Reliability of salivary cortisol levels in toddlers

Wesarg, C.; Van den Akker, A.L.; Hoeve, M.; Wiers, R.W.; Oei, N.Y.L.

DOI
10.1080/17405629.2021.1926230

Publication date
2022

Document Version
Final published version

Published in
European Journal of Developmental Psychology

EUROPEAN JOURNAL OF DEVELOPMENTAL PSYCHOLOGY
2022, VOL. 19, NO. 3, 419–437
https://doi.org/10.1080/17405629.2021.1926230

Reliability of salivary cortisol levels in toddlers

Christiane Wesarg [a,b], Alithe L. Van den Akker[a,c], Machteld Hoeve[a,c], Reinout W. Wiers [a,b,d,e] and Nicole Y. L. Oei[a,b,d]

[a]Research Priority Area (RPA) Yield, University of Amsterdam, Amsterdam, The Netherlands; [b]Developmental Psychology (Addiction Development and Psychopathology ADAPT-lab), University of Amsterdam, Amsterdam, The Netherlands; [c]Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, The Netherlands; [d]Amsterdam Brain and Cognition (ABC), University of Amsterdam, Amsterdam, The Netherlands; [e]Institute of Urban Mental Health, University of Amsterdam, Amsterdam, The Netherlands

ABSTRACT

The choice of cortisol sampling times in early childhood studies varies widely. Given that recommendations on sampling protocols are largely based on adults, the present study aimed to broaden current knowledge by examining how reliably cortisol measures obtained at different daytimes would reveal between-individual differences in toddlers' cortisol levels. Parents were instructed to take 10 saliva samples consecutively (five per day) from their toddler ($N = 19$; $M_{age} = 15.8$ months, $SD_{age} = 4.2$ months). Intra-class correlations (ICCs) were computed to evaluate cortisol reliability. Cortisol samples taken in the morning between 30 and 80 min after awakening and bedtime samples were most reliable in differentiating between children (ICCs ≥ .80). Wake-up cortisol samples taken within the first 30 min after awakening and afternoon samples showed moderate reliabilities (ICCs = .64), whereas the reliability of noon samples was poor (ICC = .43). Therefore, when investigating cortisol in young children while being restricted to a few samples only, assessing cortisol in the morning (at least 30 min after awakening) and at bedtime would be advisable.

ARTICLE HISTORY Received 11 November 2020; Accepted 28 April 2021

KEYWORDS Cortisol; reliability; sampling protocol; toddlers; early childhood

Introduction

The study of cortisol in early childhood has recently witnessed an increase in interest. A growing body of research points towards associations of cortisol with aspects of psychological development (Saridjan et al., 2014) and contextual factors including poverty (Zalewski et al., 2012), family instability (Suor et al., 2015) and maltreatment (Cicchetti et al., 2011). As

CONTACT Christiane Wesarg c.wesarg@uva.nl University of Amsterdam, Nieuwe Achtergracht 129, Amsterdam, WS 1018, The Netherlands

cortisol levels vary during the day, investigators often envisage cortisol to be sampled at the same time from every participant. However, the choice of cortisol sampling time(s) in early childhood varies widely, ranging from one single sample to multiple samples over multiple days. In order to inform researchers interested in measuring between-individual differences in young children's basal cortisol levels, more knowledge on the reliability of cortisol measures is needed.

Cortisol is a steroid hormone that is known as the 'end product' of the human hypothalamic–pituitary–adrenal (HPA) axis. This neuroendocrine system is involved in adaptation to external and internal challenges by inducing physiological and behavioural changes (Tsigos & Chrousos, 2002). Cortisol levels follow a 'diurnal rhythm' characterized by a cortisol peak about 30 min after awakening, followed by a sharp decline over the next hour or two, and a more gradual decline during the rest of the day (Fries et al., 2009; Kirschbaum & Hellhammer, 1989). Other factors such as food intake, sleep, daily hassles, and physical condition further influence cortisol levels (De Weerth & Van Geert, 2002). Food intake, for instance, leads to increased HPA axis activation (high cortisol peaks of short duration; De Weerth et al., 2003) but also alters the oral environment which affects the assessment of cortisol in saliva (Hanrahan et al., 2006). Daytime naps, in turn, may lower cortisol levels, with a return to pre-nap levels observed by 45 min post-nap in infants (Larson et al., 1991; De Weerth & Van Geert, 2002). As toddlers have multiple feeding moments and nap once or several times during the day, day-to-day variation in situational factors is likely to lead to considerable within-individual variation in cortisol levels, entailing that an individual's cortisol level assessed at a specific daytime varies across days.

