RUNNING HEAD: LTC During Pregnancy

Latent trait cortisol (LTC) during pregnancy: Composition, continuity, change, and concomitants

Gerald F. Giesbrecht,[1-3] Crystal I. Bryce,[3,4] Nicole Letourneau[1,5],
Douglas A. Granger [3,6-8] and the APrON study Team

[1]Department of Paediatrics, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[2]Alberta Children's Hospital Research Institute, Calgary, AB, Canada
[3]Institute for Interdisciplinary Salivary Bioscience Research, Arizona State University, Tempe, AZ, USA
[4]T. Denny Sanford School of Social and Family Dynamics, Arizona State University, Tempe, AZ, USA
[5]Faculty of Nursing, University of Calgary, Calgary, AB, Canada
[6]The Johns Hopkins University School of Nursing, Baltimore, MD, USA
[7]The Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA
[8]The Johns Hopkins University School of Medicine, Baltimore, MD, USA

Corresponding author: Address correspondence to Gerry Giesbrecht, PhD, Dept. of Paediatrics, University of Calgary, 355, 3820-24th Avenue NW, Calgary, Alberta Canada, T3B 2X9; email ggiesbre@ucalgary.ca.

Abstract

Individual differences in the activity of the hypothalamic pituitary adrenal (HPA) axis are often operationalized using summary measures of cortisol that are taken to represent stable individual differences. Here we extend our understanding of a novel latent variable approach to latent trait cortisol (LTC) as a measure of trait-like HPA axis function during pregnancy. Pregnant women (n=380) prospectively collected 8 diurnal saliva samples (4 samples/day, 2 days) within each trimester. Saliva was assayed for cortisol. Confirmatory factor analyses were used to fit LTC models to early morning and daytime cortisol. For individual trimester data, only the daytime LTC models had adequate fit. These daytime LTC models were strongly correlated between trimesters and stable over pregnancy. Daytime LTC was unrelated to the cortisol awakening response and the daytime slope but strongly correlated with the area under the curve from ground. The findings support the validity of LTC as a measure of cortisol during pregnancy and suggest that it is not affected by pregnancy-related changes in HPA axis function.

Keywords: latent trait cortisol; pregnancy; stability; cortisol awakening response; area under the curve, daytime slope

1. Introduction

In the past two decades the reactivity and regulation of the hypothalamic pituitary adrenal (HPA) axis has received considerable empirical attention as a mechanism for mediating individual differences in health and human development. Particularly in the area of fetal programming, pregnant women's HPA axis function has generated considerable interest (Mina and Reynolds, 2014). This area of inquiry has been bolstered by the development of reliable salivary cortisol assays that enable researchers to assess HPA axis function using minimally-invasive saliva collection procedures. The surge of research on salivary cortisol has been accompanied by the development of analytic tactics that allow researchers to model different aspects of HPA axis function, such as reactivity and recovery, the cortisol awakening response (CAR), the daytime slope or total cortisol productions (i.e., area under the curve from ground; AUCg). These analytic approaches have been accompanied by sampling strategies that recognize and capture the inherent variability of cortisol by repeating sample collection at various times of the day and over multiple days. Although it is seldom explicitly stated, the goal of all this sampling and summarizing across samples and days is to derive stable estimates of HPA axis function that reliably characterize individual differences. Unfortunately, these approaches have achieved only modest success toward this goal.

Recent studies suggest that only about 14%-18% of variance in measures such as the CAR, daytime slope and total daily output can be attributed to stable individual difference whereas the majority of variance is non-stable or state-specific (Hellhammer et al., 2007; Ross et al., 2014; Shirtcliff et al., 2012). Although incremental improvements can be made by collecting samples over a period of weeks (Hruschka et al., 2005), the burden in terms of cost and participant effort is unrealistic for the majority of studies. Moreover, the sampling protocol

requires a high level of participant compliance and even small deviations, especially for early morning samples, significantly reduce reliability (DeSantis et al., 2010; Kudielka et al., 2003).

1.1. Sources of Variability in HPA Axis Function

Activity of the HPA axis is driven by both exogenous and endogenous factors (Kertes and van Dulmen, 2012). Exogenous influences include a variety of situational factors that stimulate physiologic changes in response to changes in the environment. Most notably, the HPA axis is acutely sensitive to situations that are perceived as stressful and, in general, experiences of stress stimulate short-term increases in HPA axis activity (Dickerson and Kemeny, 2004). Endogenous factors, most notably the circadian rhythm - which itself is influenced by exogenous factors such as dark and light (Turek, 1985), also generate HPA axis fluctuations. The diurnal pattern typically includes a rapid increase in cortisol after waking and gradual declines for the remainder of the day (Hucklebridge et al., 2005). Together, the exogenous and endogenous factors determine cortisol levels at any given moment.

The HPA axis is an effective stress response system precisely because it can respond to the external and internal environments. Nevertheless, just as we expect some degree of consistency in behavior over time (i.e., personality), some degree of individual continuity in HPA axis function is expected. Physiological systems, like behavioral systems, tend to adopt a set point around which they operate (Berridge, 2004), and indeed, the HPA axis is normally regulated by a negative feedback loop. During pregnancy, however, the fetal-placental unit alters both the set point and the regulation of the HPA axis.

1.2. Stability and Change in Cortisol During Pregnancy

Human pregnancy offers a natural experiment in which to test the continuity and change in the HPA axis because of dramatic but non-pathological increases in cortisol as a function of

advancing gestation. During pregnancy, regulation of the maternal HPA axis is altered by a positive feedback loop with the fetal-placental unit (Sandman et al., 2006). Whereas circulating cortisol reduces HPA axis output in non-pregnant humans, during pregnancy the placenta secretes increasing amounts of CRH that stimulates the maternal HPA axis to increase cortisol production. Increasing levels of cortisol further stimulates the placenta to produce more CRH and this cycle continues until parturition. Pregnancy-related increases in cortisol thus provide an opportunity to evaluate the extent to which the unique physiologic state of pregnancy contributes to stable or trait-like individual differences. During pregnancy, the overall increase in cortisol production is large (3-4 fold) but the diurnal variation remains largely unchanged and HPA axis sensitivity to exogenous stimulation may be attenuated (Jung et al., 2011). Furthermore, placental CRH increases dramatically but does not exhibit a circadian rhythm (Magiakou et al., 1996; Wadhwa et al., 1997). These factors have the potential to increase the proportion of variance in salivary cortisol that can be attributed to stable individual differences.

