581]
- Lederman RP, Lederman E, Work B, McCann DS. The relationship of maternal anxiety, plasma catecholamines, and plasma cortisol to progress in labor. *American Journal of Obstetrics and Gynecology*. 1978; 132(5):495–500. [PubMed: 717451]
- Marin TJ, Martin TM, Blackwell E, Stetler C, Miller GE. Differentiating the Impact of Episodic and Chronic Stressors on Hypothalamic-Pituitary-Adrenocortical Axis Regulation in Young Women. *Health Psychology*. 2007; 26:447–455. [PubMed: 17605564]
- McCool WF, Dorn LD, Susman EJ. The relation of cortisol reactivity and anxiety to perinatal outcome in primiparous adolescents. *Research in Nursing & Health*. 1994; 17:411–423. [PubMed: 7972919]
- Parker EA, Baldwin GT, Israel B, Salinas MA. Application of health promotion theories and models for environmental health. *Health Education & Behavior*. 2004; 31(4):491–509. [PubMed: 15296631]
- Pico-Alfonso MA, Garcia-Linares MI, Celda-Navarro N, Herbert J, Martinez M. Changes in cortisol and dehydroepiandrosterone in women victims of physical and psychological intimate partner violence. *Biological Psychiatry*. 2004; 56:233–240. [PubMed: 15312810]
- Ranjit N, Young EA, Kaplan GA. Material hardship alters the diurnal rhythm of salivary cortisol. *International Journal of Epidemiology*. 2005; 34:1138–1143. [PubMed: 15951357]
- Ranjit N, Young EA, Raghunathan TE, Kaplan GA. Modeling cortisol rhythms in a population-based study. *Psychoneuroendocrinology*. 2005; 30:615–624. [PubMed: 15854778]
- Resnick HS, Kilpatrick DG, Dansky BS, Saunders BE, Best CL. Prevalence of civilian trauma and posttraumatic stress disorder in a representative national sample of women. *Journal of Consulting and Clinical Psychology*. 1993; 61:984–991. [PubMed: 8113499]

- Rich-Edwards J, Krieger N, Majzoub J, Zierler S, Lieberman E, Gillman M. Maternal experiences of racism and violence as predictors of preterm birth: rationale and study design. *Paediatric Perinatal Epidemiology*. 2001; 15(2):124–135.
- Rosmond R, Holm G, Björntrop P. Food-induced cortisol secretion in relation to anthropometric, metabolic and haemodynamic variables in men. *International Journal of Obesity and Related Metabolic Disorders*. 2000; 24(4):416–422. [PubMed: 10805497]
- Rowley DL. Closing the gap, opening the process: Why study social contributors to preterm delivery among Black women. *Maternal and Child Health Journal*. 2001; 5(2):71–74. [PubMed: 11573841]
- Schlotz W, Schulz P, Hellhammer J, Stone AA, Hellhammer DH. Trait anxiety moderates the impact of performance pressure on salivary cortisol in everyday life. *Psychoneuroendocrinology*. 2006; 31:459–472. [PubMed: 16377094]
- Selye, H. *The Stress of Life*. New York: McGraw-Hill; 1956.
- Seng JS, Kane, Low LM, Ben Ami D, Liberzon I. Peak cortisol and perinatal outcomes in pregnant women with PTSD: A pilot study. *Journal of Midwifery & Women's Health*. 2005; 50:392–398.
- Therrien F, Drapeau V, Lalonde J, Lupien SJ, Beaulieu S, Tremblay A, Richard D. Awakening cortisol response in lean, obese, and reduced obese individuals: effect of gender and fat distribution. *Obesity*. 2007; 15(2):377–385. [PubMed: 17299111]
- Tull ES, Sheu YT, Butler C, Cornelious K. Relationships between perceived stress, coping behavior and cortisol secretion in women with high and low levels of internalized racism. *Journal of the National Medical Association*. 2005; 92(2):206–212. [PubMed: 15712783]
- van Eck M, Berkhof H, Nicolson N, Sulon J. The Effects of Perceived Stress, Traits, Mood States, and Stressful Daily Events on Salivary Cortisol. *Psychosomatic Medicine*. 1996; 58:447–458. [PubMed: 8902896]
- Vining RF, McGinley RA, Maksvytis JJ, Ho KY. Salivary cortisol: a better measure of adrenal cortical function than serum cortisol. *Annals of Clinical Biochemistry*. 1983; 20:329–335. [PubMed: 6316831]
- Wittchen HU. Reliability and validity of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research*. 1994; 28:57–84. [PubMed: 8064641]
- Wolfe, J.; Kimerling, R. Assessing psychological trauma and PTSD. New York: Guilford Press; 1997. Gender issues in the assessment of posttraumatic stress disorder. In J. P. Wilson, & T. M. Keane (Eds.); p. 192-238.
- Wright RJ, Subramanian SV. Advancing a multilevel framework for epidemiologic research on asthma disparities. *Chest*. 2007; 132(5):757–769. [PubMed: 17400684]
- Yehuda R. Advances in understanding neuroendocrine alterations in PTSD and their therapeutic implications. *Annals of New York Academy of Science*. 2006; 1071:137–166.
- Young EA, Breslau N. Saliva cortisol in posttraumatic stress disorder: A community epidemiologic study. *Biological Psychiatry*. 2004; 56:205–209. [PubMed: 15271590]
- Young EA, Tolman R, Witkowski K, Kaplan G. Salivary cortisol and posttraumatic stress disorder in a low-income community sample of women. *Biological Psychiatry*. 2004; 55:621–626. [PubMed: 15013831]

