Psychoneuroendocrinology xxx (2018) xxx-xxx

Contents lists available at ScienceDirect

Psychoneuroendocrinology

journal homepage: www.elsevier.com

Stability of diurnal cortisol measures across days, weeks, and years during middle childhood and early adolescence: Exploring the role of age, pubertal development, and sex

Kate Ryan Kuhlman[a, b, *], Theodore F. Robles[b, c], Leah Dickenson[c], Bridget Reynolds[d], Rena L. Repetti[c]

[a] Department of Psychological Science, University of California Irvine, Irvine, CA 92697, United States
[b] Cousins Center for Psychoneuroimmunology, Semel Institute for Neuroscience & Human Behavior, University of California Los Angeles, Los Angeles, CA 90095, United States
[c] Department of Psychology, University of California Los Angeles, Los Angeles, 90095, United States
[d] National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD 20782, United States

ARTICLE INFO

ABSTRACT

Effective regulation of the hypothalamic-pituitary-adrenal axis (HPA-axis) has been linked to numerous health outcomes. Within-person variation in diurnal measures of HPA-axis regulation assessed over days, months, and years can range between 50–73% of total variation. In this study of 59 youth (ages 8–13), we quantified the stability of the cortisol awakening response (CAR), the diurnal slope, and tonic cortisol concentrations at waking and bedtime across 8 days (2 sets of 4 consecutive days separated by 3 weeks), 3 weeks, and 3 years. We then compared the stability of these indices across three key developmental factors: age, pubertal status, and sex. Youth provided 4 saliva samples per day (waking, 30 min post-waking, before dinner, and before bedtime) for 4 consecutive days during the 3rd week of an ongoing 8-week daily diary study. Youth repeated this same sampling procedure 3 weeks and 3 years later. Using multi-level modeling, we computed the amount of variance in diurnal HPA-axis regulation that was accounted for by nesting an individual's diurnal cortisol indices within days, weeks, or years. Across days, diurnal slope was the most stable index, whereas waking cortisol and CAR were the least stable. All indices except bedtime cortisol were similarly stable when measured across weeks, and all indices were uniformly stable when measured across 3 years. Boys, younger participants, and youth earlier in their pubertal development at study enrollment exhibited greater HPA-axis stability overall compared with females and older, more physically mature participants. We conclude that important within- and between-subjects questions can be answered about health and human development by studying HPA-axis regulation, and selection of the index of interest should be determined in part by its psychometric characteristics. To this end, we propose a decision tree to guide study design for research in pediatric samples by longitudinal timeframe and sample characteristics.

1. Introduction

Salivary indices of diurnal HPA-axis regulation have been used for the past two decades to better understand human development (Gunnar and Quevedo, 2007; Gunnar and Donzella, 2002) and the neurobiological underpinnings of health and disease (McEwen, 2013; Pariante and Lightman, 2008). More recently, HPA-axis indices have emerged as a useful indicator of intervention effectiveness in targeted high-risk groups (Slopen et al., 2014). These HPA-axis indices are most commonly measured through salivary cortisol, which is cost effective and amenable to collection in almost any setting (Saxbe, 2008). Yet, up to 50% of variability in diurnal indices of HPA-axis functioning can be attributed to day-to-day changes in mood, sleep, diet, sampling error, and other factors (Ross et al., 2014; Segerstrom et al., 2014). Some indices are more stable over time than others (e.g., diurnal slope appears to be more stable than CAR) (Doane et al., 2015; Shirtcliff et al., 2012) and are therefore more appropriate for between-subjects hypothesis testing while other indices may be better suited to within-subjects hypothesis testing. The purpose of this study was to determine the stability of these different indices of diurnal cortisol regulation across days, weeks, and years in a well-characterized, longitudinal sample of youth,

* Corresponding author at: Department of Psychological Science, University of California Irvine, 4201 Social & Behavioral Sciences Gateway, Irvine, CA, 92697, United States.
Email address: krkuhl@uci.edu (K.R. Kuhlman)

https://doi.org/10.1016/j.psyneuen.2018.09.033
Received 29 June 2018; Received in revised form 30 August 2018; Accepted 24 September 2018
Available online xxx
0306-4530/

and to examine the roles of age, pubertal status, and sex in the stability of these indices. The results have the potential to inform study design decisions for investigators interested in optimizing for within- or between-subject variability when integrating salivary measures of diurnal HPA-axis regulation in longitudinal studies of varying lengths of time.

Much research using diurnal HPA-axis measures to date has been conducted in the context of understanding the links between health and disease. To this end, investigators have focused largely on determining how well-established health-relevant risk factors are associated with differences in trait-like indices of HPA-axis regulation (Bower et al., 2005; Cohen et al., 2006; Doane et al., 2015, 2013; Kuhlman et al., 2015). Another line of research has pursued insight into within-person sensitivity to short-term changes in life stress, mood, and social interactions (Adam et al., 2006; Almeida et al., 2009; Bai et al., 2017; Kuhlman et al., 2018, 2016, 2016). Further, the importance of investigating the effects of psychosocial interventions on biological processes is promising (Rauch et al., 2017; Slopen et al., 2014) and growing (Zalta and Shankman, 2016). Integrating diurnal measures of HPA-axis functioning into intervention studies will necessitate sensitive measures of both within-subject and between-group differences. Thus, trait-like and state-like indices of diurnal HPA-axis functioning have value in our pursuit of the social, environmental, and psychological factors that influence health and development. Yet, the vast majority of studies operationalize diurnal HPA-axis functioning in the same way, regardless of the within- or between-subject nature of the research questions. This is, in part, due to lack of clarity about the appropriateness of different HPA-axis indices for capturing within- and between-subject effects.