If cortisol levels are assessed from different individuals at different times, a measure of cortisol will contain between- and within-person variability and measurement error. Although it is known that measurement error introduces only little variability when cortisol assays are performed competently (Kertes & van Dulmen, 2012), there is still uncertainty about the relative contribution of within- and between-person variability to measures of cortisol. A methodological study (Rotenberg et al., 2012) in children and adolescents aged 9–18 years old found that samples taken in the morning or at bedtime were less reliable in revealing between-individual differences in youth cortisol levels as compared to wake-up, lunch, and dinner samples. As far as we know, no similar study has been conducted in young children.

In the present study, we investigate how reliable salivary cortisol assessments at different sampling times detect between-individual differences in toddlers' basal cortisol levels. To this end, we asked parents to take 10 saliva samples from their toddler over consecutive days, yielding five samples per 24 hours. We hereby aim to inform researchers who are restricted to one or a few samples in choosing a cortisol sampling protocol in which variability arising from between-individual differences in cortisol levels is maximized, while variability due to within-individual variation is minimized.

Method

Participants

Recruitment was based on flyers distributed around campus and by approaching families directly on the street. Further, information about the study was sent to families from a list of former participants from the BabyLab of Amsterdam. The study was approved by the Review Board from the Department of Child Development and Education at the University of Amsterdam (2018-CDE-9752). An informed consent was obtained from parents prior to screening.

Families with a toddler aged between 13–15 months and 25–27 months could take part in the study. These age groups were chosen for practical reasons, as the present study's aim was to determine the most reliable time points for a future intervention study in which children would be tested at age 14 months and up to 1 year later. Caregivers had to be at least 18 years old and fluent in Dutch. Pregnancy duration had to have lasted over 32 weeks. Exclusion criteria regarding the child's health assessed in an online screening were a serious medical condition or developmental disorder, cognitive, sensory, or motor impairments, asthma, and intake of oral steroid medication.

All 21 families who expressed their interest in the study were eligible to participate. Two families decided to stop participation after the first home visit due to current stress, leaving 19 families. In all families, the participating caregiver was the biological mother, and in almost all families (n = 18), the child's biological parents lived together. Toddlers (11 girls and 8 boys) were aged between 14.1 and 26.6 months (M = 15.8 months, SD = 4.2 months): sixteen toddlers (11 girls and 5 boys) were aged around 14 months, and three toddlers (all boys) were aged around 26 months. Fifteen (78.9%) of the children were identified by their mother as Dutch,

one child as Antilles-Dutch, and three children as multi-ethnic. All children were medication-free. Mothers were predominantly highly educated (n = 15 with a university degree). Median monthly family income was above 3,200 euros (range: <800 euros to >3,200 euros).

Procedure

Home visits were scheduled on non-childcare days, so that the parent would be at home with the child on the following day as well. The experimenter arrived around 11:30 h to instruct the parent on how to take the saliva samples of their child using the SalivaBio Children's Swab (SCS, Salimetrics®). The SCS is a suitable device for saliva collection in children aged between 6 months and 6 years old. The synthetic swab is placed in the child's mouth where it takes up saliva during 60–90 seconds. The SCS is longer than the one commonly used for adults (125 vs. 30 mm; see Salimetrics website), eliminating any choking hazard.

Parents were instructed to take 10 saliva samples from their toddler: Sampling started around 12:00 h (noon sample) in the presence of the experimenter. The following sampling times were around 15:00 h (afternoon sample), shortly before going to bed (bedtime sample), in the morning directly after awakening (wake-up sample), and 30 min later (morning sample; as in the 1-day sampling protocol of Saridjan et al., 2010). Parents were asked to repeat the saliva collection during the consecutive 24 hours, sticking as close as possible to the sampling times of their first five samples.

Families were free to follow their normal daily routines on sampling days. Parents were instructed to take samples at least 30 min after a meal and the afternoon sample at least 1 hour after their child awoke from a nap. The evening sample was planned as close to sleeping as possible, such that dinner and tooth brushing took place at least 30 min prior to sampling. Parents who indicated that their child would go to bed directly after dinner were instructed to conduct the saliva sampling directly before the meal. With regard to the two morning samples, parents were asked to take the first one directly when the child woke up (wake-up sample) and the second one 30 min later (morning sample). Parents who indicated that their child would receive breast milk immediately after awakening were told to postpone the wake-up sample to 30 min after the meal. Consequently, the morning sample was advised to be taken 30 min after the wake-up sample.

Parents received a sampling diary which contained all the instructions relevant for each sampling time. On this diary, parents were instructed to immediately note down the exact collection times before placing the samples in their home freezer. Apart from demographic questions, parents were asked to fill out questions about their child's day, including information on sleeping, feeding, and daily hassles, for each day of the saliva collection.

During a short second home visit, the experimenter picked up the saliva samples in a cool bag and brought them to the university where they were stored at −20°C in the freezer.

