

WestminsterResearch

<http://www.westminster.ac.uk/westminsterresearch>

**Assessment of the cortisol awakening response: expert
consensus guidelines**

**Stalder T., Kirschbaum C., Kudielka B.M., Adam E.K., Pruessner
J.C., Wust S., Dockray S., Smyth N., Evans P., Hellhammer D.H.,
Miller R., Wetherell M.A., Lupien S. and Clow A.**

NOTICE: this is the authors' version of a work that was accepted for publication in Psychoneuroendocrinology. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Psychoneuroendocrinology, PNEC (63) 414-432 0306-4530.

Psychoneuroendocrinology is available online at:

<https://dx.doi.org/10.1016/j.psyneuen.2015.10.010>

© 2015. This manuscript version is made available under the CC-BY-NC-ND 4.0 license

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners.

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch: (<http://westminsterresearch.wmin.ac.uk/>).

In case of abuse or copyright appearing without permission e-mail repository@westminster.ac.uk

Assessment of the cortisol awakening response: expert consensus guidelines

Tobias Stalder^{a†}, Clemens Kirschbaum^a, Brigitte M. Kudielka^b, Emma K. Adam^c, Jens C. Pruessner^d, Stefan Wüst^b, Samantha Dockray^e, Nina Smyth^f, Phil Evans^f, Dirk H. Hellhammer^g, Robert Miller^a, Mark A. Wetherell^h, Sonia J. Lupienⁱ, Angela Clow^f

^a Department of Psychology, TU Dresden, Dresden, Germany

^b Department of Psychology, University of Regensburg, Regensburg, Germany

^c School of Education and Social Policy, Northwestern University, Evanston, U.S.A.

^d Department of Psychiatry, McGill University, Montreal, Canada

^e School of Applied Psychology, University College Cork, Cork, Ireland

^f Department of Psychology, University of Westminster, London, U.K.

^g Department of Psychology, Trier University, Trier, Germany

^h Department of Psychology, Northumbria University, Newcastle, U.K.

ⁱ Department of Psychiatry, University of Montreal, Montreal, Quebec, Canada

† Correspondence: Dr. Tobias Stalder

Technische Universität Dresden

Department of Psychology

Zellescher Weg 19

01069 Dresden, Germany

Tel: +49 351 463 -39678 Fax: +49 351 463 -37274

E-mail: tobias.stalder@tu-dresden.de

Word count: 2182

Number of figures: 3

Number of tables: 6

Running head: ISPNE CAR consensus

Keywords: CAR; measurement; saliva; adherence; covariates; guidelines

Summary

The cortisol awakening response (CAR), the marked increase in cortisol secretion over the first 30–45 min after morning awakening, has been related to a wide range of psychosocial, physical and mental health parameters, making it a key variable for psychoneuroendocrinological research. The CAR is typically assessed from self-collection of saliva samples within the domestic setting. While this confers ecological validity, it lacks direct researcher oversight which can be problematic as the validity of CAR measurement critically relies on participants closely following a timed sampling schedule, beginning with the moment of awakening. Researchers assessing the CAR thus need to take important steps to maximize and monitor saliva sampling accuracy as well as consider a range of other relevant methodological factors. To promote best practice of future research in this field, the International Society of Psychoneuroendocrinology initiated an expert panel charged with (i) summarizing relevant evidence and collective experience on methodological factors affecting CAR assessment and (ii) formulating clear consensus guidelines for future research. The present report summarizes the results of this undertaking. Consensus guidelines are presented on central aspects of CAR assessment, including objective control of sampling accuracy/adherence, participant instructions, covariate accounting, sampling protocols, quantification strategies as well as reporting and interpreting of CAR data. Meeting these methodological standards in future research will create more powerful research designs, thus yielding more reliable and reproducible results and helping to further advance understanding in this evolving field of research.

Contents.

1. Introduction
2. Cortisol awakening response
 - 2.1 Description and distinctive features
 - 2.2 Main components
3. Inaccurate sampling: Prevalence and impact
 - 3.1 Delay between awakening and initiation of sampling
 - 3.2 Inaccurate post-awakening sampling
4. Strategies for dealing with inaccurate sampling
 - 4.1 Objective monitoring of sampling accuracy
 - 4.1.1 Methods for verifying awakening and sampling times
 - 4.1.2 Dealing with verified inaccurate data
 - 4.2 Are there viable strategies without the use of objective measures?
 - 4.2.1 Forced awakening
 - 4.2.2 Increasing statistical power
 - 4.2.3 Exclusion of CAR non-responders
 - 4.2.4 Informational strategies
5. Maximizing adherence
6. Dealing with covariates
 - 6.1 Instructions about post-awakening behavior
 - 6.2 Control variables
 - 6.2.1 State covariates
 - 6.2.2 Trait covariates
 - 6.3 Exclusion criteria
7. Procedural and design considerations
 - 7.1 Sampling times
 - 7.2 Number of study days (cross-sectional research)
 - 7.3 Sample storage and cortisol analysis
 - 7.4. Statistical considerations
8. Summary and guidelines

1. Introduction

Abnormal secretion of the glucocorticoid hormone cortisol as the final product of the hypothalamus-pituitary-adrenal (HPA) axis is considered a crucial factor in linking the experience of chronic psychosocial stress to adverse effects on health (Chrousos, 2009). Besides reactivity to acute stressors, changes to the circadian regulation of cortisol secretion are considered important in this context (Kondratova and Kondratov, 2012; Menet and Rosbash, 2011; Nader et al., 2010). An aspect of cortisol regulation that is of special interest to psychoneuroendocrinological (PNE) inquiry is the cortisol awakening response (CAR), which describes the marked increase in cortisol levels across the first 30–45 min following morning awakening (Clow et al., 2004, 2010; Elder et al., 2014; Kudielka and Wüst, 2010). The CAR was first systematically described in the mid-1990s (Pruessner et al., 1997) and soon gained attention as a favorable biomarker in PNE research due to several methodological advantages over previously employed cortisol assessment strategies (see section 2). These advantages together with evidence showing unique associations of the CAR with psychosocial, psychiatric and health-related parameters have resulted in a rapid increase in publications over the past 15 years (see Figure 1a).

The CAR combines features of a reactivity index (response to awakening) with aspects tied to circadian regulation (occurring roughly at the same time every 24 hours) making it a fascinating research topic. However, precisely these features also make accurate assessment of the CAR a challenging task. When relying on CAR data acquired by participants themselves (usually through saliva sampling), validity critically relies on participants closely following a timed sampling schedule, beginning with the moment of awakening. Inaccurate sample timing can occur easily and can substantially bias CAR estimates. Furthermore, a number of other methodological factors, such as accounting for covariates, the number and nature of study days and the timing of sampling, can markedly affect CAR data. While not all questions regarding the role and regulation of the CAR have been adequately solved until now, several careful investigations have examined the impact of methodological factors on accurate CAR assessment and have recommended strategies for dealing with them (described below). Unfortunately, such recommendations have not been widely implemented in published CAR research. Figure 1b provides an overview of methodological characteristics of such studies, published between 2013 and 2014. It

can be seen that the employed methodological standards varied widely between investigations, with a high number of studies falling short of previous recommendations for best practice in CAR research (e.g., objective control of sampling times: Broderick et al., 2004; Kudielka et al., 2003).

To address this, the International Society of Psychoneuroendocrinology (ISPNE) has initiated an expert panel to summarize relevant evidence and collective experience on methodological factors affecting CAR assessment. The goal of this initiative was to formulate clear consensus guidelines based on current knowledge for future studies in this evolving field of research. The present report summarizes the results of this undertaking. As a large proportion of CAR research uses salivary cortisol assessments in participants' domestic setting, a particular focus is put on methodological challenges in this research context. Given the importance of sample time accuracy, the first three sections are devoted to an in-depth discussion of this topic, including strategies to increase sampling accuracy by maximizing participant adherence. In the subsequent sections a range of further methodological factors are covered. In the final section, the derived consensus guidelines are outlined and explained.

[Please insert Figure 1 about here]

2. Cortisol awakening response

The CAR is expressed as part of normal, healthy human circadian physiology. Deviations from a typical CAR pattern are assumed to mark maladaptive neuroendocrine processes. A general review of psychosocial, psychiatric and health-related correlates of the CAR is beyond the scope of this article (reviews: Chida and Steptoe, 2009; Clow et al., 2004; Fries et al., 2009; Kudielka and Wüst, 2010). However, as a prerequisite for interpreting such data, some distinct features of the CAR need to be acknowledged. It is important to emphasize that, although historically the term 'CAR' has been used to describe different aspects of post-awakening cortisol secretion (including overall levels), only the *dynamic* of post-awakening cortisol secretion is accurately referred to as the 'CAR', i.e., cortisol *changes* occurring due to the awakening *response* (Clow et al., 2010). A rationale for this recommended terminology is laid out in section 2.2.