There are several ways to measure functioning of the HPA-axis that offer unique insight into the daily regulation of the HPA-axis, with the vast majority of studies today measuring cortisol assayed from saliva. Cortisol concentrations are relatively high in the morning, exhibit a dramatic increase within one hour of waking, and decline throughout the day. While some studies have identified cortisol concentrations at specific times of day that may be relevant to health (Goodyer et al., 2000; Kirschbaum et al., 1990), indices computed from multiple samples in a day, across multiple days, are currently the standard in the field (Adam and Kumari, 2009). The most common of these computed indices are cortisol awakening response (CAR) and the diurnal slope. CAR is the magnitude of cortisol increase that occurs immediately upon waking, is thought to index adrenal sensitivity to adrenocorticotropic hormone (Clow et al., 2010), and appears to be sensitive to short-term changes in the social environment (Adam et al., 2006; Chida and Steptoe, 2009). Importantly, CAR has emerged as a reliable prospective predictor of depression (Adam et al., 2010; Kuhlman et al., 2017b; Vrshek-Schallhorn et al., 2013). The diurnal slope is used to capture the circadian down-regulation of cortisol from waking to bedtime. Individuals with different health conditions and poor prognoses also tend to lack a steady decline in cortisol throughout the day (Bower et al., 2005; Schrepf et al., 2015; Sephton et al., 2013, 2000). There are two common approaches to calculating diurnal slope. One approach regresses sampling time on all available cortisol measurements that day to yield an estimated slope of change across the day (See Bower et al., 2005; DeSantis et al., 2007; Doane et al., 2013; Ross et al., 2014 for examples using this method). However, daily experiences such as stress, exercise, and diet can contribute to variability in slopes that are distinct from the circadian rhythm. For this reason, another common approach to computing the diurnal slope is to subtract the cortisol concentration at waking from the cortisol concentration at bedtime, and divide that value by the number of hours between waking and bedtime (Hoyt et al., 2016; Kuhlman et al., 2017b, 2016). Each of these indices and their computation likely represent different underlying neurobiological processes, and therefore may be differentially sensitive to changes in an individual's life over time. The present study sought to determine the stability of different diurnal HPA-axis indices in order to provide clarity on which index can be used to optimally address within-subjects versus between-subjects research questions in pediatric samples while also balancing scientific ambitions with participant burden.

Several seminal papers have considered aspects of this issue, although they have largely focused on identifying the best trait measure of individual differences in diurnal HPA-axis regulation. Cortisol concentrations at waking and at 30 min after waking appear to be useful as trait-like indicators of individual differences in diurnal HPA-axis regulation (Doane et al., 2015; Kirschbaum et al., 1990). In a seminal study, 357 youth were followed from age 9 to age 15 to determine a trait measure of diurnal cortisol regulation that models all samples across the day simultaneously (Shirtcliff et al., 2012). The investigators determined that while 13% of the variation in cortisol levels was due to trait-like factors, the circadian rhythm of cortisol was 72% trait-like across the transition from childhood to adolescence (Shirtcliff et al., 2012). For comparison, in personality research, "trait-like" self-report measures demonstrate stability coefficients of .41 and higher, with stability lower in childhood than any other developmental phase and stability higher across shorter timeframes (e.g., days and weeks versus years) (Caspi et al., 2005; Roberts and DelVecchio, 2000). Additionally, Ross et al. (2014) compared the stability of multiple diurnal HPA-axis indices (CAR, total cortisol across the day (AUC), and diurnal slope) using data from three samples that ranged from childhood to middle adulthood with cortisol assessed multiple times for up to two years. They concluded that diurnal cortisol indices have some trait-like properties, but that diurnal HPA-axis indices are better indicators of short-term (day-to-day), rather than long-term (months and years), stress and health processes (Ross et al., 2014).

Shirtcliff et al. (2012) also showed that gender and development played important roles in the maturation of the HPA-axis (Shirtcliff et al., 2012). Specifically, females had higher cortisol and steeper circadian rhythms than males, and adolescents had lower cortisol concentrations in the morning and higher concentrations in the evening compared to concentrations when they were 9 years old (Shirtcliff et al., 2012). Importantly, these developmental changes were attributable to age rather than puberty (Shirtcliff et al., 2012).

Taken together, the studies cited here indicate that there are HPA-axis indices that can be used as trait measures of underlying neurobiological development and health over multiple years (e.g., circadian rhythm or diurnal slope) while others are more state-like (e.g., CAR) and better suited to assessing an individual's response to short-term changes in the environment across days and weeks. Yet, to our knowledge, no guidelines exist for selecting diurnal indices of HPA-axis regulation based on differences in measure stability. This paper approaches the question of stability by assessing the extent to which within-person variance contributes to total variability in waking cortisol, CAR, the diurnal slope, and bedtime cortisol in a sample of youth across three different timeframes: days, weeks, and years. Further, we will characterize how the different stability estimates vary according to three developmental factors: age, pubertal status, and sex.

2. Method

2.1. Participants

Participants in the present study were part of a study of the daily lives of children (8–13 years) and their families across an initial 8 week daily diary assessment period and a 3 year follow-up assessment (Reynolds et al., 2015; Robles et al., 2013). Participants were recruited from public schools, pediatric clinics, community centers, newspaper advertisements, and direct mailings between 2009 and 2012. A total of 59 youth enrolled in the study (47 target children and 12 siblings). Families were eligible for the study if they were currently living in two-

parent households, with at least one biological parent willing to participate. Participants were excluded for medical conditions or use of medications known to confound endocrine functioning, including chronic lung conditions (e.g., asthma), endocrine disorders (e.g., Cushing disease), metabolic disease, immunodeficiency, and heart disease or chronic heart conditions. Median personal income reported by parents in this study was within the $31,850 to $82,400 tax bracket, and 57% of mothers (59% of fathers) attained at least a bachelor's degree or higher. Youth were 46.8% non-Hispanic white, 23.4% Latino/ Hispanic, 23.4% African-American, 4.3% Asian, and 2.1% "Other". Of the initial 59 youth, 67.8% ($n = 40$) participated in the follow-up study. There were no significant differences in pubertal status, $F(1,57) = 0.15$, $p = .70$, age, $F(1,57) = 0.57$, $p = .45$, sex, $\chi^2 = 1.43$, $p = .23$, or ethnic group, $\chi^2 = 1.88$, $p = .76$, between youth who did and did not return for the follow-up study.

2.2. Procedures

This study was approved by the UCLA Institutional Review Board; parents provided written consent to study procedures, and youth provided assent. For 8 weeks, participating youth provided daily diary reports of physical symptoms, social interactions, mood, and other psychosocial factors. During the 3rd and 6th week of this diary phase, participants provided saliva samples. Three years after the 8-week daily diary phase, youth were invited to participate in a follow-up study. See Fig. 1 for data collection procedures.

2.3. Measures

2.3.1. Diurnal HPA-axis functioning

During a home visit a research assistant provided training on saliva sample collection procedures. Participants were given twist-cap bottles containing 4-inch straws to aid the passive acquisition of drool. Participants were instructed to record the time of each sample in a paper diary using a programmed electronic time stamper. All participants were informed that their compliance with saliva collection instructions would be monitored with electronic monitoring devices, and MEMS caps were used for a subset of participants (29%) to assess adherence to sampling instructions. Participants were instructed to store completed saliva samples in the freezer until retrieved by study staff and stored at −20 °C in the lab.