The need for stable measures of HPA axis function is especially important to research examining the effects of prenatal cortisol on children's development. This is particularly true of the 'developmental origins of health and disease' literature that seeks to understand how individual differences in maternal HPA axis function during pregnancy are related to birth and development outcomes (Cottrell and Seckl, 2009; Glover, 2011; O'Connor, 2003; Sandman et al., 2012). A fundamental assumption has been that maternal stress biology communicates an intelligible signal that the fetus can use to alter its development (Glover, 2011; Gluckman et al., 2005). There is accumulating evidence that the transmission of such signals across the placenta has the ability to program developmental outcomes in offspring (Bale et al., 2010). Although fluctuation in maternal cortisol may provide useful information to the fetus, it seems likely that

the fetus would ignore many of the momentary fluctuations and focus instead on more stable components of the maternal HPA axis as the preferred signal by which to predict and adapt to the environment in which it will live. Accordingly, there is a need to develop analytic techniques that can maximize the stable signal within prenatal cortisol.

1.3. Latent Trait Cortisol (LTC)

Recently, several studies (Doane et al., 2015; Kertes and van Dulmen, 2012) reported that it is possible to greatly increase estimates of trait-like salivary cortisol, compared to other measures such as the CAR or daytime slope, by modeling latent trait cortisol (LTC) using confirmatory factor analysis. These studies found that 20% to 65% of variance in single cortisol measure was attributable to LTC. This represents a substantial increase over the estimated 14%-18% of variance in the CAR and daytime slopes that could be attributed to stable individual differences (Hellhammer et al., 2007; Ross et al., 2014; Shirtcliff et al., 2012). Of note, Doane and colleagues (2012) also showed that LTC was highly stable over a nine month period among adolescents.

LTC is measured using a latent variable approach. Specifically, a latent variable is modeled using a minimum of three cortisol samples. Because the cortisol samples share communalities (i.e., they are measuring the same underlying system), they are intercorrelated and can be included as indicators of LTC (Brown, 2006). With this approach, each indicator of cortisol is partitioned into two parts, a common and unique variance. Common variance is the portion of variance that is shared among the cortisol indicators whereas unique variance is the portion that is unique to each sample and excludes any random error (Brown, 2006). LTC is a latent measure of cortisol that reflects only the shared variance among the measured cortisol samples.

A clearer definition of trait-like aspects of maternal cortisol during pregnancy has the potential to improve our understanding of the associations between prenatal cortisol exposure and children's birth and developmental outcomes. The maternal HPA axis is assumed to mediate many of the effects of maternal stress on children's development and it therefore has become the focus of a large literature assessing the developmental origins of health and disease. Specifically, maternal stress and cortisol during pregnancy have been associated with infant birthweight (Bolten et al., 2011; Field et al., 2006; Kivlighan et al., 2008) – a factor which itself is a key predictor of many health outcomes including coronary heart disease, hypertension, and type 2 diabetes (Barker, 2004). Because of the positive association between maternal cortisol and infant birthweight, prenatal cortisol is expected to have broad and long-lasting effects on children's health and development.

1.4. Present Study

Building upon recent advances in the concept of LTC (Doane et al., 2015), our objective was to model LTC during pregnancy as a way to identify the stable portion of maternal prenatal cortisol. Specifically, our aims were to identify acceptable models of LTC during pregnancy, to determine whether LTC increases as a function of advancing gestation, to determine the associations between LTC and the diurnal pattern (i.e., CAR, daytime slope, AUCg), and to determine the association between LTC and infant birthweight. We assess the relationship between LTC and birthweight as a proof of concept to evaluate the validity of this analytic approach. We address these aims with data from a prospective cohort study in which diurnal suites of saliva were obtained in early, mid and late gestation.

2. Methods

2.1. Participants

Three hundred eighty women were recruited as early in pregnancy as possible and all prior to 22 weeks gestation. Eight women subsequently miscarried and were not included in these analyses. Participants were enrolled in the Alberta Pregnancy Outcomes and Nutrition study (www.apronstudy.ca), a prospective longitudinal study that is following a community sample of healthy volunteers. Women were excluded if they smoked or consumed alcohol during pregnancy, had a non-singleton pregnancy, or were taking synthetic glucocorticoids. The sample represents a relatively well-educated (89% had education beyond high school), married or common-law (98%), mature (mean age = 31.1 years, range 20-40 years) and primiparious (41%) group of women. The majority of the sample (82%) had a household income greater than $70,000 CAD/annum (according to Statistics Canada the median household income within the recruitment region is $98,030). The majority were non-Hispanic white (82%); the remainder were Asian (5%), Latin American (3%), Chinese (3%), Arab (2%), or other (5%). Participants provided informed consent prior to data collection. The study protocol was approved by the University of Calgary Health Research Ethics Board.

2.2. Procedure

Participants collected diurnal saliva at two or three time points in pregnancy (depending on time of recruitment): T1 = two days of diurnal samples at 6-14 weeks gestation, T2 = two days at 15-27 weeks gestation, and T3 = two days near the end of pregnancy (~ 32 weeks gestation). At each time point, women self-collected saliva at home during regular weekdays on the following schedule supported by a Personal Digital Assistant (PDA, i.e. Palm^{TM}): upon waking, 30 minutes after waking, at 1130h, and at 2030h. On each sampling day, participants collected the waking sample as soon after waking as practically possible and they initiated a 30 minute timer on the PDA. This procedure allowed for precise timing of the waking plus 30 minute sample while also

allowing for individual waking times. With the exception of the waking sample, the PDA rang to indicate that a sample was to be collected.

Each time the PDA rang, it first provided a code corresponding to a pre-labeled saliva tube and instructed the participant to place the saliva roll (Salivabio Oral Swab, Carlsbad, CA) under her tongue. The time of each assessment was recorded by the PDA, permitting precise modeling of diurnal patterns. Participants were asked to continue saliva collection until the swab was completely soaked (mean collection time = 1.9 min). To facilitate adherence to the study protocol, the PDA was programmed to allow a 20 minute response window following the signal, after which data were considered missing.

2.3. Salivary Cortisol

Participants were asked to refrain from consuming food, caffeine, citric drinks and dairy, to avoid vigorous exercise (e.g., running) or brushing teeth in the 30 minutes prior to saliva sample collection, and to report adherence to these guidelines. Whole saliva was obtained from under the tongue. Saliva samples were temporarily stored in participant home freezers (usually only 1-2 days) until they could be transported frozen to the laboratory in insulated hard-shell cases and freezer packs provided to the participants. Samples were stored at -80 C until they were shipped frozen to Salimetrics (State College, PA). All samples were assayed for salivary cortisol without modification to the manufacturer's instructions.

The cortisol test has a lower limit of sensitivity of 0.007 µg/dL, standard curve range from 0.012 to 3.0 µg/dL, and average intra-and inter-assay coefficients of variation 3.5% and 5.1% respectively. Method accuracy, determined by spike and recovery, and linearity, determined by serial dilution are 100.8% and 91.7%. In the current study, a random 10% of samples were

assayed in duplicate to confirm reliability; the intra-assay coefficient of variation was 3.5%. The mean value from duplicate samples was used for data analysis.