Salivary cortisol

Out of the 190 distributed samples, 176 samples were returned and sent by post to the laboratory of the Department of Biological Psychology at the Technical University of Dresden, where they were stored at −20°C until analysis. After thawing, salivettes were centrifuged at 3,000 rpm for 5 min, which resulted in a clear supernatant of low viscosity. Salivary concentrations were measured using a commercially available chemiluminescence immunoassay with high sensitivity (IBL International, Hamburg, Germany). The intra- and inter-assay coefficients for cortisol were below 9%.

Data analysis

First, the data was – if necessary – excluded according to the following rules: the wake-up sample had to be taken within the first 30 min after wake-up and the morning sample between 30 and 90 min after wake-up. Although we first intended to assess cortisol levels at wake-up and exactly 30 min later, the fact that sampling occurred too late (i.e., >10 min after the intended time) for 30 samples led to the post-hoc decision to extend both time categories.

The reliability of cortisol measures per sampling time was examined through intra-class correlation coefficients (ICCs). Cortisol values were log-transformed to correct for non-normality. Subsequently, a series of multilevel models were estimated using R (R Core Team, 2017). For each sampling time, separate two-level models were calculated, in which Level-1 represented 'days' and Level-2 'individuals'. Based on variance estimates from intercept-only models, ICCs were computed using the

method described by Shrout and Fleiss (1979) in which the ICC represents the ratio of between-individual variance (τ^2) to total variance (between-individual variance τ^2 plus within-individual (or day) variance σ^2):

$$ICC = \frac{\tau^2}{\tau^2 + \sigma^2}$$

The ICC ranges from 0 to 1; high values indicate that a single cortisol measure reliably reflects true between-individual differences (Hruschka et al., 2005). ICCs were interpreted according to the guidelines of Koo and Li (2016); values <.50 indicate poor reliability, between .50 and .75 moderate reliability, between .75 and .90 good reliability, and >.90 excellent reliability. An ICC was further computed for individual mean cortisol levels. To this end, mean levels were calculated for every child for both sampling days by including only samples for which cortisol was determinable on both days at the same sampling time.

In order to examine whether the difference between sampling times within individuals (e.g., the afternoon sample taken 1 hour later on the second day of collection) would influence ICC estimates, the variable 'time difference' was introduced into the fixed part of the models. Models were then compared through the likelihood ratio χ^2 test by subtracting the −2 loglikelihood. If model fit improved significantly, ICCs were recomputed.

As cortisol levels may be more tightly related to the wake-up time compared to the actual time, this procedure was repeated by introducing the variable 'delay' (indicating the time difference between wake-up and sampling time) into the models on wake-up and morning cortisol levels. Given that delays in sample collection may obscure the course of the diurnal cortisol rhythm particularly in the early morning during which cortisol levels are expected to show the greatest changes (Bäumler et al., 2013; Gribbin et al., 2012), variation in delay may influence ICC estimates more than time differences unrelated to wake-up.

Results

Descriptive analyses

Of the 190 samples that were distributed to families (10 samples × 19 families), 14 (7.4%) samples from eight different families were not returned, of which six were wake-up samples and four were morning samples. Of the 176 returned samples, 8 (4.5%) samples from seven

different families lacked the required saliva volume to perform the analysis, of which five were wake-up samples and one was a morning sample. One outlier sample was excluded because of an extreme cortisol level (398.3 nmol/L), probably reflecting saliva sample contamination. Further, two wake-up and three morning samples were excluded as they were taken too late by the parent. Thus, the final sample consisted of 162 cortisol samples.

In line with the cortisol levels reported in other studies (e.g., Watamura et al., 2004), toddlers' mean cortisol levels were 7.30 nmol/L, *SD* = 8.26 nmol/L (for toddlers aged around 14 months old: *M* = 7.29, *SD* = 9.42; for toddlers aged around 26 months old: *M* = 7.32, *SD* = 7.59). Cortisol levels per sampling time are shown in Table 1. Wake-up samples were taken within 30 min post-awakening (*M* = 12 min, *SD* = 9 min). Morning samples were taken between 30 and 80 min post-awakening (*M* = 45 min, *SD* = 14 min). All children napped once or twice during the day (*M* = 1.3 naps, *SD* = 0.5 naps). A morning nap was taken by five toddlers on the first day and by four toddlers on the second day of saliva collection. For these toddlers, noon samples were taken around 51 min post-nap on average (*SD* = 61 min). Twelve toddlers had napped before the afternoon sample on the first day of saliva collection and ten toddlers on the second day. For these toddlers, afternoon samples were taken around 71 min post-nap on average (*SD* = 44 min). Figure 1 visualizes the

Table 1. Cortisol sampling times and levels.