2.1 Description and distinctive features

The CAR represents a sharp increase in cortisol levels across the first 30–45 min following morning awakening. In healthy adults, the magnitude of the CAR was found to range between a 50–156% increase in salivary cortisol levels (Clow et al., 2004). Figure 2a depicts post-awakening cortisol profiles of children, adolescents and elderly adults from the first systematic description of the CAR by Pruessner et al. (1997).

The initial finding suggesting that the awakening process stimulates cortisol secretion provided an explanation for previous data of poor test-retest reliability in clock-time based cortisol assessments during the early morning hours (e.g., Coste et al., 1994; Schulz and Knabe, 1994) and suggested that alignment of cortisol sampling with awakening would provide a more reliable measure. This notion was soon supported by data showing improved test-retest stability of awakening-aligned post-awakening cortisol levels across a broad age range (r s between .39 and .67; Pruessner et al., 1997). These initial data were viewed as indicating that the CAR could be used as a reliable trait biomarker and, hence, investigations over the following years mainly focused on inter-individual variability in CAR profiles using cross-sectional designs. Such research revealed the CAR to be related to various physical and mental health variables, albeit with some inconsistency (review: Chida and Steptoe, 2009). More recent evidence illustrated that the CAR also exhibits considerable intra-individual variability (see 7.2). Indeed, although twin studies consistently found a moderate heritability of the CAR (Kupper et al., 2005; Wüst et al., 2000a), the expression of the CAR on a particular day has been estimated to be more influenced by state factors than by stable, trait-like influences, including genetic factors (Almeida et al., 2009; Hellhammer et al., 2007; Stalder et al., 2010b). Building on these data, research also increasingly set out to investigate state correlates of the CAR (review: Law et al., 2013).

[Please insert Figure 2 about here]

The CAR period is embedded within a well-described circadian pattern of cortisol secretion, characterized by a cortisol increase prior to awakening, the CAR period and a decline of mean cortisol levels over the remaining diurnal phase (Veldhuis et al., 1989; Weitzman et al., 1971). Importantly, there is converging evidence suggesting that the CAR is relatively distinct from earlier and later components of

circadian cortisol secretion. Sleep laboratory research revealed that the CAR is not a mere continuation of the pre-awakening cortisol increase but comprises a superimposed *response* to awakening (Wilhelm et al., 2007; see Figure 2b). In addition, the CAR was found to be unrelated to cortisol levels during the remainder of the day (Edwards et al., 2001a; Maina et al., 2009) or to a latent trait cortisol factor inferred from various diurnal samples (Doane et al., 2015).

Evidence showing that in healthy humans the CAR, but not later diurnal cortisol secretion, is sensitive to light exposure further illustrates its distinct nature: morning awakening in darkness or dim light reduces the dynamic of the CAR relative to awakening in light (Figueiro and Rea, 2012; Scheer and Buijs, 1999). Furthermore, winter awakening using a dawn simulator (gradually increasing light levels before awakening) has been associated with increased post-awakening cortisol production (Thorn et al., 2004). In rodent studies, light-induced effects on glucocorticoid secretion were absent following lesions of the suprachiasmatic nucleus (SCN) or in SCN-intact mice with adrenal sympathetic denervation (Buijs et al., 2003). Thus, an SCN-mediated extra-pituitary pathway has been implicated in regulation of the CAR, but is unlikely to affect cortisol secretion over the remainder of the day (review: Clow et al., 2010).

Besides an uncoupling from basal circadian cortisol secretion, research also consistently revealed the CAR to be unrelated to cortisol reactivity to experimentally-induced psychological stress (Bouma et al., 2009; Schmidt-Reinwald et al., 1999). This has important implications for the interpretation of CAR data, i.e., indicating its distinctness from cortisol reactivity to acute psychological stress. Interestingly, in an early study, the CAR was found to be closely related to the cortisol rise following ACTH-challenge ($r = .63$), suggesting that its expression may be influenced by the maximal capacity of the adrenal cortex to produce cortisol (Schmidt-Reinwald et al., 1999).

Overall, these data highlight the distinct nature of the CAR and suggest that its assessment provides added information that may not be derived from other cortisol measures. This together with findings of unique associations with psychosocial, cognitive and health-related parameters makes the CAR an interesting measure in PNE research. Conversely, these data also illustrate that when interpreting respective findings, researchers need to be careful not to mistake the CAR for either

a marker of general HPA axis activity/basal cortisol secretion or stress-reactive cortisol changes (Clow et al., 2010).

2.2 Main components

The CAR is a dynamic phenomenon triggered by the process of morning awakening. Strategies for quantifying post-awakening cortisol secretion need to address two main underlying components: First, the starting point of the CAR period, i.e., the first sample synchronized with the moment of awakening (S1). Second, the actual dynamic of the cortisol increase after awakening, i.e., the CAR itself, assessed at set intervals after awakening. Importantly, the two components (S1 and CAR) are often found to be inversely related, with a lower CAR following a higher S1, and vice versa (Adam et al., 2006; Bäumlner et al., 2013; Huber et al., 2006; Stalder et al., 2009, 2010b; Wilhelm et al., 2007; Wüst et al., 2000b). This relationship can likely be interpreted as an illustration of the law of initial value (Wilder, 1962).

Clow et al. (2010) reviewed neurophysiological evidence on the regulation of morning cortisol secretion and concluded that differential processes are likely to be important for the pre-awakening cortisol increase and for the CAR (Clow et al., 2010). They thus suggested that separate results should be reported for S1 (i.e., the endpoint of the pre-awakening increase) and estimates of the CAR. This is also recommended as part of this consensus report (see section 7.4). In addition, quantification strategies are in use, which combine information of S1 *and* the CAR, thus providing an index of overall cortisol secretion over the post-awakening period (e.g., the AUC_G , Pruessner et al., 2003). When using such measures, it is important to acknowledge that they are influenced by both underlying components (S1 and the CAR). Hence, it is important to refer to respective measures as reflecting total 'post-awakening cortisol concentrations' or similar, but not as measures of the CAR (Clow et al., 2010). In line with this reasoning, graphical illustrations in the present article focus on depicting S1 and a measure of the CAR, in this instance the area under the curve with respect to increase (AUC_I , Pruessner et al., 2003). A discussion of statistical approaches to quantifying the CAR is provided in section 7.4.

3. Inaccurate sampling: Prevalence and impact

The validity of CAR data critically relies on the temporal accuracy of saliva sampling across the post-awakening period. A typical sampling schedule involves taking a first

sample immediately after awakening followed by repeated assessments at specified times, e.g., at 10 or 15 min intervals over the subsequent 30–60 min. Figure 3a illustrates an exemplary CAR sampling schedule. Failure to comply with such a schedule can occur in multiple ways. In the following, we distinguish between participants (i) failing to correctly report their awakening time and/or delaying the initiation of sampling in relation to the moment of awakening and (ii) not adhering to the specified time intervals for later sampling.

3.1. Delay between awakening and initiation of sampling

The commencement of sampling immediately after awakening is crucial for accurately capturing the CAR. Table 1a provides an overview of studies examining the impact of delayed initiation of sampling after awakening. This research employed a range of methods, such as actigraphy, electrocardiography (ECG) or polysomnography (PSG), to verify participants' self-reported times of awakening (see 4.2.1 for a description of methods). In addition, two recent studies also used electronic monitoring devices to verify times of sample collection (Griefahn and Robens, 2011; Smyth et al., 2013).

[Please insert Table 1 about here]

The first description of awakening time-related sampling inaccuracy was made in a post-hoc analysis carried out on a subgroup of individuals (13.1% of the total sample) who failed to show any evidence of a positive CAR (Kupper et al., 2005). By utilizing available ECG and actigraphy data, it was revealed that these participants showed a mean delay of 42 min (range: 10–135 min) between verified and self-reported awakening times. By contrast, participants with regular CAR profiles mostly showed good correspondence between self-reported and verified awakening times (Kupper et al., 2005). Following these initial data, subsequent research confirmed that failure to correctly report the time of awakening and/or to delay the beginning of sampling after awakening is relatively common and profoundly impacts CAR estimates (see Table 1a). Across studies, mean verified awakening times preceded mean self-reported awakening times by 3.3–6.2 min and mean self-reported times of collecting S1 by 7.1–24.8 min (DeSantis et al., 2010; Dockray et al., 2008; Okun et al., 2010). A particularly striking illustration of the potential extent of such inaccuracy was reported by Griefahn and Robens (2011) who accumulated data from three studies, each

employing careful objective verification of awakening *and* sampling times across 6–8 days per individual. They found that participants delayed collecting S1 by 3–30 min on 19.3% of sampling days and by even >30 min on 14.0% of sampling days (Griefahn and Robens, 2011).

However, delaying the collection of S1 after awakening by more than 15 min results in false-high estimates of S1 and false-low estimates of the CAR. This pattern emerged both from research relying on self-reports of S1 timing (DeSantis et al., 2010; Dockray et al., 2008; Okun et al., 2010) and from studies objectively monitoring sampling times (Griefahn and Robens, 2011). Figures 3c and d exemplify the impact of 20 and 40 min sampling delays, respectively, on estimates of S1 and the CAR (AUC_i).