2.4. Infant Birthweight

Infant birthweight was extracted from the birth record and used to obtain birth weight percentiles, adjusted for sex and gestational age at birth, according to growth charts derived from the 1999-2000 US Natality Datasets (Oken et al., 2003).

2.5. Analytic Plan

Analyses were conducted using Mplus 7.11 (Muthén and Muthén, 1998-2015). Prior to analyses, data were screened to ensure sample time adherence and examine normality of data. Waking samples were recoded as missing if the sample was taken more than 15 minutes after waking. Similarly, waking+30 samples were recoded as missing if the recorded sample time was more than 59 minutes after waking. If the 1130h sample was taken less than 60 minutes after the waking time, the sample was recoded as missing. All cortisol samples were log transformed and samples that were four or more standard deviations from the mean were winsorized.

As is common in ambulatory studies, there was some missing data due to insufficient amounts of saliva or missed samples. In the final data set, 6 mothers only had first trimester data, 18 mothers only had second trimester data, 81 mothers had data at both first and third trimesters, 185 mothers had data at both second and third trimesters, and 82 mothers had data at all three trimesters. A full information maximum likelihood (FIML) estimator was used to accommodate missing data. Prior to all analyses, descriptive statistics and correlations among cortisol samples and sample time were examined (Table 1). Next, confirmatory factor analyses (CFA) were used to examine the measurement models of 12 different latent cortisol variables in order to determine which models best fit the data (Table 2) and to examine the factor loadings of the different latent

cortisol models (Table 3). The measurement models were estimated in order to specify a LTC variable and to identify and understand how samples of salivary cortisol are related to the LTC variable (Brown, 2006). Initially, no correlations among the errors of the cortisol samples were allowed to examine initial model fit. Then, we examined and utilized some of the modification indices provided by Mplus 7.11 (e.g., allow correlated errors between 11am samples) to improve model fit and these modifications were included in the final models.

To determine if LTC increases during pregnancy, we used the three daytime CFA models from the measurement modeling stage (Models 4-6, Table 2) to compare how much variance within each cortisol sample was accounted for by LTC (Figure 1). Furthermore, the stability of daytime LTC between first, second, and third trimester was examined using structural equation modeling (SEM; Figure 2) and correlations among LTC at each trimester (Table 4). To determine the relation between LTC and diurnal measures of cortisol, we assessed correlations between LTC, CAR, daytime slope, and AUCg (Table 5). For these analyses, the CAR (area under the curve increase) and AUCg were calculated with the trapezoid method (Pruessner et al., 2003) using samples collected at waking and 30 min post waking for CAR and samples collected at waking, 30 min post waking, 1130h and 2030h for AUCg. Daytime slope was calculated as 2030h – waking / total time since waking (Fekedulegn et al., 2007). Last, SEM was used to examine how LTC predicted infant birthweight percentile in both adjusted models (including ethnicity as a covariate) and unadjusted models (Table 6). The model fit in all CFA and SEM models was evaluated by examining the chi-square test of model fit, standardized root mean square residual (SRMR), root mean square error of approximation (RMSEA), and comparative fit index (CFI). We concluded that model fit was adequate when at least three of the four criteria

were met: (1) the *p*-value for the chi-square test of model fit >.05, (2) SRMR ≤ .08, (3) RMSEA ≤ .06, (4) CFI ≥ .95 (Brown, 2006; Hu and Bentler, 1998).

3. Results

Preliminary analyses were conducted to examine the Pearson product-moment correlations among cortisol samples (waking, waking +30, 1130h, 2030h) across all three trimesters. All correlations among samples taken at the same time were significant within each trimester and ranged between $r = .39$ and $r = .63$, with correlations in the first trimester generally stronger, mean *r* values = .57, .48, and .49 for 1st, 2nd, and 3rd trimesters, respectively (see Table 1). Correlations for samples taken at different times of the day were modest, mean *r* values = .24, .15, and .21 for 1st, 2nd, and 3rd trimesters, respectively. In addition, correlations between sample and time-since-waking for each sample were examined to determine if subsequent models should control for the time-since-waking. Given preliminary data screening approaches to ensure sampling time adherence, it is not surprising that time-since-waking was correlated to only two of the waking+30 samples (*r* values = .13 and .14, *p*s < .05) and the 1130h samples (*r* values ranged from -.32 to -.25, *p*s < .05). Time-since-waking was therefore included as a control variable in all analyses that included the 1130h samples.

3.1. Identifying Acceptable LTC Models

Next, we conducted CFAs to examine the LTC measurement models. Based on initial correlations among cortisol samples (see Table 1), a total of 12 CFA models were estimated: morning and daytime models were estimated separately for each of the 3 trimesters (Models 1-6, Table 2), then all samples collected at the same time of day and across all trimesters were combined (Models 7-10, Table 2), and finally all morning and daytime samples across all trimesters were aggregated to estimate overall models (Models 11-12, Table 2). According to

modification indices, some indicator errors were allowed to correlate in each model (see Table 2). Because measurement models were only retained if at least three of the four indices of model fit were adequate, four models were not included in subsequent analyses. Specifically, poor fit was observed for all morning models for the three trimesters (Models 1-3, Table 2) and the across trimester morning model (Model 11, Table 2). Although the daytime model for the second trimester had marginal fit (Model 5, Table 2), it was retained to create a complete understanding of daytime LTC across pregnancy. For the remaining models, the model fit was adequate (Table 2). Standardized factor loadings for the retained LTC models were generally greater than .45, with a few factor loadings ranging from .31 to .40 (Table 3) and all were significant (ps < .05).

3.2. Stability of LTC During Pregnancy

To determine whether LTC increases over the course of pregnancy we examined each trimester's daytime LTC to compare the variance accounted for in each by the cortisol samples (see Figure 1). Note that 1130h samples were allowed to covary across days within trimester and time-since-waking was regressed on each 1130h sample. Thus, the *R*-square for these samples is the proportion of variance accounted for by the latent factor, time-since-waking, and the covariance between 1130h samples. For the 2030h samples, the *R*-square is the proportion of variance accounted for by the latent factor. Overall, all LTC models accounted for a significant proportion of variance within each sample indicator and the amount of variance accounted for was relatively consistent over time, average R^2 = .44, .29, and .48 for 1st, 2nd and 3rd trimesters, respectively (see Figure 1).

To further investigate the stability of LTC over the course of pregnancy we examined an SEM in which first trimester daytime LTC predicted second trimester LTC, and in turn, second trimester LTC predicted third trimester LTC (see Figure 2 for parameter estimates). The overall

model fit was marginal, χ^2 (115) = 183.11, $p < .01$; RMSEA = .04 (90% CI; .03, .05); CFI = .91; SRMR =.08. First trimester daytime LTC significantly predicted 62% of the variance in second trimester daytime LTC, and in turn, second trimester daytime LTC significantly predicted 46% of the variance in third trimester daytime LTC.