		Sampling time (hh:mm)				Untransformed cortisol levels (in nmol/L)	
Sample	Number of samples (*N* = 162)	M	SD	Min	Max	M	SD
Wake-up							
Sample 1 (+ 0 hrs)	14	07:26	00:55	06:00	09:24	14.14	5.99
Sample 2 (+ 24 hrs)	13	07:26	00:42	06:32	08:45	15.42	13.31
Morning							
Sample 1 (+ 0 hrs)	14	07:51	01:03	06:30	10:02	9.32	5.44
Sample 2 (+ 24 hrs)	12	07:51	00:39	07:05	09:20	9.60	8.19
Noon							
Sample 1 (+ 0 hrs)	19	12:02	00:34	11:08	13:26	7.61	8.19
Sample 2 (+ 24 hrs)	17	12:24	00:50	11:12	14:30	5.44	7.74
Afternoon							
Sample 1 (+ 0 hrs)	19	15:46	01:09	14:19	18:45	5.69	8.52
Sample 2 (+ 24 hrs)	17	16:05	01:17	14:25	18:33	4.98	5.72
Bedtime							
Sample 1 (+ 0 hrs)	19	19:17	00:46	17:45	20:40	2.97	4.44
Sample 2 (+ 24 hrs)	18	19:27	00:57	17:30	21:15	2.90	4.74

diurnal cortisol profile for each day of collection. A full correlation matrix of cortisol levels at different sampling times (Table A1) can be found in the Appendix.

Reliability analyses

Good reliability was observed for the morning sample (ICC = .81) and for the bedtime sample (ICC = .80). Wake-up and afternoon samples had a moderate reliability (ICCs = .64). A poor reliability was observed for the noon sample (ICC = .43).[1] Table 2 shows the multi-level models for cortisol

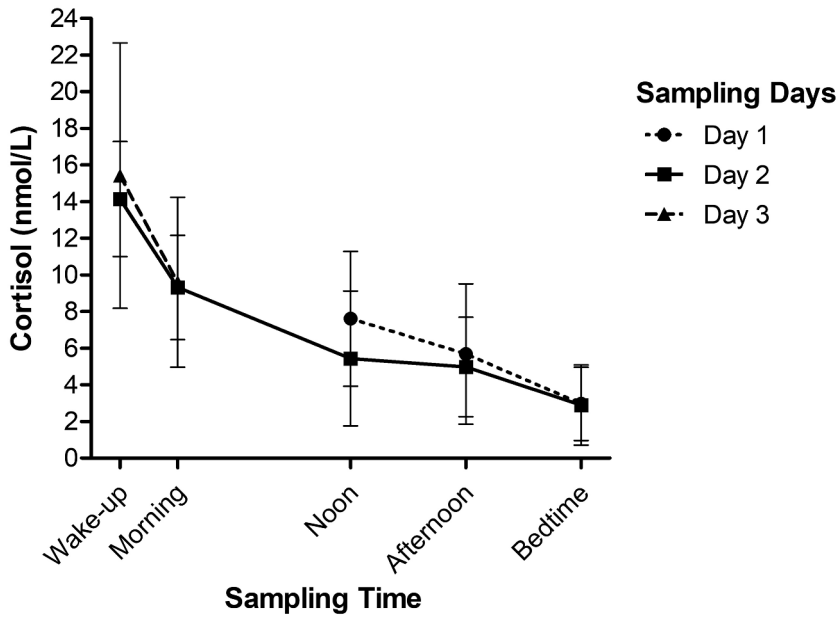

Figure 1. Mean salivary cortisol concentrations (in nmol/L) during the day by sampling day. Sampling was spread over three consecutive days, starting at noon on Day 1. Error bars represent the 95% confidence intervals.

[1]Based on a sensitivity analysis including only the 14-month-old toddlers (n = 16), the rank-order of reliability estimates did not change and interpretation of reliabilities remained the same except for the reliability of bedtime cortisol levels which was moderate instead of good (for multi-level models and ICCs of the sample consisting of 14-month-old toddlers, see Table A2).

Table 2. Estimates for multi-level models of log-transformed cortisol data.

	Wake-up		Morning		Noon		Afternoon		Bedtime	
Number of samples	27		26		36		36		37	
Fixed effects	Coefficient	*SE*	Coefficient	*SE*	Coefficient	*SE*	Coefficient	*SE*	Coefficient	*SE*
γ_{00} = intercept	1.0707	0.0678	0.8978	0.0710	0.6128	0.0779	0.4833	0.0870	0.1870	0.0951
Random effects	Parameter	*SD*	Parameter	*SD*	Parameter	*SD*	Parameter	*SD*	Parameter	*SD*
Between-individual variance: τ^2	0.0540	0.2324	0.0659	0.2568	0.0673	0.2594	0.1096	0.3310	0.1524	0.3904
Within-individual variance: σ^2	0.0308	0.1755	0.0154	0.1239	0.0889	0.2981	0.0629	0.2508	0.0370	0.1922
ICC	0.64		0.81		0.43		0.64		0.80	

The intra-class correlation coefficient (ICC) represents the ratio of between-individual variance (τ^2) to total variance (between-individual variance τ^2 plus within-individual variance σ^2) estimated from multi-level models of log-transformed cortisol data.

at each sampling time. Good reliability (ICC = .86) was observed for mean cortisol levels.