Saliva samples were collected by participants via passive drool four times per day: upon waking, 30 min after waking, before dinner, and at bedtime. Saliva collection occurred for 4 consecutive days in each sampling period on Saturdays through Tuesdays. These sampling phases occurred in week 3 and 6 of the 8 week daily diary phase, and during the 3-year follow-up assessment. Thus, each youth had the opportunity to provide 32 samples if they did not participate in the follow-up study and 48 samples if they did. Compliance with saliva sample completion was high. Of 2,528 total possible saliva samples, 87.2% (2,204 samples) were returned for analysis. On average, participants provided 42.56 samples (SD = 7.50) and the minimum number of samples any participant provided was 30. Fidelity to the sample timing instructions was defined as < 30 min difference between the electronic time stamp on the paper recording form and the MEMS cap time record. Fidelity for this sample was 97.6%.

Saliva samples were frozen after collection and shipped on dry ice to the TUD Biopsychology Laboratory in Dresden, Germany directed by Dr. Clemens Kirschbaum. Cortisol concentrations were measured using a commercially available enzyme immunoassay. All samples from the same person were assayed on the same plate (50 µl saliva required; minimum detection limit <.003 µg/dL, mean intra- and interassay coefficients of variance (CV) below 10%).

2.3.2. Age, pubertal status, and sex

Mothers provided participant age and sex as part of the baseline enrollment questionnaires. Mothers completed the Pubertal Development Scale (PDS) (Petersen et al., 1988) in terms of growth spurt, skin changes, and facial/body hair growth. For sons, mothers reported on voice changes. For daughters, mothers reported on breast growth and menstruation. Responses on the 4-point scale (1 = "has not yet begun," 2 = "has barely begun," 3 = "has definitely begun," 4 = "has been completed") were averaged. The PDS is a valid and reliable measure that correlates with physician reports of pubertal development (Dorn and Biro, 2011). Reliability for the mother-reported PDS within our sample was good, $\alpha = .81$.

2.4. Data analysis

All analyses were conducted using SPSS version 24. All variables were examined for normality and heteroscedasticity. There were 2,204 salivary cortisol samples included in these analyses, $M_{cortisol} = 11.60$, $SD_{cortisol} = 14.34$, in a distribution that was skewed and kurtotic, skewness = 6.76, kurtosis = 120.38. There were 9 samples with extreme high values, 6 of which were the 30-minute post-waking sample. We winsorized these to 4 standard deviations from the mean which brought skew and kurtosis within acceptable ranges, skewness = 1.67, kurtosis = 3.56. The mean cortisol concentration across these 2,204 samples after winsorizing the 9 extreme values was 11.39 nmol/L, $SD = 11.99$. All saliva sampling times were recoded in reference to the participant's reported waking time on the sampling day. On days where a waking sample was not taken ($n = 31$ days across all participants and days), the time of waking was imputed from that participant's average waking time on sampling days within that sampling week. Hours since waking was then computed by subtracting the waking sample time from any given sample time on the same day.

Using linear mixed models, we examined predictors of cortisol concentrations as a function of time since waking in 5 models with outcomes that reflected 1) cortisol at waking, 2) CAR, 3) the diurnal slope using waking, afternoon, and bedtime samples, 4) the diurnal slope using only waking and bedtime samples, and 5) cortisol concentrations at bedtime. In an effort to hold the total variance across models somewhat consistent, each model included 8 saliva sampling days. Cortisol at waking was modeled using only waking samples as a function of the time of waking. CAR was estimated as the linear change in cortisol from waking to 30 min post-waking as a function of the minutes between waking and the second sample. This two-sample CAR assessment is consistent with the expert consensus guidelines in pediatric popula-

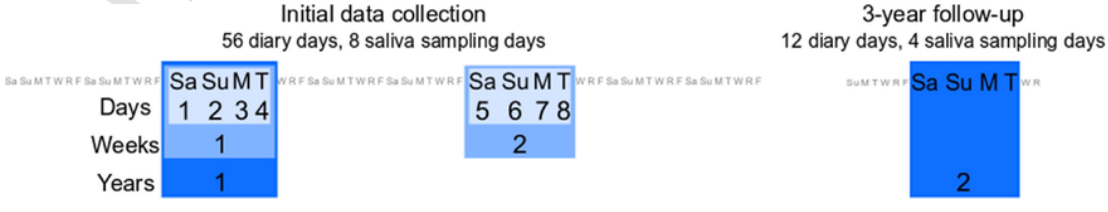

Fig. 1. Study procedure timeline and days included in models assessing stability over days, weeks, and years.

tions (Stalder et al., 2016). Diurnal slope using three samples (diurnal slope 3) was estimated as the linear change in cortisol concentration across the day as a function of the time each sample was provided (waking, afternoon, and bedtime) and waking that day. Diurnal slope using only two samples was estimated using the linear change in cortisol concentrations across the day as a function of the hours between waking and the bedtime sample (diurnal slope 2). Cortisol at bedtime was modeled using only bedtime samples as a function of the time of the bedtime sample.

Models of each of these indices over days, weeks, and years used data from 8 sampling days. For models used to estimate stability over days, cortisol concentrations obtained on the 1st 8 sampling days were nested within days within participants (See Fig. 1 for sampling procedures). For models used to estimate stability over weeks, cortisol concentrations obtained on the 1st 8 sampling days were nested within weeks within participants (4 days in one week and 4 days in a second week, three weeks later). Models estimating stability across years used data from the subsample of participants who returned for the follow-up study. Their cortisol values on the 1st 4 days in the initial data collection and on the 4 sampling days at the follow-up (three years later) were nested within years within participants.

We then ran each of the above models stratified by age (mean split), sex, and mother-reported pubertal status at study enrollment (mean split) to determine the role of each of these factors in the stability of each diurnal cortisol index.

The purpose of our analyses was to estimate the stability of each index of HPA-axis functioning across different epochs of time (i.e., days, weeks, years). We computed intra-class coefficients (ICCs) by dividing the variance remaining after nesting samples within epochs and individuals ($\sigma_{Between\ +\ \varepsilon}$) by the total variance in the model ($\sigma_{Between+\varepsilon}$ + $\sigma_{Within\ +\ \varepsilon}$):

$$ICC = \frac{\sigma_{Between+\varepsilon}}{\sigma_{Between+\varepsilon} + \sigma_{Within+\varepsilon}}$$

The resulting value can range from 0 to 1, with a lower ratio indicating that a greater proportion of the total variance is due to variability within individuals. Thus, for any cortisol index, a lower ICC indicates greater instability, and a higher ICC indicates more stability, across a particular timeframe (days, weeks, or years). We then computed 95% confidence intervals (95%CI) for each ICC in order to determine whether the stability of two models reliably differ from one another using the prop.test approach in R (Newcombe, 1998a, 1998b; Wilson, 1927). Confidence intervals were computed according to the following formula such that $\hat{\rho}$ = ICC, $z = 1.96$, and $n = \sigma_{\text{Between}\ +\ \varepsilon}$ + $\sigma_{\text{Within}\ +\ \varepsilon}$:

$$95\%CI = \hat{\rho} \pm z \bullet \sqrt{\frac{\hat{\rho}\left(1-\hat{\rho}\right)}{n}}$$

If the confidence intervals for two model ICCs do not overlap, it can be inferred with 95% confidence that they are reliably different.