Finally, we examined correlations among LTC over pregnancy by saving the LTC factor scores (FSCORE) for the trimester-specific daytime models and the overall daytime model. Correlations are shown in Table 4. Two observations are noteworthy. First, correlations among the trimester-specific factor scores were strong, suggesting that individual differences in LTC were stable over time (first and second trimesters $r = .66$; first and third trimesters $r = .51$; second and third trimesters $r = .47$). Second, correlations between trimester-specific daytime models and the overall daytime model decreased with advancing gestation, suggesting that LTC measured in early pregnancy is a stronger estimate of overall LTC during pregnancy compared to LTC measured later in pregnancy.

3.3. Associations between LTC, CAR, Daytime Slope and AUCg

Correlations among LTC factor scores (FSCORE), CAR, daytime slope, and AUCg are reported in Table 5, with the exception of values for the daytime slope, which were consistently non-significant ($r = -.03$ to .04). The highest correlations were evident among daytime LTC factor scores and AUCg ($r = .82$ to .89). Daytime LTC models were not significantly correlated with CAR. LTC models that included waking morning samples were negatively correlated with CAR ($r = -.32$ to -.41) and positively correlated with AUCg ($r = 24$ to .33). LTC models for the waking+30 samples were positively correlated with CAR ($r = .30$ to .37) and AUCg ($r = .33$ to.40).

3.4. Associations with Birthweight

Finally, 8 SEM models were examined to determine if LTC significantly predicted infant's birthweight percentile. In preliminary analyses we considered including the following covariates: maternal age, household income, gestational diabetes, pre-pregnancy BMI and ethnicity. Only ethnicity (0=non-white, 1=white) was significantly associated with birthweight, $r = -.14$, $p < .05$. In the final models birthweight percentile, which adjusts for gestational age at birth and sex, was regressed on LTC and ethnicity (Table 6). Because our interest was in the unique effect of LTC on birthweight, we report separately in in Table 6 the R^2 for unadjusted models (LTC only) and for adjusted models (LTC and ethnicity). The daytime 3rd trimester, waking, waking+30, 1130h, and 2030h models all had adequate fit. Of these five models, both daytime 3rd trimester and the 2030h model across trimesters significantly predicted birth weight percentile. LTC was negatively associated with children's birthweight percentile. Comparing the final 2 columns of Table 6 suggests that ethnicity accounted for approximately 1-3% of the variance in birthweight percentiles whereas LTC accounted for 3%-11% of birthweight variance.

4. Discussion

The aims of this study were to model LTC during pregnancy, determine its stability over the course of pregnancy, determine its association with the diurnal pattern (i.e., CAR, daytime slope, AUCg), and evaluate its association with infant birthweight. Acceptable models of LTC were identified for both morning and daytime samples. LTC was longitudinally stable during pregnancy and, contrary to our expectation, did not increase. No associations were observed between LTC and the daytime slope, but significant correlations were observed for the CAR and AUCg. In particular, daytime LTC models were strongly correlated with AUCg. Significant associations between LTC and birthweight were observed for daytime models and these models

explained a remarkably large portion of variance in infant birthweight. The findings support the reliability, validity, and stability of LTC during pregnancy.

In keeping with previous studies, we identified a reliable measure of LTC for both early morning and daytime cortisol (Kertes and van Dulmen, 2012; Kirschbaum et al., 1990). Early morning models had adequate fit only when samples across pregnancy were combined. In contrast, the daytime models had adequate fit (marginally adequate for 2nd trimester) within each trimester. Poor performance of the morning models may have resulted from greater variability in morning cortisol. Whereas daytime samples, especially evening samples, are tightly clustered around the sample mean, early morning samples display considerable scatter (see, for example, figure 1 of Shirtcliff et al., 2012). We note that all previous studies reporting early morning LTC have sampled on 3 separate days (at each occasion) in contrast to 2 days in the present study, and we speculate that the early morning models require 3 or more days in order to achieve adequate fit. This suggestion is supported not only by the previous studies but also by the fact that combining morning samples across pregnancy (i.e., 6 days) allowed us to achieve adequate model fit.

Daytime models generally had adequate fit and provided a basis on which to evaluate the change in LTC over pregnancy. Nevertheless, differences in model fit across pregnancy suggest that pregnancy-related adjustments in the maternal HPA axis may require special consideration for modeling LTC. Model fit was adequate for daytime LTC in the first and third trimesters but only marginally adequate in the second trimester. The reasons for marginal fit in the second trimester are not clear but may indicate a unique phase of pregnancy in which cortisol is a less stable construct. We note that correlations among individual cortisol samples (see Table 1) were lower in the second (mean r = .23) compared to the first (mean r = .32) and third (mean r = .28)

trimesters and because of the importance of inter-item correlations for LTC this may have contributed to the marginal fit in the second trimester. The reasons for these lower correlations in the second trimester are not clear, however changes in the ratio of free and total (free and bound) cortisol may contribute. Whereas total cortisol increases linearly over the course of pregnancy, there is a large increase (~44%) in free cortisol that occurs about mid-gestation, perhaps related to a reduction in HPA axis sensitivity to negative feedback (Demey-Ponsart et al., 1982). Whatever the reason for the reduction in correlation between cortisol samples in the second trimester, one implication for present purposes may be the need for additional days of sampling (beyond 2) to more adequately estimate LTC at mid-gestation. By including multiple days of samples, there may be more common variance among the cortisol samples that can be captured within the LTC factor.

Contrary to our expectation of increased variance accounted for by LTC as a function of advancing gestation, LTC was stable over the 3 trimesters. On average, variance in single cortisol measures attributable to the latent variable was 44% in first and 48% in third trimester, with second trimester somewhat reduced, perhaps for reasons described above. These findings are comparable to the 20%-65% previously reported in samples with children (Doane et al., 2015; Kertes and van Dulmen, 2012). Within the SEM model, the amount of variance in the LTC accounted for by previous trimester LTC was also substantial, 64% and 46% for second and third trimesters, respectively. The trimester specific daytime LTCs were highly correlated with the overall daytime LTC across pregnancy, suggesting strong stability during pregnancy. These findings suggest that LTC assessed within any one of the trimesters may be a reasonable proxy for an overall trait-like cortisol component. This may be especially true for first trimester LTC, which was correlated at $r = .84$ with the overall daytime LTC. It is important to replicate these

findings before any firm recommendations can be made, but if a reasonable estimate of overall LTC can be achieved with sampling at a single time point then it may be possible to reduce study costs and participant burden by sampling at only one time point in pregnancy.