Including the difference between sampling times within individuals did not improve model fit for any of the five sampling times. Including the time difference between wake-up and sampling time (variable 'delay') improved model fit for wake-up cortisol levels (Table A3). Delay ranged from 0 to 30 min post awakening (M = 13.17 min, SD = 13.08 min). A significant reduction in deviance was observed when controlling for delay (deviance = 2.883, $\chi^2 = 5.77$, $df = 1$, $p = .016$). When sampling took place later within the first 30 min after awakening, cortisol levels were higher ($t = 2.561$, $df = 24.626$, two-sided $p = .017$). After inclusion of delay, re-computation of the ICC still yielded moderate reliability for measures of wake-up cortisol (ICC = .59).

Discussion

In a sample composed of 19 toddlers, we determined how reliable saliva samples taken at different daytimes would reveal between-individual differences in cortisol. We found that samples taken either in the morning (between 30 and 80 min after awakening) or shortly before bedtime were most reliable in differentiating participants. Samples taken within the first 30 min after awakening or in the afternoon showed a moderate reliability, whereas the reliability of noon samples was poor. Aggregating individuals' cortisol levels obtained at different times to mean levels yielded the highest reliability.

Our morning reliability findings are in line with studies on school-aged children and adolescents revealing lower test–retest correlations between cortisol levels assessed directly upon awakening as compared to 30 min (Brosnan et al., 2009; O'Connor et al., 2005) or 60 min later (Michels et al., 2012). Cortisol levels are tightly related to an individual's waking time, showing steep increases until reaching a peak around 30 min after awakening (Gribbin et al., 2012). We therefore examined whether including the delay in sampling time in the multilevel model would change the reliability estimate. After controlling for sampling delay, the ICC for wake-up cortisol slightly decreased, which is not uncommon when predictors are added to empty models (Baguley, 2012). Considering that both multilevel models revealed a moderate reliability of wake-up cortisol, we conclude that the delay in sampling time is not a plausible explanation for the lower reliability of wake-up compared to morning samples.

Another possible explanation could relate to intra-individual variation in sleep, as the length and quality of sleep during the night have previously been associated with toddlers' wake-up cortisol levels (Scher et al., 2010). Furthermore, wake-up cortisol levels have been observed to be inversely related to cortisol increases up to 30 min later (Bäumler et al., 2013). Given this, it is possible that cortisol levels measured 30 min after awakening are less susceptible to intra-individual variation in sleep length and quality, rendering them more reliable to measure between-individual differences. As we did not assess sleep length and quality in our study, the influence of these factors on cortisol reliability in toddlers needs to be further investigated.

It is striking that samples taken around noon were the least reliable in distinguishing between our participants. Given that the noon sample was taken by parents about half an hour after the experimenter's arrival, it is possible that children showed a cortisol reaction due to this potentially stressful situation. Cortisol levels rise upon confrontation with a stressor, reaching their peak between 20 and 40 min after stressor onset (Goldberg et al., 2003). Indeed, noon cortisol levels were higher on the first day of collection as compared to the second day. Thus, the poor reliability of noon samples may apply to the specific circumstances of our study and may not be generalizable. Nonetheless, our findings indicate that if one wishes to reliably assess between-individual differences in children's cortisol levels, it may be advisable to discard 'instructional samples' taken in the presence of an experimenter and only use samples that are later taken by the parent. Alternatively, samples should be taken immediately after the arrival of the experimenter.

With regard to the afternoon cortisol levels showing moderate reliability, we assume that day-to-day differences in children's routines may have affected cortisol levels in a day-specific way. Although we aimed to minimize the impact of cortisol-influencing variables such as food intake and naps through careful design – asking parents to wait 30 min after a meal and 1 hour after a nap before sampling – not all parents could adhere to these instructions. In some cases, the afternoon sample was taken less than an hour after a nap, and differences in sampling times with regard to naps could have lowered the reliability of afternoon samples. Furthermore, it is likely that other factors such as physical activity (Kertes & Gunnar, 2004) or arousing/stressful experiences (Goldberg et al., 2003) contributed to variance stemming from the sampling occasion. Given that sampling mainly took place during the

weekend, it is possible that children's activities differed considerably between sampling days. This day-to-day variation may be higher in the middle of the day as compared to the early morning or around bedtime, during which families likely follow routines, which could in turn lead to lower reliabilities of noon and afternoon samples. In support of this hypothesis, a lower reliability of noon and afternoon samples as compared to morning and bedtime cortisol samples has been observed in school-aged children and college students (O'Connor et al., 2005; Zhang et al., 2017).