3. Results

Diurnal cortisol exhibited a typical regulatory profile across all samples in this study. Individuals had more elevated cortisol in the morning at waking than at bedtime and demonstrated a significant increase in cortisol in the 30 min after waking, $b = 3.30$, $SE = 0.90$, $p < .001$. See Table 1 for sample characteristics and each diurnal cortisol index at the initial wave of data collection and the 3-year follow-up.

Table 1
Sample characteristics and diurnal cortisol indices at wave 1 ($n = 59$) and at 3-year follow-up ($n = 40$).

	M (SE)	% (*n*)
Demographic characteristics		
Age at enrollment (range 8–13)	11.02 (0.20)	
Female		59.3 (35)
Pubertal status at enrollment (range 1–4)	2.09 (0.11)	
Diurnal cortisol indices in initial sample		
Waking cortisol (nmol/L)	16.51 (0.77)	
Cortisol awakening response (Δ nmol/L per hour)	1.49 (1.05)	
Diurnal slope (using waking, afternoon, and bedtime samples) (Δ nmol/L per hour)	−1.04 (0.03)	
Diurnal slope (using only waking and bedtime samples) (Δ nmol/L per hour)	−1.00 (0.04)	
Bedtime cortisol (nmol/L)	2.53 (0.29)	
Diurnal cortisol indices (nmol/L) at 3-year follow-up		
Waking cortisol (nmol/L)	22.23 (1.27)	
Cortisol awakening response (Δ nmol/L per hour)	6.05 (2.60)	
Diurnal slope (using waking, afternoon, and bedtime samples) (Δ nmol/L per hour)	−1.27 (0.07)	
Diurnal slope (using only waking and bedtime samples) (Δ nmol/L per hour)	−1.24 (0.08)	
Bedtime cortisol (nmol/L)	4.03 (0.31)	

3.1. Stability of diurnal cortisol indices across days

Fig. 2 illustrates the stability estimates (ICCs) and *95%CI* for each diurnal cortisol index over days, weeks, and years. Table S1 presents the variance components and *95%CI* for each model. Daily stability in the diurnal cortisol indices ranged from ICC = .46 to .77. Within-subject variability accounted for approximately 50% of the total variance in cortisol at waking, bedtime, and CAR, but less than 30% of the total variance in both approaches to computing diurnal slope. Indeed, both of the diurnal slope estimates were significantly more stable than waking, CAR, or bedtime cortisol. Thus day-level variance in diurnal HPA-axis function may be more detectable in waking, bedtime, and CAR indices; diurnal slope may be less sensitive to factors that change from day-to-day.

Female participants exhibited higher day-to-day stability of the 3 sample diurnal slope compared to males. There were no significant differences in day-to-day stability of any of our HPA-axis indices by age, or pubertal status. See Table 2 for ICCs by pubertal status, sex, and age across days. See Table S2-4 for the variance components and *95%CI* for each model by age, pubertal status, and sex.

3.2. Stability of diurnal cortisol indices across weeks

When looking at variance at the week level (saliva collected during one week and then another 3 weeks later), the range in stability was almost identical to the range at the day level: from .46 (bedtime cortisol) to .76 (diurnal slope computed with 2 samples). Within-subject variability across weeks accounted for less than 50% of the total variance in bedtime cortisol, but less than 1/3 of the total variance in waking cortisol, both approaches to diurnal slope, and CAR. Bedtime cortisol was the most sensitive index to week-to-week variability, whereas the other HPA indices were relatively stable. However, the only statistically significant difference in stability was between diurnal slope (both approaches) and bedtime cortisol.

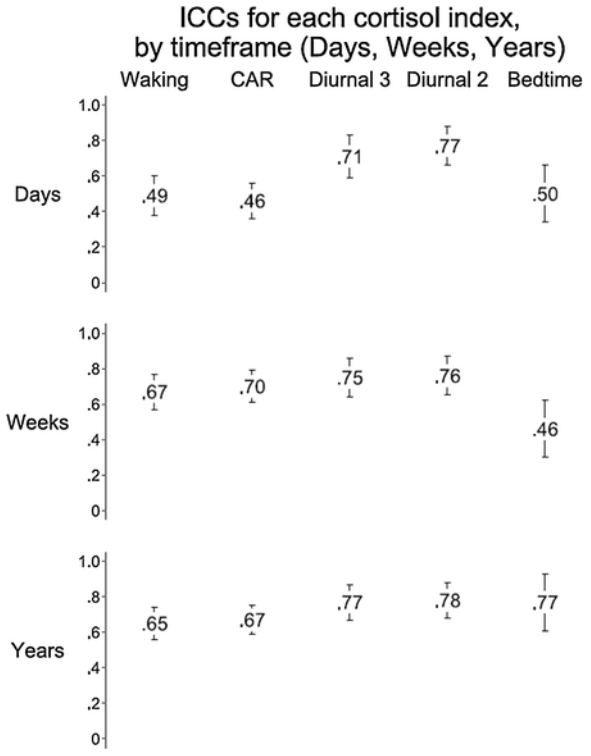

Fig. 2. Intraclass correlation coefficients (ICC) of each cortisol index by timeframe (days, weeks, years).Note: Diurnal 3 = diurnal cortisol slope computed using waking, afternoon, and bedtime samples. Diurnal 2 = diurnal cortisol slope computed using waking and bedtime saliva samples.

The week-level stability of the diurnal cortisol indices (measured over the span of 3 weeks using 8 sampling days) varied significantly by participant pubertal status and sex, but not age. See Table 2 for ICCs by pubertal status, sex, and age across weeks. Specifically, female participants exhibited less stable waking cortisol and CAR than male participants. The effect of more advanced pubertal development paralleled the effect of being female. Participants who, at study enrollment, were ahead of their peers' pubertal development exhibited significantly less stable waking cortisol and CAR than did participants who were still pre-pubertal at study enrollment. There were no significant differences in stability of diurnal slope or bedtime cortisol by age, pubertal status, or sex.

3.3. Stability of diurnal cortisol indices across years

Models estimating stability across years were based on data from the subsample of participants who returned for the Time 2 follow-up study. Their cortisol values on the 1st 4 days at Time 1 and on the 4 sampling days at Time 2 (three years later) were nested within years within participants. At the year level, diurnal cortisol indices ranged in stability (ICCs) from .65 (waking cortisol) to .78 (diurnal slope 2). Notably, less than 40% of variance in all of the diurnal HPA-axis indices was accounted for by within-subject variability across years, and no indices were significantly more or less stable than the others. Thus, over the span of years, diurnal HPA-axis regulation varies more between individuals than within individuals, even across early adolescence.