Our study suggests that the stable portion of cortisol during pregnancy is largely independent of pregnancy-related increases in cortisol. Several factors may contribute to this finding. Although placental CRH stimulates overall increases in salivary cortisol during pregnancy it does not have a circadian rhythm (Magiakou et al., 1996; Wadhwa et al., 1997). Instead, the circadian variability in cortisol is most likely driven by arginine vasopressin (AVP) secreted into the portal system by parvocellular neurons of the PVN in a pulsatile fashion (Magiakou et al., 1996). This situation, increased level but not increased diurnal variability, may be akin to adding a constant to a distribution of scores. Adding a constant certainly increases the mean level but it does nothing to change the variability. Given the stability of correlations among diurnal cortisol samples over pregnancy (see Table 1) and the stability in variance accounted for by LTC (see Figure 1), these data suggest that the unique physiological state of pregnancy does not fundamentally alter trait-like cortisol.

The associations we observed between LTC and other well-established measures of cortisol lend support to its construct validity and help to define its potential use in research. The fact that early morning LTC was significantly correlated with the CAR and daytime LTC was correlated with AUCg but not CAR supports the discriminant validity of morning and daytime LTC. Although early morning LTC was also associated with AUCg, the correlations were smaller than those observed for the CAR. The modest associations between early morning LTC and CAR suggest that they are related but unique measures. These findings support the validity of the LTC method but also suggest that LTC may provide a better estimate of total cortisol

secretion than of change over time. Further support for this suggestion can be found in the very low correlations between daytime LTC and daytime slopes. LTC may therefore serve as an overall measure of cortisol production, perhaps similar to the estimate that can be derived from hair cortisol. It will be important in future studies to evaluate the association between hair cortisol and salivary LTC to help clarify what aspect of HPA axis function is assessed by LTC.

It is important to note that no studies to date, including our own, have been able to generate an LTC model with adequate fit that spans the entire day. This difficulty identifying an overall model is probably related to low correlations between morning and evening samples, even within the same persons on the same day. For this reason, and supported by numerous studies showing that early morning and daytime cortisol identify unique components of HPA axis function, we propose that separate estimates of LTC for early morning and daytime cortisol are not only required from a measurement perspective but also reflect the fact that morning and daytime cortisol provide largely unique insight into HPA axis function.

These findings suggest that daytime LTC overlaps strongly with AUCg but LTC has some distinct advantages over AUCg. First, it focusses on the trait-like aspect of cortisol, which is of primary interest for researchers examining the effects of prenatal cortisol exposure in offspring outcomes. Second, it can produce unbiased parameter estimates that are representative of the common variance among cortisol samples, even in the presence of missing data, which are inevitable in field studies. AUCg cannot be calculated for individuals with even a single missing data point resulting in a significant loss of information and power. It may also be possible to produce reliable estimates of LTC with fewer samples and days than are required for measures such as AUCg, which requires many days of sampling to achieve moderate reliability (Hruschka et al., 2005).

The association between LTC and infant birthweight percentile demonstrates the utility of LTC for investigating the effects of maternal cortisol on fetal and child growth and development. Of note, only daytime LTC was significantly associated with birthweight percentile. This is in contrast to several studies (Bolten et al., 2011; Kivlighan et al., 2008) who observed significant associations between early morning cortisol and birthweight. Both of those studies used measures that capture the dynamic change in cortisol after waking whereas LTC capture the stable portion of cortisol. These measurement differences may have contributed to differences in the pattern of findings. Although daytime LTC was associated with 11% of the variance in infant birthweight percentile, it is also important to note that because LTC was stable over the course of pregnancy, it may not be a suitable approach for those researchers who are interested in determining how the timing of exposures affect development.

Several strengths and limitations of our study should be noted. As mentioned previously, we sampled cortisol on two days within each trimester and this appears to be insufficient to achieve adequate model fit for early morning cortisol and for all aspects of cortisol in the second trimester. A large sample and prospective assessment in three trimesters are significant strengths but our sample under-represents women at the lower end of the socioeconomic scale and ethnic minorities. This may limit generalizability of these findings to those groups. We carefully screened our early morning data to ensure that our LTC models were based upon only those samples that were compliant.

Although the concept of LTC is not new, its properties remain relatively unknown. Studies that model LTC are sparse and our study adds substantial knowledge assessing its utility as a measure of HPA axis function during pregnancy. Based upon the current and previous studies, we offer some tentative conclusions about the measurement of LTC and its properties

during pregnancy. First, the validity of LTC as a measure of a stable trait-like component during pregnancy was supported. Second, LTC itself is stable during pregnancy and not affected by pregnancy-related changes in HPA axis function. Finally, we demonstrated the utility of LTC as a measure of maternal cortisol on a primary developmental outcome, fetal growth.

References

Bale, T.L., Baram, T.Z., Brown, A.S., Goldstein, J.M., Insel, T.R., McCarthy, M.M., Nemeroff, C.B., Reyes, T.M., Simerly, R.B., Susser, E.S., Nestler, E.J., 2010. Early life programming and neurodevelopmental disorders. Biol. Psych. 68, 314-319.

Barker, D.J., 2004. The developmental origins of adult disease. J. Am. Coll. Nutr. 23, 588S-595S.

Berridge, K.C., 2004. Motivation concepts in behavioral neuroscience. Physiol. Behav. 81, 179-209.

Bolten, M.I., Wurmser, H., Buske-Kirschbaum, A., Papousek, M., Pirke, K.M., Hellhammer, D., 2011. Cortisol levels in pregnancy as a psychobiological predictor for birth weight. Arch. Women's Mental Health 14, 33-41.

Brown, T.A., 2006. Confirmatory factor analysis for applied research. Guilford Press; US, New York, NY.

Cottrell, E.C., Seckl, J.R., 2009. Prenatal stress, glucocorticoids and the programming of adult disease. Front. Behav. Neurosci. 3, 19.

Demey-Ponsart, E., Foidart, J.M., Sulon, J., Sodoyez, J.C., 1982. Serum CBG, free and total cortisol and circadian patterns of adrenal function in normal pregnancy. J. Steroid. Biochem. 16, 165-169.

DeSantis, A.S., Adam, E.K., Mendelsohn, K.A., Doane, L.D., 2010. Concordance between self-reported and objective wakeup times in ambulatory salivary cortisol research. Int. J. Behav. Med. 17, 74-78.

Dickerson, S.S., Kemeny, M.E., 2004. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. Psychol. Bull. 130, 355-391.

Doane, L.D., Chen, F.R., Sladek, M.R., Van Lenten, S.A., Granger, D.A., 2015. Latent trait cortisol (LTC) levels: reliability, validity, and stability. Psychoneuroendocrinology 55, 21-35.

Fekedulegn, D.B., Andrew, M.E., Burchfiel, C.M., Violanti, J.M., Hartley, T.A., Charles, L.E., Miller, D.B., 2007. Area under the curve and other summary indicators of repeated waking cortisol measurements. Psychosom. Med. 69, 651-659.

Field, T., Diego, M., Hernandez-Reif, M., 2006. Prenatal depression effects on the fetus and newborn: a review. Inf. Behav. Dev. 29, 445-455.

Glover, V., 2011. Annual research review: Prenatal stress and the origins of psychopathology: an evolutionary perspective. J. Child Psychol. Psych. 52, 356-367.