We further found that samples taken shortly before bedtime showed a good reliability, which accords with previous reports on higher correlations between bedtime compared to wake-up cortisol levels in children (Dozier et al., 2006; Michels et al., 2012). Opposite findings were reported by Rotenberg et al. (2012) who observed that youth bedtime cortisol levels were less reliable than wake-up and dinner cortisol levels. However, participants in the latter study were older (9–18 years old), and sampling mainly occurred during weekdays when children and adolescents attended school, which renders their findings less generalizable to our study in which sampling in toddlers occurred on non-childcare days.

There are limitations to our study which need to be taken into consideration when interpreting the results. Our sample size was small, and some samples were missing, particularly wake-up and morning samples. If morning cortisol is the focus of the study, it is therefore advisable to ask parents to repeat sampling in the morning to compensate for potential missing data. Additionally, we only assessed salivary cortisol on two consecutive days. Future studies on cortisol reliability might, for instance, endeavour sampling on an additional third day. Given that in our study, parents – and possibly toddlers as well – experienced the collection of 10 samples spread over 48 hours as a burden, such a future methodological study may consider having less daily samples on more than two sampling days. Further, our sample was composed of highly educated families, with parents willing and able to perform saliva sampling at home. Thus, caution must be applied in generalizing these findings, and replication based on a larger, socio-economically heterogeneous sample representative of the general population is required.

Given that we relied on parental reports for toddlers' wake-up times and all sampling times, inaccuracies in these reports might have biased our findings. Studies suggest that rates of compliance with sampling

protocols are higher based on parental self-reports compared to objective measures, such as those obtained with electronic monitoring containers which detect and store the times of container openings (Smith & Dougherty, 2014; Valentino et al., 2017). Considering that these additional technologies are associated with financial costs, it is likely that future studies with lower budgets will also rely on self-report, which renders our findings particularly informative for such studies. Further work should be undertaken to investigate how effective highly monitored procedures are in improving the reliability of cortisol measures.

Comparing the reliability estimates obtained in our study, it stands out that morning samples (taken between 30 and 80 min post-awakening) and bedtime samples were most reliable in revealing between-individual differences in toddlers' cortisol levels. We thus recommend these sampling times on non-childcare days for future studies that intend to assess basal cortisol levels in very young children with just few sampling opportunities.

Acknowledgements

This work was supported by the UvA-FMG Research Priority Area Yield.

Disclosure Statement

The authors declare that they have no relevant or material financial interests relating to the research described in this paper.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Funding

This work was supported by the Research Priority Area (RPA) Yield.

ORCID

Christiane Wesarg http://orcid.org/0000-0001-7861-1723
Reinout W. Wiers http://orcid.org/0000-0002-4312-9766

References

Baguley, T. (2012). *Serious stats: A guide to advanced statistics for the behavioral sciences*. Palgrave Macmillan.

Bäumler, D., Kirschbaum, C., Kliegel, M., Alexander, N., & Stalder, T. (2013). The cortisol awakening response in toddlers and young children. *Psychoneuroendocrinology*, *38* (11), 2485–2492. https://doi.org/10.1016/j.psyneuen.2013.05.008

Brosnan, M., Turner-Cobb, J., Munro-Naan, Z., & Jessop, D. (2009). Absence of a normal Cortisol Awakening Response (CAR) in adolescent males with Asperger syndrome (AS). *Psychoneuroendocrinology*, *34*(7), 1095–1100. https://doi.org/10.1016/j.psyneuen.2009.02.011

Cicchetti, D., Rogosch, F. A., Toth, S. L., & Sturge-Apple, M. L. (2011). Normalizing the development of cortisol regulation in maltreated infants through preventive interventions. *Development and Psychopathology*, *23*(3), 789–800. https://doi.org/10.1017/S0954579411000307

De Weerth, C., & Van Geert, P. (2002). A longitudinal study of basal cortisol in infants: Intra-individual variability, circadian rhythm and developmental trends. *Infant Behavior and Development*, *25*(4), 375–398. https://doi.org/10.1016/S0163-6383(02)00141-8

De Weerth, C., Zijl, R. H., & Buitelaar, J. K. (2003). Development of cortisol circadian rhythm in infancy. *Early Human Development*, *73*(1–2), 39–52. https://doi.org/10.1016/S0378-3782(03)00074-4

Dozier, M., Manni, M., Gordon, M. K., Peloso, E., Gunnar, M. R., Stovall-mcclough, K. C., Eldreth, D., & Levine, S. (2006). Foster children's Diurnal production of cortisol: An exploratory study. *Child Maltreatment*, *11*(2), 189–197. https://doi.org/10.1177/1077559505285779

Fries, E., Dettenborn, L., & Kirschbaum, C. (2009). The cortisol awakening response (CAR): Facts and future directions. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *72*(1), 67–73. https://doi.org/10.1016/j.ijpsycho.2008.03.014