[Please insert Figure 3 about here]

The impact of smaller delays in sampling S1 (<15 min) has been more difficult to capture. Earlier research indicated no differences in CAR estimates between fully accurate individuals (delays <1 min) and those with 1–15 min delays (Dockray et al., 2008; Okun et al., 2010). Other studies, however, reported a trend for an attenuated CAR in individuals with 5–15 min delays (DeSantis et al., 2010) or suggested that CAR estimates already started to decrease with delays exceeding ~10 min (Griefahn and Robens, 2011). An important addition to these data comes from recent research by Smyth et al. (2013, 2015): employing careful control of awakening and sampling times in healthy participants (sampling at 5 min intervals), their findings revealed that minor delays (5–15 min) yielded estimates of an *increased* CAR and an earlier peak. This has been accounted for by the observation that cortisol levels remained relatively unchanged over the first 5–10 min post-awakening ('latent period'), with a significant increase first being detectable in the 15 min sample. These data suggest that moderate sampling delays shift the examined time window closer to the actual increase component by removing the latent period from the analysis, thus resulting in higher CAR estimates with an earlier peak component (Smyth et al., 2013, 2015). Figure 3b illustrates this notion for an exemplary 8 min sampling delay.

Inaccuracy in the commencement of sampling immediately after awakening can arise from a range of scenarios, including non-adherence due to motivational reasons (avoidance of discomfort, attending to other responsibilities, etc.). However,

observations by Griefahn & Robens (2011) and Smyth et al. (2013) suggested that non-motivational factors might also influence awakening-related inaccuracies. In both studies, considerable sampling delays occurred for S1 even though participants took later samples in close accordance with the protocol. This suggests that delayed sampling after awakening may be the primary cause of inaccurate CAR assessment and arise in well-intentioned and otherwise conscientious participants (Smyth et al., 2013). A potential explanation for this is the occurrence of sleep inertia in the immediate post-awakening period, i.e., a state of reduced cognitive and motor performance (Tassi and Muzet, 2000). Sleep inertia may increase the difficulty of adhering to requested timings and/or may impede the precise determination of the moment when one is fully awake (Clow et al., 2010; Smyth et al., 2013).

In sum, recent evidence suggests that even well-intentioned participants may not always be able to precisely identify their awakening moment. This can lead to moderate delays in collecting S1 after awakening that are sufficient to substantially bias CAR estimates. In light of this, objective verification of awakening time is necessary for obtaining valid CAR data.

3.2. Inaccurate post-awakening sampling

Besides failure to collect S1 immediately on awakening, inaccuracy may also arise from delays at subsequent sampling times. Table 1b summarizes data on the correspondence between self-reported and objectively verified times of saliva sampling in ambulatory CAR research. Two landmark studies focused on sampling accuracy during the post-awakening and the remaining diurnal sampling period (Broderick et al., 2004; Kudielka et al., 2003). Inaccurate saliva sampling occurred relatively frequently and was associated with an underestimation of the CAR. This general pattern was later confirmed by research specifically focusing on the CAR in adults (Kudielka et al., 2007b), in parents obtaining CAR samples of their preschool-age children (Smith and Dougherty, 2014) and in a large multi-ethnic sample (Golden et al., 2014). An important qualitative extension of these data was provided by findings showing that the accuracy of saliva sampling can be considerably improved by informing participants about the fact that they are being objectively monitored (Broderick et al., 2004; Kudielka et al., 2003). The potential implications of this latter finding are discussed in detail in section 4.2.4.

4. Strategies for dealing with inaccurate sampling

The following sections describe available objective monitoring strategies for ambulatory CAR research, discuss ways for dealing with identified inaccurate data and look into potential strategies in lieu of objective measures.

4.1 Objective monitoring strategies

Accurate assessment of the CAR requires objective verification of awakening and sampling times (Adam and Kumari, 2009). Ideally, such objective methods should be employed in combination with a diary log system to record self-report data of awakening and sampling times (besides other factors, such as potential covariates; see 6.3).

4.1.1 Methods for verifying awakening and sampling times

Several methods have been used to verify awakening times. Polysomnography (PSG) is considered the gold standard in sleep research (Van De Water et al., 2011) and has been used for verifying awakening times in CAR research (e.g., Gribbin et al., 2012; Griefahn & Robens, 2010, 2011; Okun et al., 2010). However, PSG is costly, labor-intensive and disruptive to participants' normal routines (Van de Water et al., 2011). Wrist actigraphy might be a more readily obtainable method as it is minimally-disruptive, relatively inexpensive and well-validated against PSG for assessing sleep parameters (e.g., Cole et al., 1992; Lichstein et al., 2006). Wrist actigraphy has been successfully used in ambulatory CAR research across several studies (e.g., Smyth et al., 2013). Another actigraphy-based approach is the use of chest-worn motility monitors that additionally record heart inter-beat-interval (IBI) data (CAR research: Kupper et al., 2005; Stalder et al., 2011). As arousing from sleep is associated with an increase in heart rate (e.g., Huikuri et al., 1994) and rapid cardiovascular activation lasting around ten heart beats (Trinder et al., 2001, 2003), available IBI data (together with actigraphy data) could further help to more precisely determine the awakening moment. Still, each of the above described methods appears suitable for the verification of awakening time in CAR research. In addition, future research may explore the use of recently developed smartphone-linked or consumer-brand devices as potential low cost alternatives for objective awakening time verification (review: Kelly et al., 2012). However, this strongly rests on the successful validation of such devices against a well-established method, such as PSG, which to date is mostly still lacking (Kelly et al., 2012; Meltzer et al., 2015).

Concerning the verification of sampling times in ambulatory research, the commonly used electronic monitoring systems have proven useful. These typically use screw top bottles that record times of bottle openings, however, boxes that record time stamps have also been devised. By storing saliva sampling devices inside the bottle and instructing participants to restrict openings to the times of

sample taking, the respective time stamps provide an indirect index of sampling times. Clearly, the use of such systems cannot fully protect against intentional misuse (e.g., participants may still take out samples from the bottle without performing the saliva sampling) but it does present the current best practice. Alternatives to this approach might arise as a consequence of modern technology. For example, smartphones with built-in cameras could be used to obtain time-stamped self-photographs ('selfies') by participants when collecting a sample to verify sampling accuracy in future studies. If adequately developed, such a strategy can equally be recommended for the verification of sampling accuracy as electronic monitoring systems.

An alternative approach that removes the need for sampling time verification is the use of automated sampling methods to assess the CAR. For one, intravenous blood sampling, when coupled with stationary PSG assessment (e.g., Wilhelm et al., 2007), ensures the accuracy of CAR assessment. However, as this research is typically restricted to the sleep laboratory, its artificial setting is associated with reduced ecological validity. This may be prevented by recently developed systems for automated sampling of subcutaneous tissue free cortisol (Bhake et al., 2013). Although clearly more demanding and invasive for participants than the self-collection of saliva samples, this approach could potentially allow the assessment of the CAR in participants' home settings in some future research.

4.1.2 Dealing with verified inaccurate data

Once information on sampling accuracy has been obtained, it can be used to reduce bias on CAR estimates through (i) data exclusion strategies and (ii) statistical modeling approaches.

For data exclusion strategies, the extent of inaccuracy is usually first calculated as the discrepancy between the scheduled and the actual/verified sampling time (Δt). The Δt -value of individual sampling times is then compared to a predetermined accuracy margin (e.g., 5, 10 or 15 min) and, in case any Δt exceeds this margin, CAR data for the respective sampling day are excluded from subsequent analyses (e.g., Kudielka et al., 2003). When using such an approach, deciding on the most suitable accuracy margin for data exclusion is difficult. Proceeding from the above reviewed findings (particularly: Smyth et al., 2013, 2015), even small time discrepancies may entail substantial bias on CAR estimates, unless they become

negligible (i.e., $\Delta t \approx 0$ min). However, narrowing accuracy margins causes a growing loss of (putatively informative) data. Consequently, any consensus about a fixed accuracy margin is necessarily a trade-off between scientific precision and practical feasibility. For example, previous research employing awakening time verification by wrist actigraphy suggests that specifying an accuracy margin of $\Delta t = 0 \pm 5$ min for S1 will yield data loss of 26–46% (DeSantis et al., 2010; Dockray et al., 2008). In addition, the selective exclusion of participants with inaccurate sampling may result in potential selection bias and reduced generalizability of results. In order to keep the percentage of classified inaccurate data (and thus data loss and potential bias) as low as possible, it is crucial to employ a full range of measures to maximize adherence (see section 5).