Gluckman, P.D., Hanson, M.A., Spencer, H.G., 2005. Predictive adaptive responses and human evolution. Trends. Ecol. Evol. 20, 527-533.

Hellhammer, J., Fries, E., Schweisthal, O.W., Schlotz, W., Stone, A.A., Hagemann, D., 2007. Several daily measurements are necessary to reliably assess the cortisol rise after awakening: state- and trait components. Psychoneuroendocrinology 32, 80-86.

Hruschka, D.J., Kohrt, B.A., Worthman, C.M., 2005. Estimating between- and within-individual variation in cortisol levels using multilevel models. Psychoneuroendocrinology 30, 698-714.

Hu, L.-t., Bentler, P.M., 1998. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychol. Methods 3, 424-453.

Hucklebridge, F., Hussain, T., Evans, P., Clow, A., 2005. The diurnal patterns of the adrenal steroids cortisol and dehydroepiandrosterone (DHEA) in relation to awakening. Psychoneuroendocrinology 30, 51-57.

Jung, C., Ho, J.T., Torpy, D.J., Rogers, A., Doogue, M., Lewis, J.G., Czajko, R.J., Inder, W.J., 2011. A longitudinal study of plasma and urinary cortisol in pregnancy and postpartum. J. Clin. Endocrinol. Metab. 96, 1533-1540.

Kertes, D.A., van Dulmen, M., 2012. Latent state trait modeling of children's cortisol at two points of the diurnal cycle. Psychoneuroendocrinology 37, 249-255.

Kirschbaum, C., Steyer, R., Eid, M., Patalla, U., Schwenkmezger, P., Hellhammer, D.H., 1990. Cortisol and behavior: 2. Application of a latent state-trait model to salivary cortisol. Psychoneuroendocrinology 15, 297-307.

Kivlighan, K.T., DiPietro, J.A., Costigan, K.A., Laudenslager, M.L., 2008. Diurnal rhythm of cortisol during late pregnancy: associations with maternal psychological well-being and fetal growth. Psychoneuroendocrinology 33, 1225-1235.

Kudielka, B.M., Broderick, J.E., Kirschbaum, C., 2003. Compliance with saliva sampling protocols: electronic monitoring reveals invalid cortisol daytime profiles in noncompliant subjects. Psychosom. Med. 65, 313-319.

Magiakou, M.A., Mastorakos, G., Rabin, D., Margioris, A.N., Dubbert, B., Calogero, A.E., Tsigos, C., Munson, P.J., Chrousos, G.P., 1996. The maternal hypothalamic-pituitary-adrenal axis in the third trimester of human pregnancy. Clin. Endocrinol. (Oxf) 44, 419-428.

Mina, T.H., Reynolds, R.M., 2014. Mechanisms linking in utero stress to altered offspring behaviour. Curr. Top. Behav. Neurosci. 18, 93-122.

Muthén, L.K., Muthén, B.O., 1998-2015. Mplus user's guide, Sixth edition ed. Muthén & Muthén, Los Angeles.

O'Connor, T.G., 2003. Early experiences and psychological development: conceptual questions, empirical illustrations, and implications for intervention. Dev. Psychopathol. 15, 671-690.

Oken, E., Kleinman, K.P., Rich-Edwards, J., Gillman, M.W., 2003. A nearly continuous measure of birth weight for gestational age using a United States national reference. BMC Pediatrics 3, 6.

Pruessner, J.C., Kirschbaum, C., Meinlschmid, G., Hellhammer, D.H., 2003. Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. Psychoneuroendocrinology 28, 916-931.

Ross, K.M., Murphy, M.L., Adam, E.K., Chen, E., Miller, G.E., 2014. How stable are diurnal cortisol activity indices in healthy individuals? Evidence from three multi-wave studies. Psychoneuroendocrinology 39, 184-193.

Sandman, C.A., Davis, E.P., Buss, C., Glynn, L.M., 2012. Exposure to prenatal psychobiological stress exerts programming influences on the mother and her fetus. Neuroendocrinology 95, 8-21.

Sandman, C.A., Glynn, L., Schetter, C.D., Wadhwa, P., Garite, T., Chicz-DeMet, A., Hobel, C., 2006. Elevated maternal cortisol early in pregnancy predicts third trimester levels of placental corticotropin releasing hormone (CRH): priming the placental clock. Peptides 27, 1457-1463.

Shirtcliff, E.A., Allison, A.L., Armstrong, J.M., Slattery, M.J., Kalin, N.H., Essex, M.J., 2012. Longitudinal stability and developmental properties of salivary cortisol levels and circadian rhythms from childhood to adolescence. Dev. Psychobiol. 54, 493-502.

Turek, F.W., 1985. Circadian neural rhythms in mammals. Annu. Rev. Physiol. 47, 49-64.

Wadhwa, P.D., Sandman, C.A., Chicz-DeMet, A., Porto, M., 1997. Placental CRH modulates maternal pituitary adrenal function in human pregnancy. Ann. N Y Acad. Sci. 814, 276-281.

Table 1. *Correlations among Cortisol Samples within Trimester*

Day One	Day Two			
	Waking	Waking +30	1130h	2030h
First Trimester				
Waking	0.63*	0.41*	0.25*	0.29*
Waking +30	0.32*	0.54*	0.04	0.16*
1130h	0.17*	0.12*	0.58*	0.30*
2030h	0.20*	0.23*	0.36*	0.51*
Second Trimester				
Waking	0.57*	0.39*	0.07	0.10*
Waking +30	0.36*	0.46*	-0.01	0.09
1130h	0.07	0.17*	0.46*	0.21*
2030h	0.08	0.09	0.19*	0.41*
Third Trimester				
Waking	0.58*	0.36*	0.18*	0.22*
Waking +30	0.38*	0.48*	0.17*	0.15*
1130h	0.12*	0.09	0.39*	0.36*
2030h	0.13*	0.11*	0.24*	0.52*

Note. $n = 372$, * $p < .05$

Table 2. Descriptions of Latent Trait Cortisol (LTC) Measurement Models and Model Fit