The stability of these diurnal cortisol indices measured over the span of 3 years using 8 sampling days varied significantly by participant pubertal status, age, and sex. See Table 2. Specifically, participants older than 11 at study enrollment exhibited significantly less stable waking and CAR values than participants age 11 or younger at enrollment. Similarly, participants whose mothers reported that their child was further along in pubertal development at study enrollment exhibited significantly less stable waking cortisol and CAR across years than peers reported to be earlier in pubertal development. Finally, male participants exhibited more stable waking cortisol and CAR values than female participants. There were no significant differences in the stability of diurnal slope or bedtime cortisol by any of the 3 grouping variables we examined.

4. Discussion

Consistent with previous studies (Platje et al., 2013; Ross et al., 2014; Shirtcliff et al., 2012), diurnal indices of HPA-axis functioning in this sample exhibited a high degree of variability within individuals over time. Overall, diurnal cortisol indices appeared to be more consistently stable as sampling took place over longer periods of time, suggesting that the greatest source of within-subject variability is day-to-day, not necessarily across longer phases of important human development. Diurnal slope, independent of how it was calculated, was the most stable index across all timeframes and therefore may be useful when understanding how stable individual differences in HPA-axis functioning relate to health, development, and behavior. In comparison, over short periods of time CAR, waking, and bedtime cortisol were most variable within participants, and therefore may be more useful in studies investigating the impact of recent changes in the environment or behavior on functioning of the HPA-axis. Further, these stability indices varied meaningfully by age, pubertal status, and sex at the time of study enrollment.

Based on our results, the most commonly computed indices of HPA-axis functioning are quite vulnerable to instability within individuals across time. Thus, studies testing between-subjects research questions using these indices may be more susceptible to type I error, such as a study exploring differences in diurnal HPA-axis regulation between individuals exposed to high or low daily stress. Indeed, the well-documented day-to-day variability in HPA-axis regulation as measured by salivary cortisol has inspired a great deal of discourse on this possibility (Boggero et al., 2017; Segerstrom et al., 2014). As a result, the recommended sample size and number of sampling days for studies of diurnal cortisol functioning continue to grow. Unfortunately, this undermines the value of salivary HPA-axis measures as cost-effective and feasible across research, clinical, and community settings. Instead, we argue that a more nuanced approach can be taken to selecting diurnal HPA-axis measures. For example, the most stable measures of HPA-axis function are ideal for testing questions of how individual differences in the tonic functioning of this biological stress response system predict health outcomes or moderate treatment responses.

Fig. 3 depicts a decision tree that aims to guide researchers to useful diurnal HPA-axis indices based on their research questions, study design, and sample characteristics. To illustrate, a randomized controlled trial (RCT) of a 12-week intervention may effectively use diurnal slope, a relatively trait-like measure of HPA-axis functioning, to determine the effect of their intervention on trait functioning of a key neurobiological system linking stress to health. However, due to their stability, these indices may not be sensitive enough to detect the influence of short-term, daily changes in the social or physical environment that have been so crucial to our growing understanding of biopsychosocial processes. That same RCT, may also look at within-person changes in bedtime cortisol from pre- to post-intervention to detect whether the intervention influenced a more sensitive and relatively state-like index of diurnal cortisol regulation. Relatively unstable measures, in comparison, demonstrate enough variability to invite the interrogation of other influences on HPA-axis functioning, such as fluctuations in health and health behaviors, changes in social relationships, systematic measurement error, and random error over short timeframes. Given that some diurnal HPA-axis indices, namely CAR and diurnal slope, have been prospectively linked to important health out-

Table 2
Intraclass correlation coefficients (ICC) of diurnal cortisol indices across days, weeks, and years by age, pubertal status, and sex.

	ICC									
	Total	Age			Pubertal Status			Sex		
		≤ 11	> 11	$p <$ .05.	≤ 2	> 2	$p <$ .05.	Male	Female	$p <$ .05
Days										
Waking	0.49	0.50	0.50		0.50	0.50		0.50	0.50	
CAR	0.46	0.49	0.38		0.58	0.40		0.57	0.44	
Diurnal slope (3)	0.71	0.66	0.75		0.68	0.74		0.60	0.77	*
Diurnal slope (2)	0.77	0.77	0.76		0.79	0.77		0.71	0.80	
Bedtime	0.50	0.50	0.50		0.50	0.50		0.50	0.50	
Weeks										
Waking	0.67	0.86	0.58		0.75	0.59	*	0.89	0.61	*
CAR	0.70	0.85	0.66		0.81	0.67	*	0.86	0.67	*
Diurnal slope (3)	0.75	0.77	0.72		0.78	0.72		0.76	0.73	
Diurnal slope (2)	0.76	0.79	0.73		0.79	0.74		0.76	0.76	
Bedtime	0.46	0.64	0.68		0.57	0.81		0.58	0.80	
Years										
Waking	0.65	0.79	0.61	*	0.78	0.57	*	0.76	0.61	*
CAR	0.67	0.84	0.64	*	0.81	0.60	*	0.77	0.63	*
Diurnal slope (3)	0.77	0.80	0.76		0.82	0.74		0.79	0.76	
Diurnal slope (2)	0.78	0.88	0.75		0.87	0.74		0.83	0.76	
Bedtime	0.77	0.84	0.59		0.73	0.82		0.74	0.77	

Note: Darker shading indicates higher stability. Diurnal slope 3 = diurnal cortisol slope computed using waking, afternoon, and bedtime samples. Diurnal slope 2 = diurnal cortisol slope computed using waking and bedtime saliva samples.

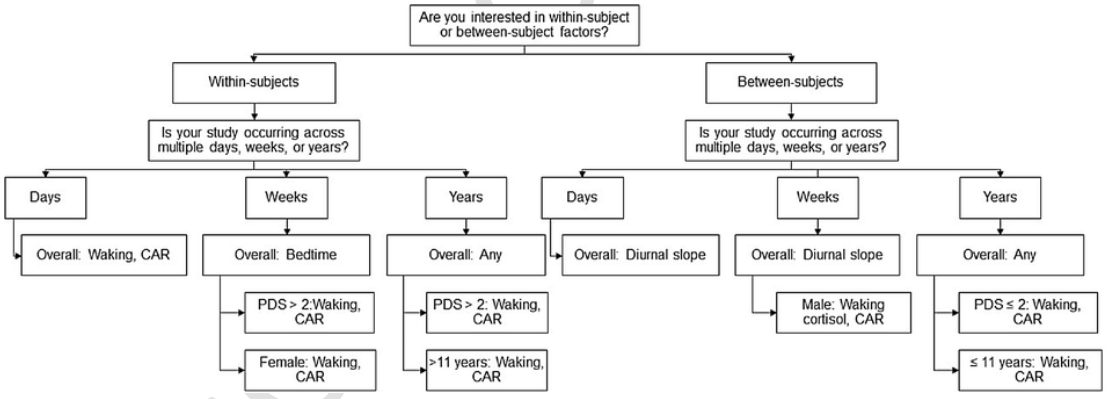

Fig. 3. Summary decision tree for selecting diurnal HPA-axis indices for youth aged 8–13 based on study design and sample characteristics.