Goldberg, S., Levitan, R., Leung, E., Masellis, M., Basile, V. S., Nemeroff, C. B., & Atkinson, L. (2003). Cortisol concentrations in 12- to 18-month-old infants: Stability over time, location, and stressor. *Biological Psychiatry*, *54*(7), 719–726. https://doi.org/10.1016/S0006-3223(03)00010-6

Gribbin, C. E., Watamura, S. E., Cairns, A., Harsh, J. R., & LeBourgeois, M. K. (2012). The cortisol awakening response (CAR) in 2- to 4-year-old children: Effects of acute nighttime sleep restriction, wake time, and daytime napping. *Developmental Psychobiology*, *54*(4), 412–422. https://doi.org/10.1002/dev.20599

Hanrahan, K., McCarthy, A. M., Kleiber, C., Lutgendorf, S., & Tsalikian, E. (2006). Strategies for salivary cortisol collection and analysis in research with children. *Applied Nursing Research*, *19*(2), 95–101. https://doi.org/10.1016/j.apnr.2006.02.001

Hruschka, D. J., Kohrt, B. A., & Worthman, C. M. (2005). Estimating between- and within-individual variation in cortisol levels using multilevel models. *Psychoneuroendocrinology*, *30*(7), 698–714. https://doi.org/10.1016/j.psyneuen.2005.03.002

Kertes, D. A., & Gunnar, M. R. (2004). Evening Activities as a Potential Confound in Research on the Adrenocortical System in Children. *Child Development*, *75*(1), 193–204. https://doi.org/10.1111/j.1467-8624.2004.00663.x

Kertes, D. A., & van Dulmen, M. (2012). Latent state trait modeling of children's cortisol at two points of the diurnal cycle. *Psychoneuroendocrinology*, *37*(2), 249–255. https://doi.org/10.1016/j.psyneuen.2011.06.009

Kirschbaum, C., & Hellhammer, D. H. (1989). Salivary cortisol in psychobiological research: An overview. *Neuropsychobiology*, *22*(3), 150–169. https://doi.org/10.1159/000118611

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Larson, M. C., Gunnar, M. R., & Hertsgaard, L. (1991). The effects of morning naps, car trips, and maternal separation on adrenocortical activity in human infants. *Child Development*, *62*(2), 362–372. https://doi.org/10.2307/1131009

MacArthur SES & Health network | Research. (2000). Retrieved February 5, 2020, from https://macses.ucsf.edu/research/allostatic/salivarycort.php

Michels, N., Sioen, I., De Vriendt, T., Huybrechts, I., Vanaelst, B., & De Henauw, S. (2012). Children's morning and evening salivary cortisol: Pattern, instruction compliance and sampling confounders. *Hormone Research in Paediatrics*, *77*(1), 27–35. https://doi.org/10.1159/000334412

O'Connor, T. G., Ben-Shlomo, Y., Heron, J., Golding, J., Adams, D., & Glover, V. (2005). Prenatal anxiety predicts individual differences in cortisol in pre-adolescent children. *Biological Psychiatry*, *58*(3), 211–217. https://doi.org/10.1016/j.biopsych.2005.03.032

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Rotenberg, S., McGrath, J. J., Roy-Gagnon, M.-H., & Tu, M. T. (2012). Stability of the diurnal cortisol profile in children and adolescents. *Psychoneuroendocrinology*, *37*(12), 1981–1989. https://doi.org/10.1016/j.psyneuen.2012.04.014

Saridjan, N. S., Henrichs, J., Schenk, J. J., Jaddoe, V. W. V., Hofman, A., Kirschbaum, C., Verhulst, F. C., & Tiemeier, H. (2014). Diurnal cortisol rhythm and cognitive functioning in toddlers: The generation R study. *Child Neuropsychology*, *20*(2), 210–229. https://doi.org/10.1080/09297049.2013.763921

Saridjan, N. S., Huizink, A. C., Koetsier, J. A., Jaddoe, V. W., Mackenbach, J. P., Hofman, A., Kirschbaum, C., Verhulst, F. C., & Tiemeier, H. (2010). Do social disadvantage and early family adversity affect the diurnal cortisol rhythm in infants? The generation R study. *Hormones and Behavior*, *57*(2), 247–254. https://doi.org/10.1016/j.yhbeh.2009.12.001

Scher, A., Hall, W. A., Zaidman-Zait, A., & Weinberg, J. (2010). Sleep quality, cortisol levels, and behavioral regulation in toddlers. *Developmental Psychobiology*, *52*(1), 44–53. https://doi.org/10.1002/dev.20410

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Smith, V. C., & Dougherty, L. R. (2014). Noisy spit: Parental noncompliance with child salivary cortisol sampling. *Developmental Psychobiology*, *56*(4), 647–656. https://doi.org/10.1002/dev.21133