Model	Description	χ^2 (df), *p* -value	RMSEA [90% CI]	CFI	SRMR
1. Morning 1st trimester [a]	Waking and waking +30 samples	χ^2 (1) = 4.33, *p*=.04	.14 [.03, .29]	.98	.02
2. Morning 2nd trimester [a]	Waking and waking +30 samples	χ^2 (1) = 28.65, *p*= .00	.32 [.22, .42]	.89	.05
3. Morning 3rd trimester [a]	Waking and waking +30 samples	χ^2 (1) = 26.57, *p* <.01	.28 [.19, .38]	.92	.04
4. Daytime 1st trimester	1130h and 2030h samples	χ^2 (7) = 11.58, *p* =.12	.06 [.00, .12]	.97	.04
5. Daytime 2nd trimester [b]	1130h and 2030h samples	χ^2 (7) = 15.50, *p* = .03	.07 [.02, .11]	.94	.04
6. Daytime 3rd trimester	1130h and 2030h samples	χ^2 (7) = 13.36, *p* =.06	.05 [.00, .09]	.98	.03
7. Waking	All trimesters	χ^2 (6) = 10.06, *p* =.12	.04 [.00, .09]	.99	.05
8. Waking+30	All trimesters	χ^2 (6) = 5.38, *p* =.50	.00 [.00, .06]	1.00	.03
9. 1130h	All trimesters	χ^2 (36) = 37.89, *p* =.28	.01 [.00, .06]	.99	.05
10. 2030h	All trimesters	χ^2 (6) = 14.64, *p* = .02	.06 [.02, .10]	.96	.05
11. Morning [a]	Waking and waking+30 samples, all trimesters	χ^2 (42) = 130.28, *p* < .01	.08 [.06, .09]	.91	.08
12. Daytime	1130h and 2030h samples, all trimesters	χ^2 (108) = 165.91, *p* = .12	.04 [.03, .05]	.92	.07

Note. *n* = 372, [a] Models were not retained for subsequent analyses. [b] Although the second trimester daytime model (Model 5) had marginal fit, it was included in subsequent analyses to provide a complete set of analyses for daytime LTC across trimesters.

Table 3. Standardized Factor Loadings for the Retained Latent Trait Cortisol (LTC) Measurement Models

LTC Model	Individual Cortisol Samples											
	First Trimester				Second Trimester				Third Trimester			
	Waking		Waking+30		Waking		Waking+30		Waking		Waking+30	
	Day 1	Day2	Day 1	Day2	Day 1	Day2	Day 1	Day2	Day 1	Day2	Day 1	Day2
Waking	.44 (1.00)	.72 (1.62)	--	--	.69 (1.49)	.49 (1.01)	--	--	.59 (1.28)	.63 (1.27)	--	--
Waking+30	--	--	.77 (1.00)	.77 (.97)	--	--	.56 (.73)	.64 (.78)	--	--	.48 (.58)	.56 (.67)
	1130h		2030h		1130h		2030h		1130h		2030h	
	Day 1	Day2	Day 1	Day2	Day 1	Day2	Day 1	Day2	Day 1	Day2	Day 1	Day2
Daytime 1st trimester	.56 (1.00)	.50 (.83)	.65 (1.10)	.76 (1.47)	--	--	--	--	--	--	--	--
Daytime 2nd trimester	--	--	--	--	.31 (1.00)	.35 (1.21)	.66 (2.30)	.62 (2.36)	--	--	--	--
Daytime 3rd trimester	--	--	--	--	--	--	--	--	.39 (1.00)	.51 (1.47)	.73 (2.48)	.72 (2.34)
1130h	.55 (1.00)	.63 (1.06)	--	--	.53 (.45)	.68 (.67)	--	--	.37 (1.12)	.48 (.79)	--	--
2030h	--	--	.43 (1.00)	.57 (1.50)	--	--	.63 (1.36)	.50 (1.19)	--	--	.51 (1.08)	.49 (.98)
Daytime	.57 (1.00)	.59 (.97)	.44 (.74)	.61 (1.18)	.45 (.65)	.49 (.79)	.52 (.82)	.39 (.68)	.47 (.55)	.50 (.66)	.52 (.81)	.46 (.68)

Note. Standardized factor loadings are followed by the unstandardized factor loadings in parentheses, $n = 372$. The unstandardized and standardized factor loadings for all models were significant, $p < .05$.

Table 4. Correlations among Daytime Latent Trait Cortisol (LTC) Factor Scores

	Daytime 1st trimester	Daytime 2nd trimester	Daytime 3rd trimester
Daytime 1st trimester	--		
Daytime 2nd trimester	.66*		
Daytime 3rd trimester	.51*	.47*	
Daytime (across trimesters)	.84*	.76*	.74*

Note. $n = 372$, * $p < .05$

Table 5. Correlations Between Latent Trait Cortisol (LTC) Factor Scores, CAR, and AUCg at Each Trimester

	1st Trimester		2nd Trimester		3rd Trimester		Across Trimesters	
Factor Scores	CAR	AUCg	CAR	AUCg	CAR	AUCg	CAR	AUCg
Daytime 1st trimester	-.05	.89*	-.01	.62*	-.05	.50*	-.03	.72*
Daytime 2nd trimester	.03	.57*	.06	.82*	.03	.46*	.06	.62*
Daytime 3rd trimester	-.01	.49*	.03	.49*	-.04	.85*	-.01	.62*
Waking (across trimesters)	-.41*	.29*	-.38*	.24*	-.32*	.33*	-.42*	.29*
Waking+30 (across trimesters)	.36*	.33*	.30*	.39*	.37*	.40*	.40*	.40*
1130h (across trimesters)	.04	.79*	.06	.79*	.07	.59*	.07	.73*
2030h (across trimesters)	-.01	.66*	.05	.70*	.01	.62*	.02	.69*
Daytime (across trimesters)	.02	.86*	.04	.82*	.04	.75*	.05	.84*

Note. $n = 372$, * $p < .05$

Table 6
Predicting Birth Weight from Latent Trait Cortisol (LTC) and Ethnicity

Model	χ^2 (df)	RMSEA [90% CI]	CFI	SRM R	LTC β (B)	R^2 (unadjusted)[b]	R^2 (including ethnicity)[c]
Daytime 1st	χ^2 (12) = 25.00, p = .01	.05 [.02, .08]	.93	.07	-.26 (-29.84)**	.07, p =.12	.10, p =.04
Daytime 2nd	χ^2 (12) = 24.87, p =.02	.05 [.02, .08]	.92	.04	-.25 (-61.14)*	.06, p = .09	.08, p = .04
Daytime 3rd Trimester [a]	χ^2 (12) = 12.97, p =.12	.04 [.00, .07]	.98	.03	-.18 (-42.22)*	.03, p = .15	.05, p = .05
Waking Across Trimesters [a]	χ^2 (11) = 16.32, p =.13	.04 [.00, .07]	.99	.04	.02 (3.14)	.00, p = .88	.02, p = .17
Waking+30 Across Trimesters [a]	χ^2 (11) = 9.65, p =.56	.00 [.00, .05]	1.00	.03	-.02 (-1.86)	.00, p =.87	.02, p =.16
1130h Across Trimesters [a]	χ^2 (47) = 56.29, p =.17	.02 [.00, .04]	.97	.06	-.12 (-13.43)	.02, p =.42	.03, p =.13
2030h Across Trimesters [a]	χ^2 (11) = 20.82, p =.04	.05 [.01, .08]	.96	.05	-.33 (-41.84) **	.11, p =.02	.13, p =.01
Daytime Across Trimesters	χ^2 (125) = 196.79, p <.01	.04 [.03, .05]	.90	.07	-.25 (-28.02) **	.06, p =.05	.08, p =.02

Note. Standardized estimates for LTC are followed by unstandardized estimates (in brackets), *p*-values correspond with unstandardized model results. Model fit indicators and parameter estimates are for the unadjusted model. Model fit and LTC parameter estimates for the adjusted model (not shown) were substantively the same as those presented above. n = 372. $^{**}p < .01$, $^{*}p < .05$. [a] model retained based on fit criteria, [b] models included only LTC as a predictor, [c] models included LTC and ethnicity as predictors.