comes like depression (Adam et al., 2010; Kuhlman et al., 2017b; Vrshek-Schallhorn et al., 2013) and mortality in cancer patients (Schrepf et al., 2015; Sephton et al., 2013, 2000), adoption of intensive repeated measures approaches to better understand predictors of their within-subject variability are particularly important. It should be noted that this decision tree is preliminary and warrants validation and further iterations with data from samples that vary in age, demographics, health status, as well as studies that vary in experimental design. Further, more research is needed to understand how highly variable indices contribute to changes in more stable indices of HPA-axis regulation over time.

A strength of this study was the age of the sample (ages 8–13 years) and timing of the 3 year follow-up that co-occurred with important aspects of social and physical development. There were different psychometrically optimal HPA-axis indices depending on sex, age (younger or older than 11 years), and pubertal status (less than or equal to a score of 2 on the mother-reported PDS). Most notably, cortisol concentrations in the morning were more stable across weeks and years in younger participants, boys, and those who were earlier in pubertal development at study enrollment. Age was only associated with differences in diurnal cortisol stability when sampling occasions were separated by years, such that participants 11 or younger exhibited more stable waking cortisol and CAR. The ages 11–12 represent a common transition to middle school in the United States. This transition involves an often stressful shift in school experiences to that of multiple teachers and classrooms throughout the day, and a growing social network of same-aged peers (Eccles et al., 1993). Thus, instability of morning cortisol over longer periods of assessment may reflect devel-

opmental changes in exposure to these types of social stressors. With respect to pubertal development, a mean-split at a score of 2 on the PDS reflects one group that is homogenous for the physical changes associated with puberty "barely" beginning, and one group that is homogenous for the physical changes associated with puberty that have "definitely" begun or have been completed. We found that across weeks and years, youth who began in the earlier stages of pubertal development exhibited more stable waking and CAR indices. Pubertal development is a time of neural and neuroendocrine reorganization (Kuhlman et al., 2017a; Romeo, 2010; Romeo et al., 2014, 2006). Thus instability in the more advanced group of participants may reflect changes in the functioning of the axis that occur during and after puberty. Of course, age and pubertal development are difficult to disentangle in this age group. Before age 11, only 10% of girls have reached menarche (Chumlea et al., 2003), thus we expect that the younger age group is fairly homogenous in that their pubertal development has not yet begun or is barely underway. One previous study showed similar patterns of instability that were driven largely by age, rather than pubertal status (Shirtcliff et al., 2012). However, given that differences in stability emerged across weeks and years when split by pubertal timing, and only emerged across years for age, pubertal status measures may be a more sensitive measure of these developmental changes compared to age, regardless of their cause. Finally, we observed that girls had more stable diurnal slopes across days, but that boys exhibited more stable waking and CAR indices across weeks and years. More than anything, these observations underscore the importance of either controlling for sex or stratifying analyses by sex for these indices in similarly aged samples. Of particular note, the developmentally-based group differences in HPA-axis regulation over weeks and years were limited to two indices: waking and CAR. This raises the possibility that the instability of diurnal HPA-axis functioning in these groups is a consequence of sleep or the well-established barriers to sleep that occur in adolescence (Carskadon, 2011; Hagenauer and Lee, 2013).

The results of this study should be interpreted in the context of its strengths and limitations. First, this was a study of 2,204 salivary cortisol samples taken from a relatively small sample of youth (ages 8–13 at enrollment) 3 weeks, 6 weeks, and 3 years into a longitudinal study of diverse, middle-class, urban-dwelling families. Thus, these results of this study may not generalize to samples in different phases of the human lifespan or living under different social or physical conditions. Further, only 40 youth completed the 3 year longitudinal follow-up and therefore results with respect to stability across years may be less reliable than those across days and weeks. With respect to our results related to age, pubertal status, and sex, it is important to note that we did not test these factors as moderators, but rather stability of HPA-axis indices in models homogenous for these factors. These findings should therefore be considered hypothesis-generating, rather than conclusive. Our approach in the present analyses and their subsequent interpretation inherently prioritized stability of an index over the unique underlying neurobiological processes it reflects. Certainly, in any investigation the theoretical interest in a particular aspect of HPA-axis functioning must be balanced against the psychometric characteristics of the HPA-axis index for that study design or sample. Both considerations are important and our analysis is relevant to the latter. In particular, quantifying the estimated stability of an HPA-axis index of interest can inform power analyses and sample size decisions that minimize the risk of both type I and type II error. Finally, the lack of predictors in our models leaves us without guidance on how much of the between- vs. within-subject variance in each model was due to specific sources of error, such as variability in participant compliance, relative to random, unidentifiable sources of error variance. This is particularly relevant to samples collected in relation to waking which exhibited relatively low day-to-day stability and are highly susceptible to error due to the potential delay between actual waking and sample collection (Stalder et al., 2016). As these findings begin to be tested in different studies, attention to the error variance and contributors of error variance for each diurnal cortisol index over time will be critical to our understanding and the methodological rigor of our field.

A number of recommendations exist to guide researchers assessing diurnal HPA-axis functioning. These recommendations often include collection of more samples within a day, sampling across more consecutive days, and testing questions in larger sample sizes (Boggero et al., 2017; Segerstrom et al., 2014; Stalder et al., 2016). Yet, careful interrogation of sampling procedures show that more frequent sampling is not always necessary (Hoyt et al., 2016), particularly when investigating factors that contribute to the within-person variability of HPA-axis regulation. Our paper provides researchers interested in diurnal HPA-axis regulation in youth with additional guidance regarding the maximization of within- or between-subject variability over different timeframes.

Role of the funding source

The submitted research was financially support by the William T. Grant Foundation (Research Grant 9333) and the NIMH (K08MH112773) who were not involved in study design, the collection, analysis and interpretation of data, writing of the manuscript, or the decision to submit the article for publication.

Acknowledgements

This research was supported by a Research Grant (9333) from the William T. Grant Foundation. Preparation of this article was supported by the National Institute of Mental Health (K08MH112773) awarded Dr. Kate R. Kuhlman. This research would not have been possible without the parents and children for their participation in this project and dedication to scientific progress, and we would like to thank the graduate students, laboratory staff, and research assistants on the project for their time and effort working with families.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.psyneuen.2018.09.033.