Suor, J. H., Sturge-Apple, M. L., Davies, P. T., Cicchetti, D., & Manning, L. G. (2015). Tracing differential pathways of risk: Associations among family adversity, cortisol, and cognitive functioning in childhood. *Child Development*, *86*(4), 1142–1158. https://doi.org/10.1111/cdev.12376

Tsigos, C., & Chrousos, G. P. (2002). Hypothalamic-pituitary-adrenal axis, neuroendocrine factors and stress. *Journal of Psychosomatic Research*, *53*(4), 865–871. https://doi.org/10.1016/S0022-3999(02)00429-4

Valentino, K., De Alba, A., Hibel, L. C., Fondren, K., & McDonnell, C. G. (2017). Adherence to diurnal cortisol sampling among mother-child dyads from maltreating and nonmaltreating families. *Child Maltreatment*, *22*(4), 286–294. https://doi.org/10.1177/1077559517725208

Watamura, S. E., Donzella, B., Kertes, D. A., & Gunnar, M. R. (2004). Developmental changes in baseline cortisol activity in early childhood: Relations with napping and effortful control. *Developmental Psychobiology*, *45*(3), 125–133. https://doi.org/10.1002/dev.20026

Zalewski, M., Lengua, L. J., Fisher, P. A., Trancik, A., Bush, N. R., & Meltzoff, A. N. (2012). Poverty and single parenting: Relations with preschoolers' cortisol and effortful control. *Infant and Child Development*, *21*(5), 537–554. https://doi.org/10.1002/icd.1759

Zhang, Q., Chen, Z., Chen, S., Xu, Y., & Deng, H. (2017). Intraindividual stability of cortisol and cortisone and the ratio of cortisol to cortisone in saliva, urine and hair. *Steroids*, *118*, 61–67. https://doi.org/10.1016/j.steroids.2016.12.008

Appendix

Table A1. Full correlation matrix of untransformed cortisol levels.

Sample	1	2	3	4	5	6	7	8	9
Day 1									
1 (Noon)	–								
2 (Afternoon)	.59**	–							
3 (Bedtime)	.68**	.90**	–						
Day 2									
4 (Wake-up)	.01	.33	.44	–					
5 (Morning)	.02	.88**	.84**	.40	–				
6 (Noon)	.31	.83**	.59*	.25	.80**	–			
7 (Afternoon)	.29	.82**	.64**	.33	.79**	.84**	–		
8 (Bedtime)	.47*	.95**	.80**	.33	.84**	.89**	.83**	–	
Day 3									
9 (Wake-up)	.08	.56*	.37	.43	.68*	.61*	.76**	.58*	–
10 (Morning)	.12	.82**	.56	.30	.96**	.81**	.84**	.86**	.80**

* $p < .05$; ** $p < .01$.

Table A2. Estimates for multi-level models of log-transformed cortisol data of 14-month-old toddlers.

	Wake-up		Morning		Noon		Afternoon		Bedtime	
Number of samples	22		21		30		30		31	
Fixed effects	Coefficient	*SE*	Coefficient	*SE*	Coefficient	*SE*	Coefficient	*SE*	Coefficient	*SE*
γ_{00} = intercept	1.1028	0.0731	0.9024	0.0836	0.6294	0.0792	0.4708	0.0900	0.1967	0.0832
Random effects	Parameter	*SD*	Parameter	*SD*	Parameter	*SD*	Parameter	*SD*	Parameter	*SD*
Between-individual variance: τ^2	0.0474	0.2177	0.0741	0.2721	0.0491	0.2216	0.0942	0.3069	0.0896	0.2992
Within-individual variance: σ^2	0.0352	0.1875	0.0160	0.1265	0.0941	0.3067	0.0644	0.2538	0.0400	0.2001
ICCs	0.57		0.82		0.34		0.59		0.69	

The intra-class correlation coefficient (ICC) represents the ratio of between-individual variance (τ^2) to total variance (between-individual variance τ^2 plus within-individual variance σ^2) estimated from multi-level models of log-transformed cortisol data.

Table A3. Estimates for two multi-level models of wake-up cortisol levels.

	Model 1: Intercept-only model		Model 2 including 'delay'	
Number of samples	25		25	
Fixed effects	Coefficient	*SE*	Coefficient	*SE*
γ_{00} = intercept	1.0610	0.0717	0.8907	0.091
γ_{10} = delay			0.0139	0.0054
Random effects	Parameter	*SD*	Parameter	*SD*
Between-individual variance: τ^2	0.0565	0.2376	0.0404	0.2010
Within-individual variance: σ^2	0.0322	0.1793	0.0276	0.1662
−2 log-likelihood	−2.577		0.306	
ICC	0.64		0.59	

Model 1 differs from the empty (intercept-only) model for wake-up cortisol levels described in Table 2 in that only cases with reported wake-up times (N = 25) were included to enable comparability with Model 2.