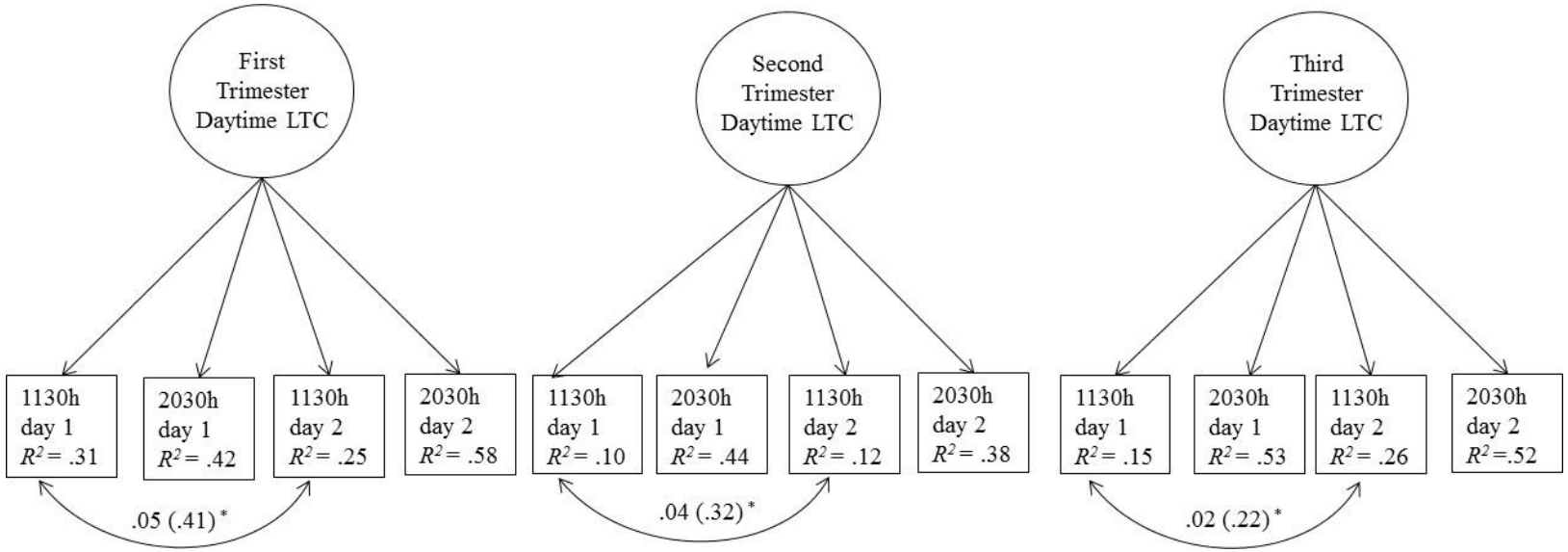

Figure 1. Model *R*-square values for latent trait cortisol (LTC) indicators by trimester. All three models were ran separately, n = 372. Time lapse since waking was regressed on the 1130h samples; thus, the *R*-square value for the 1130h indicators is a measure of the proportion of variance accounted for in the sample by the latent factor, time lapse since waking, and the 1130h covariance. * $p <$ .01

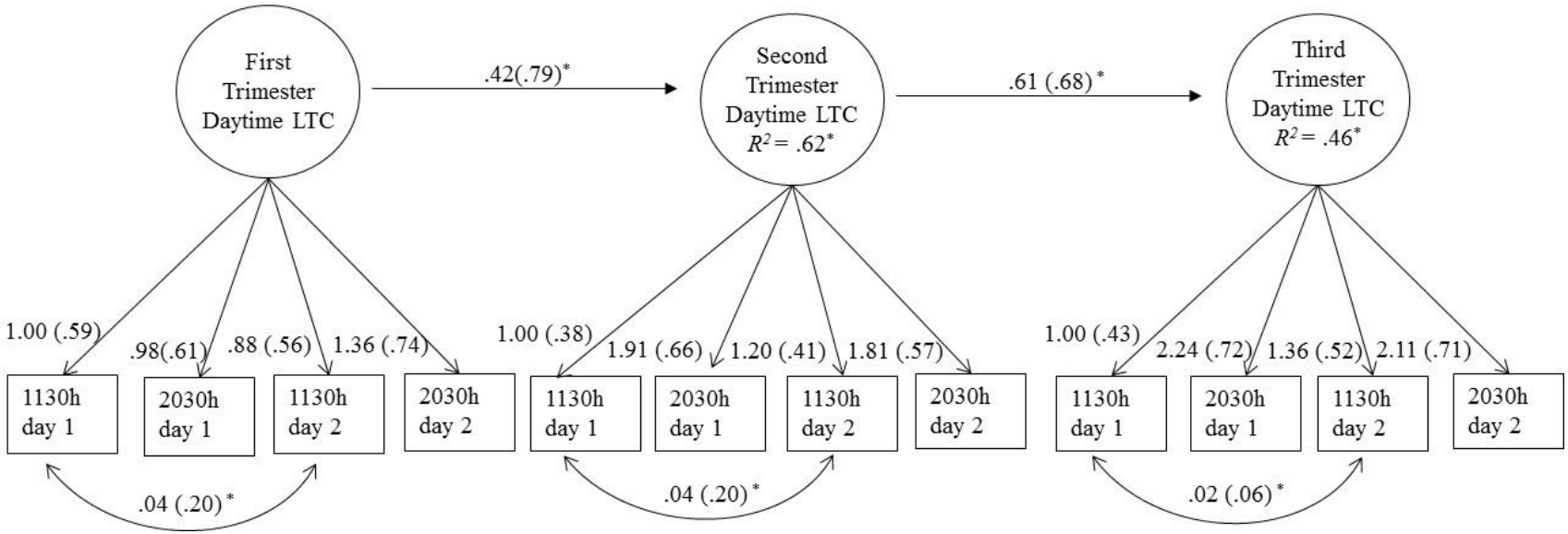

Figure 2. Predicting latent trait cortisol (LTC) across trimesters using daytime samples. Unstandardized estimates are followed by standardized estimates in parentheses, n = 372. All unstandardized factor loadings were significant. Time since waking for the 1130h samples was regressed on each corresponding 1130h sample, but is not depicted. * $p < .01$.