References

Adam, E.K., Kumari, M., 2009. Assessing salivary cortisol in large-scale, epidemiological research. Psychoneuroendocrinology 34, 1423–1436. https://doi.org/10.1016/j.psyneuen.2009.06.011.

Adam, E.K., Hawkley, L.C., Kudielka, B.M., Cacioppo, J.T., 2006. Day-to-day dynamics of experience–cortisol associations in a population-based sample of older adults. Proc. Natl. Acad. Sci. 103, 17058–17063. https://doi.org/10.1073/pnas.0605053103.

Adam, E.K., Doane, L.D., Zinbarg, R.E., Mineka, S., Craske, M.G., Griffith, J.W., 2010. Prospective prediction of major depressive disorder from cortisol awakening responses in adolescence. Psychoneuroendocrinology 35, 921–931. https://doi.org/10.1016/j.psyneuen.2009.12.007.

Almeida, D.M., Piazza, J.R., Stawski, R.S., 2009. Interindividual differences and intraindividual variability in the cortisol awakening response: an examination of age and gender. Psychol. Aging 24, 819–827.

Bai, S., Robles, T.F., Reynolds, B.M., Repetti, R.L., 2017. Children's diurnal cortisol responses to negative events at school and home. Psychoneuroendocrinology 83, 150–158. https://doi.org/10.1016/j.psyneuen.2017.05.027.

Boggero, I.A., Hostinar, C.E., Haak, E.A., Murphy, M.L.M., Segerstrom, S.C., 2017. Psychosocial functioning and the cortisol awakening response: meta-analysis, P-curve analysis, and evaluation of the evidential value in existing studies. Biol. Psychol. 129, 207–230. https://doi.org/10.1016/j.biopsycho.2017.08.058.

Bower, J.E., Ganz, P.A., Dickerson, S.S., Petersen, L., Aziz, N., Fahey, J.L., 2005. Diurnal cortisol rhythm and fatigue in breast cancer survivors. Psychoneuroendocrinology 30, 92–100. https://doi.org/10.1016/j.psyneuen.2004.06.003.

Carskadon, M.A., 2011. Sleep in adolescents: the perfect storm. Pediatr. Clin. North Am. 58, 637–647. https://doi.org/10.1016/j.pcl.2011.03.003.

Caspi, A., Roberts, B.W., Shiner, R.L., 2005. Personality development: stability and change. Annu. Rev. Psychol. 56, 453.

Chida, Y., Steptoe, A., 2009. Cortisol awakening response and psychosocial factors: a systematic review and meta-analysis. Biol. Psychol. 80, 265–278. https://doi.org/10.1016/j.biopsycho.2008.10.004.

Chumlea, W.C., Schubert, C.M., Roche, A.F., Kulin, H.E., Lee, P.A., Himes, J.H., Sun, S.S., 2003. Age at menarche and racial comparisons in US girls. Pediatrics 111, 110–113. https://doi.org/10.1542/peds.111.1.110.

Clow, A., Hucklebridge, F., Stalder, T., Evans, P., Thorn, L., 2010. The cortisol awakening response: more than a measure of HPA axis function. Neurosci. Biobehav. Rev. 35, 97–103. https://doi.org/10.1016/j.neubiorev.2009.12.011.

Cohen, S., Schwartz, J.E., Epel, E., Kirschbaum, C., Sidney, S., Seeman, T., 2006. Socioeconomic status, race, and diurnal cortisol decline in the coronary artery risk development in young adults (CARDIA) study. Psychosom. Med. 68, 41–50. https://doi.org/10.1097/01.psy.0000195967.51768.ea.

DeSantis, A.S., Adam, E.K., Doane, L.D., Mineka, S., Zinbarg, R.E., Craske, M.G., 2007. Racial/ethnic differences in cortisol diurnal rhythms in a community sample of adolescents. J. Adolesc. Health 41, 3–13. https://doi.org/10.1016/j.jadohealth.2007.03.006.

Doane, L.D., Mineka, S., Zinbarg, R.E., Craske, M., Griffith, J.W., Adam, E.K., 2013. Are flatter diurnal cortisol rhythms associated with major depression and anxiety disorders in late adolescence? The role of life stress and daily negative emotion. Dev. Psychopathol. 25, 629–642. https://doi.org/10.1017/S0954579413000060.

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. https://doi.org/10.1016/j.psyneuen.2015.01.017.

Dorn, L.D., Biro, F.M., 2011. Puberty and its measurement: a decade in review. J. Res. Adolesc. 21, 180–195. https://doi.org/10.1111/j.1532-7795.2010.00722.x.

Eccles, J.S., Midgley, C., Wigfield, A., Buchanan, C.M., Reuman, D., Flanagan, C., Iver, D.M., 1993. Development during adolescence. The impact of stage-environment fit on young adolescents' experiences in schools and in families. Am. Psychol. 48, 90.

Goodyer, I.M., Herbert, J., Tamplin, A., Altham, P.M.E., 2000. Recent life events, cortisol, dehydroepiandrosterone and the onset of major depression in high-risk adolescents. Br. J. Psychiatry 177, 499–504. https://doi.org/10.1192/bjp.177.6.499.

Gunnar, M.R., Donzella, B., 2002. Social regulation of the cortisol levels in early human development. Psychoneuroendocrinology 27, 199–220. https://doi.org/10.1016/S0306-4530(01)00045-2.

Gunnar, M., Quevedo, K., 2007. The neurobiology of stress and development. Annu. Rev. Psychol. 58, 145–173. https://doi.org/10.1146/annurev.psych.58.110405.085605.

Hagenauer, M.H., Lee, T.M., 2013. Adolescent sleep patterns in humans and laboratory animals. Horm. Behav. 64, 270–279. https://doi.org/10.1016/j.yhbeh.2013.01.013.

Hoyt, L.T., Ehrlich, K.B., Cham, H., Adam, E.K., 2016. Balancing scientific accuracy and participant burden: testing the impact of sampling intensity on diurnal cortisol indices. Stress Amst. Neth. 19, 476.

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.

Kuhlman et al., 2015 K.R. Kuhlman, E.G. Geiss, I. Vargas, N.L. Lopez-Duran, Differential associations between childhood trauma subtypes and adolescent HPA-axis functioning, Psychoneuroendocrinology 54 (2015) 103–114, https://doi.org/10.1016/j.psyneuen.2015.01.020.

Kuhlman, K.R., Repetti, R.L., Reynolds, B.M., Robles, T.F., 2016. Change in parent-child conflict and the HPA-axis: Where should we be looking and for how long?. Psychoneuroendocrinology 68, 74–81. https://doi.org/10.1016/j.psyneuen.2016.02.029.