In the second group of strategies, verified inaccurate data are not excluded but instead the objective information on actual sampling times is *utilized* for the calculation of CAR estimates (i.e., these data are incorporated into the statistical model). Hence, this provides a more economical approach, preventing unwanted data and participant loss, and resulting concerns regarding reduced generalizability. To use such a strategy, statistical models are required that adequately describe the temporal dynamics of cortisol secretion across the CAR period. Section 7.4 provides a description of such modelling approaches in CAR research.

In sum, the decision about the most adequate strategy involves trading off considerations about scientific precision against those of practical feasibility. Researchers' primary concern should be to obtain valid, unbiased data. The use of a well-specified statistical model of the CAR that incorporates verified awakening and sampling time data fulfils this criterion and should be the method of choice. When using data exclusion strategies, achieving any confidence that CAR data are not biased requires the specification of relatively strict accuracy margins which, unfortunately, is associated with data loss. To achieve comparability between studies, we recommend that future research employs a transparent approach whereby it is clearly stated whether findings emerge when applying a strict accuracy margin of $\Delta t = 0 \pm 5$ min for each post-awakening sample, either as the sole approach or in combination with researchers' own analytical strategy (i.e., as an additional sensitivity analysis).

4.2 Are there viable strategies without the use of objective measures?

The use of objective monitoring strategies increases the costs per participant and may thus reduce the number of participants from whom endocrine data can be obtained (Adam and Kumari, 2009). This may lead researchers to consider whether alternative strategies exist that yield valid CAR data without having to employ objective measures.

4.2.1 Forced awakening

A design-based strategy to counteract problems of sampling inaccuracy is to externally awaken participants (usually through study personnel). This is a work-intensive approach that has been used with participants examined in a hospital setting (e.g., Huber et al., 2006; Nicolson and Van Diest, 2000), sleep laboratory (Wilhelm et al., 2007) or quarantined as part of a larger study (Polk et al., 2005). Recently, a variation of this approach has been employed in infants and young children who were too young to sample saliva themselves and were thus woken up by their parents to ensure the accuracy of sampling initiation (Bäumler et al., 2013, 2014a, 2014b; Stalder et al., 2013). In the latter studies, this was further complemented by objective verification of awakening and sampling times.

An argument in favor of a forced awakening approach is that current evidence suggests that the CAR is unaffected by participants' mode of awakening (spontaneous vs. externally woken; e.g., by alarm clock; Stalder et al., 2009; Wüst et al., 2000b). This makes it unlikely that forced awakening leads to fundamentally different CAR profiles than spontaneous awakening. However, a remaining danger is that participants may wake up prior to the planned wake up time. Hence, a forced awakening approach should still be complemented by objective awakening time verification. Under such a condition, forced awakening may help to yield high quality CAR data, particularly if study personnel also continue to monitor the accuracy of subsequent saliva sampling. The latter possibility may then spare the use of electronic monitoring devices to verify sampling accuracy (see 4.2.1).

Besides issues related to sample timing inaccuracy, however, researchers assessing the CAR in a hospital or sleep-laboratory setting need to be aware of the possibility of state-related confounding which can induce significant bias in CAR analyses (see 6.2). Further, as mentioned before, such conditions are associated with reduced ecological validity, which is a key advantage of sample collection in participants' home settings.

4.2.2 Increasing statistical power

Researchers may wonder whether problems of sampling inaccuracy cannot be overcome by simply increasing statistical power, e.g., in large-scale epidemiological research. Indeed, this would be the case if inaccuracy occurred randomly, i.e., not systematically related to relevant participant characteristics (cross-sectional research) or situational factors (intra-individual research). Under such circumstances, inaccurate sampling would merely increase the error of CAR estimates, which could be tackled by increasing the number of observations. However, extensive data indicate that non-adherence, a factor which is likely to strongly affect sampling inaccuracy, does not occur randomly but covaries with relevant psychological factors. For example, research has shown that adherence to medical treatments regimens is influenced by individual differences in depressiveness and/or social support (meta-analyses: DiMatteo, 2004; DiMatteo et al., 2000). Research focusing on the CAR also confirmed an inverse relationship between perceived social support and sampling inaccuracy (Kudielka et al., 2007b). In a large multi-ethnic study, inaccurate sampling of a diurnal cortisol profile (including post-awakening samples) was related to lower income, education levels, and (marginally) ethnicity (Golden et al., 2014). Furthermore, effects of sampling accuracy have been found to interact with health status, i.e., female fibromyalgia patients were less influenced by being informed about the use of objective monitoring strategies than healthy controls (Broderick et al., 2004). This indicates that inaccurate sampling is likely to co-vary with parameters that are of central interest to PNE inquiry, i.e., psychological or health-related factors. Under such circumstances, failing to control for sampling accuracy poses the eminent threat that true relationships may be obscured or false relationships may be accepted. In this case, *“(i)ncreasing the N or the number of samples collected will yield the same level and direction of error. In fact, increasing statistical power would only increase the researcher’s confidence in a false result.”* (Broderick et al., 2004, p. 648).

There is less evidence on state correlates of sampling inaccuracy from intra-individual CAR research. Broderick et al. (2004) observed no differences in sampling accuracy between weekdays and weekends, a factor frequently associated with altered CAR profiles (Kunz-Ebrecht et al., 2004b; Schlotz et al., 2004). Still, it is clearly conceivable that inaccurate sampling may co-vary with state psychological

factors relevant for PNE research (e.g., state arousal or stress, prospective memory load, sleep characteristics; Law et al., 2013). Hence, it cannot be excluded that failure to objectively control for inaccurate sampling confounds intra-individual CAR data. Again, this problem cannot be alleviated by increasing statistical power.

4.2.3 Exclusion of CAR non-responders

Inaccurate sampling has often been associated with flattened or even negative CAR profiles (e.g., Broderick et al., 2004; Dockray et al., 2008; Kupper et al., 2005). This has led to the proposition that issues of sampling inaccuracy may be addressed by excluding participants who fail to show a cortisol increase from S1 to later samples as these are 'suspected non-adherents' (Thorn et al., 2006). However, this is unlikely to be a sufficient approach. The effects of inaccurate sampling on CAR estimates are likely to be non-linear and continuous (see Figures 3b–d): compared to fully accurate sampling, small delays after awakening may first result in an overestimation of the CAR which then turns into the well-documented underestimation of the CAR with longer delays (>15–20 min). The complete absence of a post-awakening increase (i.e., a negative CAR) is likely to occur only if the delay between awakening and the initiation of sampling exceeds the peak of the underlying CAR (between 30–45 min, see Figure 3d). The exclusion of negative CAR profiles would thus eliminate only extreme cases of inaccurate sampling but not mild or moderate cases, which already have the potential to substantially bias results.

In addition, the exclusion of CAR non-responders may by itself induce bias. It is still unresolved whether, given fully accurate sampling, there are genuine occurrences of participants not showing a positive CAR. Preliminary evidence suggests that this may indeed be the case: in accurately sampled data based on objective monitoring, no increase or only a minor CAR (<2.5 nmol/L, Wüst et al., 2000b; or <1.5 nmol/L; Miller et al., 2013a) emerged on 13.1% of study days in infants (Stalder et al., 2012), on 18.0% of days in toddlers and young children (Bäumler et al., 2013), in adults on 14.7% (Dockray et al., 2008), and on 19.7% of days in healthy participants (Smyth et al., 2013). Of note, patients with brain lesions, particularly in the hippocampal formation (Buchanan et al., 2004; Wolf et al., 2005), appear to show more generally attenuated or even absent CARs (see also section 6.3). Overall, these data indicate that absent CARs may represent genuine phenomena that simply form the lower end of a distribution of CAR magnitudes.

Following this assumption, absent CARs should occur more frequently within groups that as a whole exhibit a reduced CAR profile. CAR research often investigates *group differences* (e.g., between clinical patients and control subjects), thus by definition trying to prove that one group has a lower CAR than another. Hence, a CAR non-responder exclusion strategy would bias data by systematically excluding a greater percentage of low values from one group than from another.

4.2.4 Informational strategies

In research examining diurnal cortisol levels (including post-awakening sampling), participants who were informed about their sampling accuracy being verified by electronic monitoring systems showed a 76% reduction in cumulated sampling deviations compared to 'non-informed' participants (Kudielka et al., 2003) and more days with accurate sampling (informed: 90%, non-informed: 71%; Broderick et al., 2004). Importantly, besides improving the actual rates of sampling accuracy, 'informed' participants also tended to correctly self-report their sampling times, even if they had not sampled accurately (Broderick et al., 2004; Kudielka et al., 2003). Subsequent research with participants being informed about objective monitoring mostly confirmed high accuracy of post-awakening sampling in these individuals (Griefahn and Robens, 2011; Smyth et al., 2013) although this was not the case in a recent study on a large multi-ethnic sample (see Table 1b; Golden et al., 2014).