Table 1

Rate and Odds of Returning Kit Based on Participant Characteristics (n = 470)

Profile Factors	Proportion in sample	Return Rate n=223	Bivariate		Adjusted	
			OR	p	OR	p
Demographic Disadvantage Factor ^{a,*}						
African American	39.8	34.5	.823	.027	1.366	.381
Teen	19.8	16.1	.822	.060	.796	.441
Living in poverty	21.5	17.0	.799	.026	.861	.612
Low education	42.6	35.9	.784	.005	.822	.486
Central city clinic	40.2	33.6	.785	.006	.532	.102
Pregnancy Risk Factors ^{b,**}						
> 14 weeks at start of care	34.0	35.0	1.039	.684	1.606	.041
Past year domestic violence	3.0	3.6	1.233	.461	2.812	.094
Cigarette use	13.0	8.5	.914	.006	.496	.032
Alcohol use	14.3	13.0	.944	.461	.870	.623
Drug use	3.8	3.6	.728	.795	1.168	.763
Psychiatric Status Factors ^{c,***}						
Childhood abuse	23.6	21.1	.884	.218	.944	.818
Lifetime PTSD diagnosis	32.6	28.3	.842	.059	.783	.353
Current PTSD diagnosis	11.1	7.2	.729	.011	.631	.250
Current MDD diagnosis	12.6	12.1	.964	.782	.950	.878
Current GAD diagnosis	4.7	6.3	1.467	.119	2.489	.063

^aCorrelation of sum of sociodemographic disadvantages with likelihood of return: Tau-b = -.132, p = .002.^bCorrelation of sum of pregnancy risk factors with likelihood of return: Tau-b = -.050, p = .253.^cCorrelation of sum of psychiatric status factors with likelihood of return: Tau-b = -.071, p = .101.

* Significance of the demographic (first) step and model of the logistic regression was .073; Nagelkerke's R-squared .028.

** Significance of the perinatal (second) step was .023 and model, .010; Nagelkerke's R-squared .064.

*** Significance of the psychiatric (third) step was .232 and model, .012; Nagelkerke's R-squared .083.

Table 2**Mailing and Processing Factors and Assay Quality**

Indicator	%(n)
Envelope ripped in mail, tubes lost	1% (2)
Moldy tube	0.5% (1)
Problem with label	2% (4)
>14 days to freezer	3% (7)
Intra-assay variability	<5%
Inter-assay variability	<10%

Table 3

Proportion of Specimens Affected by Deviations in Collection Procedure (n=223)

Type of deviation	Percent of specimens affected by each deviation
Morning (defined as 7–9 am)	
Ate in 30 minutes preceding specimen	2.7
Smoked in 30 minutes preceding specimen	0.0
Outside the time-frame by ≤ hour	28.7
Outside the time-frame by >1 hour	8.5
Forgot/missed specimen entirely	0.9
Pre-supper (defined as 4–6 pm)	
Ate in 30 minutes preceding specimen	17.0
Smoked in 30 minutes preceding specimen	1.8
Outside the time-frame by ≤1 hour	31.4
Outside the time-frame by >1 hour	18.8
Forgot/missed specimen entirely	4.0
Bedtime (defined as 10 pm-midnight)	
Ate in 30 minutes preceding specimen	12.1
Smoked in 30 minutes preceding specimen	1.3
Outside the time-frame by ≤1 hour	11.7
Outside the time-frame by >1 hour	3.6
Forgot/missed specimen entirely	0.5