Kuhlman, K.R., Chiang, J.J., Horn, S., Bower, J.E., 2017. Developmental psychoneuroendocrine and psychoneuroimmune pathways from childhood adversity to disease. Neurosci. Biobehav. Rev. 80, 166–184. https://doi.org/10.1016/j.neubiorev.2017.05.020.

Kuhlman, K.R., Irwin, M.R., Ganz, P.A., Crespi, C.M., Petersen, L., Asher, A., Bower, J.E., 2017. Cortisol awakening response as a prospective risk factor for depressive symptoms in women after treatment for breast cancer. Psychosom. Med. 79, 763–769. https://doi.org/10.1097/PSY.0000000000000499.

Kuhlman, K.R., Repetti, R.L., Reynolds, B.M., Robles, T.F., 2018. Interparental conflict and child HPA-axis responses to acute stress: insights using intensive repeated measures. J. Fam. Psychol. https://doi.org/10.1037/fam0000437.

McEwen, B.S., 2013. The brain on stress toward an integrative approach to brain, body, and behavior. Perspect. Psychol. Sci. 8, 673–675. https://doi.org/10.1177/1745691613506907.

Newcombe, R.G., 1998. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat. Med. 17, 857.

Newcombe, R.G., 1998. Interval estimation for the difference between independent proportions: comparison of eleven methods. Stat. Med. 17, 873.

Pariante, C.M., Lightman, S.L., 2008. The HPA axis in major depression: classical theories and new developments. Trends Neurosci. 31, 464–468.

Petersen, A.C., Crockett, L., Richards, M., Boxer, A., 1988. A self-report measure of pubertal status: reliability, validity, and initial norms. J. Youth Adolesc. 17, 117–133.

Platje, E., Vermeiren, R.R., Branje, S.J., Doreleijers, T.A., Meeus, W.H., Koot, H.M., Frijns, T., van Lier, P.A., Jansen, L.M., 2013. Long-term stability of the cortisol awakening response over adolescence. Psychoneuroendocrinology 38, 271–280.

Rauch, S.A.M., King, A.P., Liberzon, I., Sripada, R.K., 2017. Changes in salivary cortisol during psychotherapy for posttraumatic stress disorder: a pilot study in 30 veterans. J. Clin. Psychiatry 78, 599–603. https://doi.org/10.4088/JCP.15m10596.

Reynolds, B.M., Robles, T.F., Repetti, R.L., 2015. Measurement reactivity and fatigue effects in daily diary research with families. Dev. Psychol. 52, 442–456. https://doi.org/10.1037/dev0000081.

Roberts, B.W., DelVecchio, W.F., 2000. The Rank-order Consistency of Personality Traits From Childhood to Old Age: a Quantitative Review of Longitudinal Studies.

Robles, T.F., Reynolds, B.M., Repetti, R.L., Chung, P.J., 2013. Using daily diaries to study family settings, emotions, and health in everyday life. J. Soc. Pers. Relatsh. 30, 179–188. https://doi.org/10.1177/0265407512457102.

Romeo, R.D., 2010. Adolescence: a central event in shaping stress reactivity. Dev. Psychobiol. 52, 244. https://doi.org/10.1002/dev.20437.

Romeo, R.D., Bellani, R., Karatsoreos, I.N., Chhua, N., Vernov, M., Conrad, C.D., McEwen, B.S., 2006. Stress history and pubertal development interact to shape hypothalamic-pituitary-Adrenal Axis plasticity. Endocrinology 147, 1664–1674. https://doi.org/10.1210/en.2005-1432.

Romeo, R.D., Minhas, S., Svirsky, S.E., Hall, B.S., Savenkova, M., Karatsoreos, I.N., 2014. Pubertal shifts in adrenal responsiveness to stress and adrenocorticotropic hormone in male rats. Psychoneuroendocrinology 42, 146–152. https://doi.org/10.1016/j.psyneuen.2014.01.016.

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. https://doi.org/10.1016/j.psyneuen.2013.09.016.

Saxbe, D.E., 2008. A field (researcher's) guide to cortisol:… – google scholar. Health Psychol. Rev. 2, 163–190. https://doi.org/10.1080/17437190802530812.

Schrepf, A., Thaker, P.H., Goodheart, M.J., Bender, D., Slavich, G.M., Dahmoush, L., Penedo, F., DeGeest, K., Mendez, L., Lubaroff, D.M., others, 2015. Diurnal cortisol and survival in epithelial ovarian cancer. Psychoneuroendocrinology 53, 256–267.

Segerstrom, S.C., Boggero, I.A., Smith, G.T., Sephton, S.E., 2014. Variability and reliability of diurnal cortisol in younger and older adults: implications for design decisions. Psychoneuroendocrinology 49, 299–309. https://doi.org/10.1016/j.psyneuen.2014.07.022.

Sephton, S.E., Sapolsky, R.M., Kraemer, H.C., Spiegel, D., 2000. Diurnal cortisol rhythm as a predictor of breast cancer survival. J. Natl. Cancer Inst. 92, 994–1000. https://doi.org/10.1093/jnci/92.12.994.

Sephton, S.E., Lush, E., Dedert, E.A., Floyd, A.R., Rebholz, W.N., Dhabhar, F.S., Spiegel, D., Salmon, P., 2013. Diurnal cortisol rhythm as a predictor of lung cancer survival. Brain Behav. Immun. 30, S163–S170.

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. https://doi.org/10.1002/dev.20607.

Slopen, N., McLaughlin, K.A., Shonkoff, J.P., 2014. Interventions to improve cortisol regulation in children: a systematic review. Pediatrics 133, 312–326. https://doi.org/10.1542/peds.2013-1632.

Stalder, T., Kirschbaum, C., Kudielka, B.M., Adam, E.K., Pruessner, J.C., Wüst, S., Dockray, S., Smyth, N., Evans, P., Hellhammer, D.H., 2016. Assessment of the cortisol awakening response: expert consensus guidelines. Psychoneuroendocrinology 63, 414–432.

Vrshek-Schallhorn, S., Doane, L.D., Mineka, S., Zinbarg, R.E., Craske, M.G., Adam, E.K., 2013. The cortisol awakening response predicts major depression: predictive stability over a 4-year follow-up and effect of depression history. Psychol. Med. (Paris) 43, 483–493. https://doi.org/10.1017/S0033291712001213.

Wilson, E.B., 1927. Probable inference, the law of succession, and statistical inference. J. Am. Stat. Assoc. 22, 209–212.

Zalta, A.K., Shankman, S.A., 2016. Conducting psychopathology prevention research in the RDoC era. Clin. Psychol. Publ. Div. Clin. Psychol. Am. Psychol. Assoc. 23, 94. https://doi.org/10.1111/cpsp.12144.