Still, the above data may indicate the possibility that merely informing participants about the use of objective control methods could provide a strategy for obtaining reliable data on sampling accuracy (i.e., through self-reports). In this context, 'mock' strategies could be considered, with participants being told that objective monitoring strategies are being used without this actually being the case. Besides ethical considerations, it is important to note that the efficacy of such an approach has not been tested yet. This is not trivial as the effectiveness of a mock compared to a real 'informed' strategy may be reduced by several routes, such as non-verbal transmission of lower expectations from experimenters to participants (given experimenters' awareness that no objective monitoring is being used) or passing on of information about the non-functionality of objective monitoring between participants (e.g., in student populations). A potential solution may be an 'open' strategy as part of which objective monitoring is employed in a random subgroup, while all participants are told that there is a chance of being monitored (Adam and Kumari, 2009). While this avoids deceiving participants, it is unclear whether

information about the mere chance of being monitored is equally effective as certainty about this fact. Hence, without firm evidence showing the effectiveness of such an approach, it is recommended that objective monitoring is employed across all participants and is not substituted with an informational strategy. Notwithstanding, evidence clearly suggests that participants should always be informed about the use of objective monitoring strategies.

5. Maximizing adherence

Irrespective of objective monitoring, it is expedient to work towards maximizing participant adherence. Such strategies are cost-efficient as they prevent data loss through the exclusion of inaccurately sampled data and increase data quality (i.e., fully adherent data are superior to statistically inferred/corrected data). Table 2 lists strategies for maximizing adherence in CAR research. Several of these strategies are derived from the authors' collective research experience, without formal published evidence on effectiveness testing.

An important opportunity for increasing participant adherence is provided by the initial face-to-face meeting. Strategies employed during this meeting may both raise participants' motivation for being adherent and help to increase the clarity of the study procedure. An important way to raise motivation is trying to engage participants with the research goals. Besides conveying the general purpose of the study, this involves explaining the importance of being adherent in CAR research and the consequences of non-adherence. To ensure that participants fully understand the whole study procedure, it is considered important that researchers go through the protocol in detail with them and practice relevant components (e.g., the saliva sampling procedure). As part of this, it should be explained precisely what is meant by the 'moment of awakening' in order to standardize this critical aspect across participants (Adam & Kumari, 2009). We recommend that such a definition should focus on the regaining of consciousness as the central characteristic of the awakening moment (e.g., "*When you are awake, i.e., you are conscious: you know who and where you are; you are in a state that is clearly different from when you were sleeping even though you may still feel tired.*"). In addition, it should be made clear that participants should not initiate sampling after premature nightly awakenings (e.g., "*If you wake during the middle of the night and plan to go back to sleep, do not begin sampling; please only begin when you are awake for the final time before you*").

plan to get up for the day.") and that they should refrain from dozing or snoozing during the CAR sampling period (e.g., *"During this study, please do not fall back to sleep or 'doze' after your initial awakening. You can stay in bed or get out of bed but please stay awake (even if you are not fully alert) during and after the saliva sampling period."*). Besides using the initial face-to-face meeting to clarify such critical questions about sample timing, it is also important that appropriate sampling dates are negotiated by the researcher and participant and agreed as 'convenient and typical'.

Besides face-to-face contact, take-home instructions in written form should be provided. For some populations, the additional use of instructional DVDs has proven useful (e.g., Stalder et al., 2013). Overall, it is important to make instructions as explicit and practice-orientated as possible, e.g., participants may be told to place the sample kit and a pen beside the bed before going to sleep to avoid post-awakening delays through having to search for the material (Adam and Kumari, 2009). Strategies that make the collection kit more user friendly and help participants organize the collection (e.g., color coding of material) are also deemed helpful. Researchers have further had positive experiences with using reminder phone calls, emails, or text messages on the evening prior to sampling (e.g., Smyth et al., 2013). Besides reminding participants of important procedures (e.g., to wear actigraphy devices to bed), such measures also signal an extra effort made by the research team, thus again highlighting the importance of accurately following the study procedure to participants. In addition, recent studies have employed methods for reminding participants about times of post-awakening sample collection. These include automated strategies, e.g., reminder watches (Franz et al., 2013), reminding through participants' mobile phones (Garcia-Banda et al., 2014) or electronic reminders, e.g., timers that are activated by participants when taking S1 and then beep/flash at the later sampling times (e.g., Doane and Adam, 2010; Griefahn and Robens, 2011). In addition, reminder phone calls or text messaging at participants' individually predicted sampling times has been employed (e.g., Oskis et al., 2009). While such strategies may increase adherence by preventing against the forgetting of sampling, they cannot provide certainty about the accuracy of sampling. Hence, they should only be viewed as complementary approaches but cannot replace objective monitoring strategies.

[Please insert Table 2 about here]

6. Dealing with covariates

Researchers have to deal with the fact that hormone secretion is related to a large number of state and trait factors (Schlotz, 2011). Depending on the research context, these covarying factors may be considered confounders, mediators, moderators or direct variables of interest (Adam and Kumari, 2009; Kudielka et al., 2012; Schlotz, 2011). If a covariate is not of main interest, the most critical question is whether it confounds observed associations (Schlotz, 2011). Confounding is given when a covariate is related to both the CAR and the variable(s) of interest, thus creating a spurious relationship between them, and needs to be addressed. However, even if a factor is only related to the CAR but not to other variables of interest, this may increase the error variance of the model and thus reduce statistical power for detecting associations with the CAR (Schlotz, 2011). Strategies for preventing unwanted influences of covariates in ambulatory PNE research can be grouped into instructional, statistical adjustment and exclusion strategies (Kudielka et al., 2012; Schlotz, 2011).

6.1 Instructions about post-awakening behavior

Besides informing participants about the necessity to collect samples in close accordance with the specified sampling times (sections 3–5), further instructions may address participant behavior over the post-awakening period. Table 3 provides an overview of factors to be considered in this context. The most common instructions have been for participants to take nil by mouth other than water, refrain from smoking and omit cleaning their teeth (to avoid abrasion and vascular leakage into saliva) until after the final sampling (Clow et al., 2004). There is support for an influence of the first two factors: Cortisol secretion is known to be acutely influenced by caffeine and nicotine intake (review: Kudielka et al., 2009), food consumption (particularly high protein foods; e.g., Gibson et al., 1999; Rosmond et al., 2000) and blood glucose levels (Rohleder and Kirschbaum, 2007). This suggests that breakfasting (incl. caffeinated or sugared drinks and/or protein-rich foods (e.g., eggs) or smoking during the post-awakening phase may affect the CAR. By contrast, tooth brushing is at least unlikely to be associated with strong group effects as salivary cortisol levels were found to be unaffected by normal dental hygiene (Gröschl et al., 2001) or even

vigorous tooth brushing, despite the latter leading to blood leakage into saliva (Kivlighan et al., 2004). Research in children, using salivary transferrin levels as a marker of blood contamination, also concurred with the general notion that blood contamination through dental hygiene is unlikely to have a strong effect on salivary cortisol levels (Granger et al., 2007).

Similarly, current findings speak against an influence of physical behavior/activity levels in the normal range on the CAR, which has been found to be unaffected by postural changes (i.e., remaining supine vs. standing/behaving normally; Hucklebridge et al., 2002; Wilhelm et al., 2007) or the level of motility over the post-awakening period (Stalder et al., 2009). However, this does not apply to physical exercising, which is known to induce cortisol reactivity when performed above a certain intensity level (e.g., Hill et al., 2008; Kirschbaum and Hellhammer, 1994). Finally, participants' mode of awakening (spontaneous vs. alarm clock) has been found unrelated to expression of the CAR (Stalder et al., 2009; Wüst et al., 2000b). Still, as the above studies did not control for the beta error, it cannot be excluded that small effects exist but were not detected in the respective study samples.

Together, an effect on the CAR is particularly suggested for eating, drinking (caffeinated or sugared beverages), smoking or engaging in physical exercise during the post-awakening period. Concerning these behaviors, in most research contexts it is recommended that researchers (i) instruct participants to abstain from these behaviors until after they have finished post-awakening sampling. Alternatively, (ii) in case researchers feel that these restrictions impose a too severe burden on participants' normal routines (i.e., reducing willingness to participate and/or ecological validity), participants may be allowed to engage in these behaviors but should then be strongly encouraged to report this systematically (e.g., through the diary log system). This is critical to facilitate subsequent statistical adjustment for such potential influences. The latter point also applies to those behaviors without a proven influence on the CAR (mode of awakening, dental hygiene, moderate physical activity), for which it is still recommended to obtain self-report data. Furthermore, in instances when participants are allowed to eat, drink and/or brush their teeth, they should be instructed to rinse their mouth afterwards and to abstain from engaging in these behaviors in the immediate period (1–2 min) before sampling.