Table 4

Effect of Deletions on Representativeness and Evening Cortisol Distribution

Factors	Sample A All evening specimen returners n=214	Sample B Those with .00 values deleted n=205	Sample C Those with diabetes, med use deleted n=198	Sample D Those at >20 weeks gestational age deleted n=168	Sample E Those using tobacco in pregnancy deleted n=152
Profile of demographic, pregnancy risk, and psychiatric status characteristics of each sample					
African American	34.6	34.6	35.4	***28.6	27.6
Teen	15.9	16.1	15.7	***11.9	12.5
Living in poverty	16.8	17.1	17.2	***13.1	11.8
Low education	36.4	36.6	36.9	***31.5	30.3
Central city clinic	35.0	34.1	34.3	***28.0	27.6
Initiates care >14 wks	35.5	35.6	35.4	***23.8 ^a	*21.1
Past year DV	3.7	3.9	4.0	3.6	3.3
Cigarette use	8.9	9.3	9.6	9.5	0
Alcohol use	13.1	12.2	12.6	13.1	*10.5
Drug use	3.7	3.4	3.5	3.0	*2.0
Childhood abuse	21.5	20.5	21.2	23.2	21.7
Lifetime PTSD	28.5	29.3	30.3	32.1	***28.9
Current PTSD	7.0	6.8	7.1	6.5	*4.6
Current MDD	11.7	11.7	11.6	11.3	***8.6
Current GAD	6.5	6.3	6.6	6.5	5.9
Description of evening cortisol concentration profiles for each sample					
Mean	.1389	.1450	.1437	.1259	.1144
SD	.1413	.1413	.1421	.1120	.0917
Skew	3.463	3.531	3.599	3.083	2.753
Kurtosis	18.35	18.66	19.06	14.35	13.83

^a This proportion is reduced by deleting all who completed the first interview after 20 weeks of gestation. T-test results for the significant comparison: Sample A-Sample E comparison, $t=1.8752$, $df=364$, $p=.062$.

* $p < .05$.

100> <000
p

p<.01
**

Table 5

Effect of Selection Bias by Sociodemographic Level on Evening Cortisol Distributions

Factors	Sample D %(n) n=168	Lower risk %(n) n=117	Higher risk %(n) n=51	Lower vs. Higher p
Profile of demographic, pregnancy and psychiatric characteristics of each sample				
African American	28.6 (48)	5.1 (6)	82.4 (42)	<.001
Teen	11.9 (20)	0.9 (1)	37.1 (19)	<.001
Living in poverty	13.1 (22)	0 (0)	43.1 (22)	<.001
Low education	31.5 (53)	7.7 (9)	86.3 (44)	<.001
Central city clinic	28.0 (47)	1.7 (2)	88.2 (45)	<.001
Initiates care >14 wks	23.8 (40)	12.8 (15)	49.0 (25)	<.001
Past year DV	3.6 (6)	0.9 (1)	9.8 (5)	<.001
Cigarette use	9.5 (16)	6.8 (8)	15.7 (8)	.072
Alcohol use	13.1 (22)	12.0 (14)	15.7 (8)	.511
Drug use	3.0 (5)	0.9 (1)	7.8 (4)	.014
Childhood abuse	23.2 (39)	19.7 (23)	31.4 (16)	.098
Lifetime PTSD	32.1 (54)	25.6 (30)	41.7 (24)	.006
Current PTSD	6.5 (11)	0.9 (1)	19.6 (10)	<.001
Current MDD	11.3 (19)	8.5 (10)	17.6 (9)	.087
Current GAD	6.5 (11)	7.7 (9)	3.9 (2)	.364
Description of evening cortisol concentration profiles for each sample				
Mean	.1259	.1048	.1742	
SD .1120	.0757	.1585		
Skew	3.083	2.112	2.426	
Kurtosis	14.35	7.201	7.813	

Note. T-test results for part B: Sample D-Lower risk group comparison, $t=1.774$, $df = 283$, $p=.077$ Sample D-Higher risk group comparison, $t=2.431$, $df = 217$, $p=.016$. Lower-Higher risk group comparison, $t=-2.981$, $df = 60.2$, $p=.004$.