[Please insert Table 3 about here]

6.2 Control variables

There are several covariates of the CAR that cannot be influenced by instructions. These include factors that are difficult to standardize in ambulatory settings (e.g., ambient light levels), natural circumstance (e.g., season, menstrual cycle phase) or dispositional factors (e.g., age, sex). Table 4 shows the most important of these covariates. These are usually dealt with through statistical adjustment or by matching of study groups accordingly. For the sake of conciseness, the following part does not go into detail on between-study inconsistencies in results but focuses on practical implications for the main parameters. Also, for the ease of reading, references are only provided in the table but not in the text. For in-depth discussions of CAR covariates interested readers are referred to the respective review articles (state factors: Law et al., 2013; trait factors: Fries et al., 2009; Chida & Steptoe, 2009; sleep-related: Elder et al., 2014). Importantly, many of these data are based on CAR assessments *without* objective monitoring strategies. Hence, it cannot be excluded that previously discussed issues of inaccurate sampling might have biased these findings.

6.2.1 State covariates

Expression of the CAR on a particular day is to a large part determined by state-related factors (Almeida et al., 2009; Hellhammer et al., 2007; Stalder et al., 2010b; see also 7.2). These are often variables of interest in CAR research. When state variables are not the central focus, they should nonetheless be measured and covaried. Table 4a shows the most important state covariates of the CAR and provides selected citations. Sleep-related factors comprise an important group, with time of awakening being particularly important (earlier awakening generally being associated with an elevated CAR). Although evidence on other sleep characteristics, e.g., sleep duration or quality, is still emerging and is less consistent, it is still recommended that they should be captured as potential covariates. Conversely, mildly disturbed sleep or sleep restrictions appear to have little effect on the CAR while an influence of sleep architecture is presently unclear (Elder et al., 2014). Some high-quality data for the assessment of sleep-related factors in CAR research may

conveniently be derived from the employed methods of objective awakening time verification (actigraphy or polysomnography; see 4.2.1) and should be reported.

Higher levels of ambient light have been related to an elevated CAR. Objective assessment of light levels would be ideal (e.g., through small and unobtrusive photosensor devices; Figueiro et al., 2012), but may not be feasible for many CAR studies. Obtaining self-report data on participants' retinal light exposure (lighting of the bedroom, use of eye masks, etc.) provides an alternative strategy in this context. Further, in non-ambulatory studies (e.g., sleep laboratory), light exposure should be kept at constant levels (Elder et al., 2014). The season of assessment may also affect CAR expression (data are inconsistent) and should be considered as a covariate in studies conducted over extended time periods.

Psychosocial factors surrounding the sampling day comprise an important group of state covariates. The CAR seems to be affected by experiences over the day *prior to* CAR sampling, with a larger CAR being found after days characterized by more negative feelings (e.g., threat, lack of control or loneliness). On the study day, anticipations of a more demanding or challenging day ahead have been related to an increased CAR. This entails evidence showing a larger CAR on weekdays vs. weekend days, on days with more naturally occurring anticipated challenges/obligations and on days with experimentally induced higher prospective memory load in children and social challenge in adults. Overall, these data concur with the hypothesis that the CAR serves a function in preparing the individual for challenges of the upcoming day, which may be modulated by post-awakening anticipatory processes (Adam et al., 2006; Fries et al., 2009; Wilhelm et al., 2007). Thus, state psychosocial factors surrounding the study day should be assessed (e.g., by using a diary log system) and their influence on results examined and, potentially, adjusted for.

It is important to emphasize that adequate addressing of state covariates is also important for cross-sectional studies. Indeed, any systematic relationship between state covariates and examined individual-level variables (sociodemographic features, clinical patient status, psychosocial stress levels, etc.) can lead to confounding (Adam and Kumari, 2009; Hellhammer et al., 2007). The danger of such state-related confounding is particularly evident for the above-discussed psychosocial parameters (Stalder et al., 2010a, 2010b). For example, despite being matched carefully on sociodemographic grounds, studied groups are often examined under

different state circumstances, e.g., hospitalized patients vs. home-based controls (Gaab et al., 2005) or largely non-working patients vs. working controls (Roberts et al., 2004). Such situational differences may covary with state psychosocial factors, such as prospective memory requirements, i.e., less planning might be required for a hospital day compared to a work day (Stalder et al., 2010a). Under such circumstances, group differences in CAR profiles could be falsely attributed to participants' clinical status when they are indeed due to the differential assessment contexts. Hence, clear awareness of the possibility of state-related confounding should guide researchers' study planning. This may include the assessment of state covariates as well as design-based choices, e.g., in the above example, researchers may choose to recruit a hospital-based control group or to postpone the assessment of clinical patients until they have left the hospital setting.

[Please insert Table 4 about here]

6.2.2 Trait covariates

Several sociodemographic and health-related parameters have been identified as trait covariates of the CAR (see Table 4b). Although data are characterized by inconsistency and the magnitude of effects is generally small, it is still recommended that these factors be considered as potential confounds. A relatively consistent finding is an influence of sex; with women exhibiting a larger and more prolonged CAR than men. Age effects on the CAR have also been reported by some research, mainly occurring during specific developmental stages, such as infancy/early childhood, the onset of menarche in female adolescents and with aging. Other potentially relevant factors that have been related to the CAR are ethnicity and/or socioeconomic status as well as health-related behaviors, such as habitual smoking and heavy drinking. Furthermore, information on body fat-related anthropometric measures (body-mass-index or waist-to-hip ratio) and, in women, the use of oral contraceptives should be obtained and considered as covariates.

6.3 Exclusion criteria

The impact of some covariates is considered so severe that elimination of affected data from analyses seems necessary, i.e., by excluding participants or by postponing CAR sampling until the covariate is no longer present. Table 5 lists variables that may be considered as exclusion criteria in CAR research. Concerning acute factors

on the testing day, it is sensible that CAR sampling is postponed for participants with an active illness (e.g., influenza, common cold) until they are in a healthy state again (Adam, 2006; Adam and Kumari, 2009). Likewise, sampling should be rescheduled in participants under the acute influence of major circadian rhythm changes, e.g., shift work or jet lag (shift work: Federenko et al., 2004; Griefahn and Robens, 2010; Harris et al., 2010; Kudielka et al., 2007a; jet lag: Doane et al., 2010), to a time when they have slept at least seven nights under a constant day-night schedule.

Another potentially important factor is the female menstrual cycle. One study has reported differences in the CAR during the short period of ovulation (Wolfram et al., 2011) but not between the follicular and luteal phase (Kudielka and Kirschbaum, 2003). While CAR sampling during the ovulatory period should thus ideally be avoided, regular incorporation of objective methods (hormonal assays or ambulatory chromatographic tests) into CAR research is not always going to be feasible. Thus, researchers may estimate the time of ovulation based on self-report data (i.e., midway through the usual menstrual cycle length) and avoid sampling around this time (e.g., ± 2 days), thus accounting for inter-cycle and inter-individual variability. Alternatively, with sufficiently large sample sizes, menstrual timing may be measured and its influence on CAR estimates statistically accounted for.

Concerning longer-term influences, the use of oral glucocorticoid (GC) medication is a frequent exclusion criterion in salivary cortisol research (review: Adam and Kumari, 2009; Granger et al., 2009). Exogenous GCs are likely to affect HPA axis activity through negative feedback induction (Granger et al., 2009) and, when administered orally, may induce false-high values by cross-reacting with antibodies in the immunoassay (e.g., Perogamvros et al., 2010). Although direct evidence on the CAR is still outstanding, systemic GC administration resulted in marked attenuation of salivary cortisol levels 30 min post-awakening (Masharani et al., 2005) and should thus be an exclusion criterion in most research circumstances. Other types of medication may also influence salivary cortisol levels via direct and indirect pathways (review: Granger et al., 2009). Medication intake should thus be assessed and decisions about participant inclusion/exclusion be made after case-by-case evaluation.

The presence of HPA axis-related endocrine disorders (e.g., Morbus Cushing, Morbus Addison) is another obvious exclusion criterion in cortisol research (Adam and Kumari, 2009). Concordantly, suffering from active Cushing's disease has been

associated with an attenuated CAR (Roa et al., 2013). A detailed discussion of other endocrine disorders is beyond the scope of this article and a case-by-case evaluation should guide the inclusion/exclusion of endocrine patients. Furthermore, there is converging evidence that brain damage, particularly in the hippocampal formation (Buchanan et al., 2004; Wolf et al., 2005), is associated with marked attenuation or absence of the CAR and thus warrants participant exclusion. Likewise, particularly late pregnancy is associated with marked basal hypercortisolemia (De Weerth and Buitelaar, 2005), with at least some notion of attenuation of the CAR (Buss et al., 2009). However, given the profound biological and psychological changes that occur throughout pregnancy, CAR profiles of pregnant women should generally not be compared to those of non-pregnant women.

Finally, there is also abundant data revealing associations between the CAR and a range of mental and physical health conditions. A review of this literature is beyond the scope of this article and interested readers are referred to the respective review articles (psychosocial factors: Chida and Steptoe, 2009; health-related factors: Kudielka et al., 2012).

[Please insert Table 5 about here]

7. Procedural and design considerations

7.1 Sampling times

Figure 1b shows that the number of post-awakening samples has varied widely between CAR studies. While basic CAR research, which is usually conducted on smaller samples, has tended to employ protocols with 4–5 post-awakening samples (e.g., Edwards et al., 2001b; Pruessner et al., 1997; Wüst et al., 2000b), large-scale epidemiological research often only uses two sampling times (typically on awakening and 30–45 min post-awakening; Adam and Kumari, 2009). The choice about the number of sampling times involves a cost/accuracy trade-off: collection and assay costs increase with number of samples but a larger number of post-awakening samples also allows for more accurate estimation of CAR profiles.

An important consideration concerning protocols with only two post-awakening samples is that, although they provide a general approximation of the underlying CAR, they cannot be sure to capture the CAR peak. Specifically, as the CAR usually peaks between 30 and 45 min post-awakening (Clow et al., 2004; Smyth et al., 2015; Wüst et al., 2000b), repeated sampling over this period is necessary to capture peak

levels. The resultant issues surrounding the use of two-sample protocols are not trivial: such protocols assume that the second sample (e.g., 30 min) will be taken near the peak of most individuals, with some *random* error in peak timing. Importantly, there is evidence that individual CAR peak times may not be randomly distributed but may itself be related to individual difference variables that can be at the focus of PNE research (e.g., Lopez-Duran et al., 2014). Specifically, systematic differences in CAR peak timing have been related to executive function (Evans et al., 2012), adolescent development (Oskis et al., 2009), menstrual cycle phase/ovulation (Wolfram et al., 2011) and, at least descriptively, sex (Pruessner et al., 1997; Schlotz et al., 2004; Wüst et al., 2000b). Given such associations, two-sample protocols may lead to erroneous conclusions that a study variable is associated with the CAR magnitude whereas, indeed, it is associated with CAR peak timing.

As knowledge on the main correlates of CAR peak timing is still limited, the use of two-sample protocols cannot be recommended. In case of financial restrictions, we suggest that CAR research on adult populations employs a protocol with a minimum of three sampling points: on awakening, 30 min and 45 min. This is considered a compromise combining relatively low costs with still sufficiently detailed information. Specifically, the chosen sampling points are likely to capture mean peak concentrations of both male (~30 min) and female (~45 min) adult populations (e.g., Pruessner et al., 1997; Schlotz et al., 2004; Wüst et al., 2000b). An exception to this guideline applies to research in children and pre-pubertal adolescents who may not yet exhibit sex-specific CAR patterns (e.g., Oskis et al., 2009). Hence, in such research the use of a two-sample protocol (0 and 30 min) may be justifiable. Still, researchers need to be aware that their data may be difficult to interpret as it remains unknown whether potential relationships are seen with CAR magnitude or differential CAR peak timing.

Overall, researchers aiming to conduct more fundamental CAR research are recommended to use a 4–5 sample protocol (e.g., 0, 15, 30, 45, 60 min) providing more in-depth information on temporal dynamics of post-awakening cortisol secretion. Notwithstanding, additional research on the impact of various sampling protocols (different sampling times) on the accuracy of CAR measurement is recommended.

7.2 Number of study days (cross-sectional research)

Initial findings on the CAR suggested moderate to high test-retest stability across repeated assessment days (e.g., Edwards et al., 2001b; Wüst et al., 2000b). These data were influential in recommending the CAR (measured on one or two days only) as a reliable trait biomarker. This is illustrated by the fact that 31.7% of recent studies obtained CAR measures on only one day while 47.1% of studies measured the CAR across two days (see Figure 1b). However, over the past decade, evidence has accumulated suggesting that earlier conceptions over-emphasized trait-specificity, as the CAR is prone to substantial intra-individual variability: Hellhammer et al. (2007) employed structural equation modeling to CAR data obtained across six consecutive days per person. Their results showed that the CAR on a single day is determined to a larger extent by situational factors (61–82%) than by longer-term, trait-like components (15–37%; Hellhammer et al., 2007). These findings were confirmed by subsequent research providing comparable state-specificity estimates of 78% (Almeida et al., 2009) and 64% (Stalder et al., 2010b). It was estimated that assessments on at least six days per person may be necessary for achieving reliable trait data on the CAR (AUC; Hellhammer et al., 2007).

Concerning practical implications, it should be noted that reduced trait-specificity *per se* does not constitute a serious problem for cross-sectional research. Indeed, if state influences occur randomly this will merely increase the measurement error, which can be counteracted by increasing the sample size. Unfortunately, it was shown that the assumption of truly random state variability is easily violated (see 6.2.1). This possibility of state-related confounding poses the main threat to the integrity of CAR results. Importantly, extending the sampling period (e.g., to six or more days) does not necessarily protect against this danger. For example, in a scenario where confounding is related to systematic differences in the examination context (e.g., hospitalized patients compared to home-based controls), extending the sampling period will not solve the problem, unless the context is changed. Indeed, extending the sampling period may only yield a more reliable estimation of the influence of examination context. Hence, when trait CAR estimates are of interest, accounting for the possibility of state-related confounding on the CAR should be of highest priority (see 6.2). If this is adequately addressed, a flexible decision about the number of study days can be made, considering feasibility and power calculations (Adam and Kumari, 2009). Researchers should know, however, that, given only modest test-retest reliability of CAR assessments, the detectable effect sizes for

associations with trait measures are severely restricted for single-day CAR data and will increase with each additional day from which trait estimates are derived (e.g., Frost and Thompson, 2000). Hence, if practically feasible, it is recommended that CAR data are obtained over two or more sampling days. Given the importance of weekday-weekend differences in the CAR (see 6.2), we recommend that days be evenly distributed across the week (e.g., sampling on two weekdays and one weekend day) in cross-sectional research.

7.3 Sample storage and cortisol analysis

Measurement of the CAR in ambulatory settings has been facilitated by the possibility to assess free cortisol levels using convenient saliva sampling (Hellhammer et al., 2009; Kirschbaum and Hellhammer, 1989, 1994). A general review of salivary cortisol assessment methods is beyond the scope of this text and can be found in the following reviews (Hellhammer et al., 2009; Kudielka et al., 2012).

Given that the CAR is often assessed in ambulatory settings, the storage of samples after collection is an important topic. Research suggests that salivary cortisol is relatively stable at room temperature for only a short time period (< 5 days) after which concentrations begin to decline (e.g., Clements and Richard Parker, 1998; Gröschl et al., 2001; Whembolua et al., 2006). It is thus recommended that participants be instructed to place saliva samples in their home freezer immediately after data collection (i.e., every morning) and to return them to the lab as soon as possible. For this, transport of samples under cooled conditions is preferable, although potential effects of short-term thawing during transport are likely to be small; e.g., effects on cortisol levels have been observed after five, but not after two thawing cycles (Gröschl et al., 2001).

Long-term storage of saliva samples should be at temperatures of -20°C or lower. Importantly, even with storage at -20°C , salivary cortisol levels have been found to decline over a period of nine months (Kudielka et al., 2012). These data suggest that particularly in longitudinal CAR research with collection periods stretched out across several years, researchers should strive for a prompt analysis of samples (ideally, within 0–6 months of sample collection).

Concerning the measurement of salivary cortisol concentrations, most analyses are performed by immunoassays. These are fast, relatively cheap and can be performed by most biochemical laboratories. Although cross-reactivity with other

analytes cannot be excluded, salivary cortisol immunoassays are generally deemed valid procedures and should thus be sufficient for assessing the CAR in most contexts. Concerning data interpretation, it should be known that immunoassays tend to overestimate cortisol levels in saliva (Jönsson et al., 2003) with substantial differences between assays (Kirschbaum and Hellhammer, 1989; Miller et al., 2013b). More accurate information on salivary cortisol levels may be obtained by use of chromatography-based methods (e.g., Gao et al., 2015). However, these are more work-intensive and expensive and are thus only indicated when the assessment of additional analytes is of interest (e.g., salivary cortisone; Perogamvros et al., 2010).

7.4 Statistical considerations

Prior to conducting inferential statistics on CAR data, cortisol values should be screened for distributional properties and outliers as a first step (see Schlotz, 2011). Such preprocessing procedures need to account for the fact that salivary cortisol data are usually positively skewed. To enable the use of general linear model-based analyses, it is thus recommended that appropriate transformation techniques are applied to approximate a normal distribution (Miller and Plessow, 2013; Schlotz, 2011). Besides distributional properties, some extreme outlying cortisol values may be present which, if unaccounted for, would disproportionately influence the results of parametric statistics (Schlotz, 2011). A useful convention for this purpose is to define such outliers as log-transformed values that are located more than three standard deviations from the mean. Such values may be dealt with by excluding these subjects or observations from analyses (*trimming*) or by replacing them with the values of a preset margin, e.g., the upper or lower 5% percentiles (*winsorizing*; Schlotz, 2011).

Following adequate preprocessing procedures, inferential statistics can be applied to CAR data. If the purpose of such analyses is to investigate group differences in cortisol profiles, a rather straightforward approach is the use of simple repeated-measures analysis of variance (rANOVA). Here, a main effect of time is usually considered as evidence for the presence of a significant CAR, whereas time-by-group interactions are examined for investigating group differences in CAR profiles. In addition, researchers are often interested in investigating continuous associations with the CAR which requires a single estimate of the CAR (to be included in regression-type models). For this purpose, CAR summary indicators are often computed, such as AUC-based measures, means or change scores

(Fekedulegn et al., 2007). Although a large number of such summary indicators can be applied to reactive cortisol data, these measures tend to be highly interrelated and have been shown to represent two main underlying components, i.e., 'total cortisol production' and 'change in cortisol levels' (Khoury et al., 2015). As explicated above (see section 2.2), only measures of the latter group, capturing the dynamic of post-awakening cortisol changes (e.g., AUC_I, mean increase, baseline-to-peak increase), should be referred to as measures of the CAR. In addition, besides such CAR-summary indicators, results for the first sample on awakening (S1) should be reported (Clow et al., 2010; see section 2.2). Furthermore, given the inverse relationship between S1 and the CAR, exploratory analyses on the CAR may be conducted with the level of S1 being statistically adjusted for.

A weakness of the above approaches (rANOVA, summary indicators) is that they suffer from case-wise data exclusion (i.e., data of whole study days need to be discarded if a single data point is missing). Concerning the CAR, this problem is further aggravated by the common use of exclusion strategies for dealing with inaccurately sampled data (see section 4.1.2). The use of multiple imputation strategies comprises a possibility for dealing with data that is missing-at-random (see Schafer and Graham, 2002). Another, more convenient method for addressing such problems is the use of hierarchical (also known as mixed-effects or multilevel) regression modeling (see Raudenbush and Bryk, 2002, for a general introduction). Hierarchical regression models are also capable of accounting for the continuous dynamics of time (so called growth curves; Llabre et al., 2004) and thus provide more flexibility in handling inaccurately sampled data or missing values. In addition, hierarchical models can adequately deal with manifest problems of heteroscedasticity and auto-correlations in the error structure of cortisol data (see also Kudielka et al., 2012) and allow for the simultaneous processing of between- and within-individual information (e.g., Hruschka et al., 2005). Furthermore, hierarchical regression models are flexibly extendible to account for different phases of cortisol secretion and/or varying cortisol peaks (piecewise growth curve models; Lopez-Duran et al., 2014; Schlotz et al., 2011; Willoughby et al., 2007), as well as for qualitative differences in cortisol secretion patterns, like non-responses to awakening (growth mixture models; Miller et al., 2013b).

8. Summary and guidelines

The present review shows that in order to derive meaningful data from CAR research, attention needs to be paid to methodological detail to prevent the danger of obtaining biased results. The variant methodological standards employed in past research (see Figure 1b) are thus likely to have contributed to the inconsistency observed in this field. Hence, it is strongly recommended that researchers follow the guidelines described here for obtaining valid CAR data in future research.

Table 6 summarizes the derived consensus guidelines. The control of sample timing accuracy takes up a chief position within this realm. Importantly, the reviewed data provides a strong case for recommending that future research should always employ objective strategies for the verification of awakening *and* sampling times to obtain valid data, even if that means testing fewer participants because of the increased per-subject costs. Further, alternative strategies that may reduce costs in future research are provided (e.g., use of time-stamped self-photographs to monitor sampling accuracy; see 4.1.1). Finally, if cost considerations preclude using objective monitoring across a whole cohort, the restriction of endocrine assessments to a random subsample of participants provides a valuable alternative (Adam and Kumari, 2009). In parallel, when aiming to reduce costs in large-scale research, compromises in other procedural aspects (e.g., the number of post-awakening samples or assessment days) are likely to have less detrimental implications for the validity of results than cutting down monitoring standards (see 7.2). Besides objective monitoring, we also recommend the complementary use of diary log systems to derive subjective reference values against which to compare objective data. Such diary log systems should also provide a place for reporting any atypical events or violations with instructions.

[Please insert Table 6 about here]

On-site strategies, such as forced-awakening and/or the monitoring of sampling by study personnel, can be useful if participants are examined in a stationary setting (e.g., hospital, sleep laboratory). However, while on-site monitoring of saliva sampling can substitute the use of electronic monitoring bottles, forced awaking cannot replace objective awakening time verification as participants may still wake up prior to the external wake-up time (see 4.1.1). Further, we do not recommend the exclusion of participants on the basis of their cortisol data (i.e., CAR non-responders) as this is

likely an insufficient strategy which, contrarily, may even in itself induce bias (see 4.1.3). For that matter, strategies for dealing with inaccurate sampling have to be rooted in the use of objective monitoring strategies. Two general approaches can be applied involving either the exclusion of data outside a predetermined accuracy margin or the use of statistical modeling techniques to incorporate objectively verified time data. Although researchers may choose their individually preferred approach, to achieve some level of consistency and comparability between studies, we recommend that results of (additional) sensitivity analyses are provided with CAR data being excluded on the basis of a strict accuracy margin of $\Delta t = 0 \pm 5$ min for each post-awakening sample.

Evidence clearly suggests that participants should be informed about the use of objective monitoring devices as this improves adherence (note: this should not be used as a substitute for actual objective monitoring). Furthermore, other strategies have been outlined based on researchers' collective experience that may maximize adherence by positively affecting participant motivation and/or improving the clarity of the assessment procedure (see 5). Instructing participants should also be used as a first step towards limiting the influence of covariates on the CAR. This particularly concerns instructing participants to refrain from eating or drinking (except for water), smoking and exercising during the morning sampling period. Conversely, current evidence does not indicate that participants need to abstain from tooth brushing or have to be instructed concerning their mode of awakening or post-awakening activity level (see 6.1 and Table 3). Notwithstanding, participants should always be encouraged to report engaging in any such post-awakening behaviors which may then be considered as potential confounds.

Adequate assessment of state and trait covariates of the CAR is crucial for preventing the possibility of confounding. Practical guidelines for dealing with the most important such covariates have been provided (see 6.2 & 6.3). An often underestimated danger is state-related confounding in cross-sectional CAR research. Researchers should consider whether systematic differences exist in the examination context of participants and adjust for this possibility in their study planning (see 6.2). In addition, the effects of some acute or long-term covariates may be prevented by the postponement of sampling (for state factors) or participant exclusion (for trait factors). Although we have made suggestions for how to deal with such parameters

(see 6.3), the decision about the most adequate strategy needs to consider the individual study context.

Concerning the number of post-awakening sampling times, protocols including at least three assessments (on awakening, 30 min and 45 min) are recommended for research in adult populations to allow distinguishing between effects on CAR magnitude and the CAR peak timing. In research on children and pre-pubertal adolescents, the use of two-sample protocols (0 and 30 min) may be justifiable. In smaller scale, fundamental research on the CAR more extended sampling protocols are recommended to provide more comprehensive data on temporal dynamics of post-awakening cortisol secretion. In cross-sectional research, assessment of the CAR on an increased number of study days seems necessary to reliably capture trait components of the CAR. Specifically, researchers deciding to assess the CAR on only a single day need to be aware of the fact that 61–82% of variability in their results is likely to be due to state-specific factors. Such variability will decrease statistical power and restrict the detectable effect sizes for associations with trait measures in between-subject analyses. However, as long as state covariates are adequately dealt with this is unlikely to bias respective results. Hence, researchers need to trade off pros and cons of increasing the number of study days against those of increasing the sample size in terms of feasibility and cost considerations. A useful strategy in many cross-sectional study contexts is to assess the CAR on two weekdays and one weekend day.

Finally, regarding the reporting of post-awakening cortisol data, we recommend that results are provided separately for the two main assumed components: S1 (the first sample on awakening) and the CAR (the dynamic of the post-awakening cortisol response). Given the inverse association between S1 and the CAR, it may further be explored to control for S1 when analyzing the CAR. Composite measures reflecting total post-awakening cortisol levels (e.g., AUC_G) may be reported as additional information but should be referred to as reflecting total 'post-awakening cortisol concentrations', or similar, but not as estimates of the CAR. The term 'cortisol awakening response' should be restricted to measures of the dynamic of the post-awakening increase (e.g., AUC_I , 0–30 delta, $MnInc$).

Overall, while the methodological considerations of this consensus statement may appear onerous, adherence to these steps is believed to create more powerful research designs that will yield reliable and reproducible data, thus increasing

researchers' confidence in the validity of their findings and helping to further advance understanding in this evolving field of research.

References:

- Adam, E.K., 2006. Transactions among adolescent trait and state emotion and diurnal and momentary cortisol activity in naturalistic settings. *Psychoneuroendocrinology* 31, 664–679. doi:10.1016/j.psyneuen.2006.01.010
- Adam, E.K., Hawkey, 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. U. S. A.* 103, 17058–17063. doi:10.1073/pnas.0605053103
- Adam, E.K., Kumari, M., 2009. Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology* 34, 1423–1436. doi:10.1016/j.psyneuen.2009.06.011
- 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. doi:10.1037/a0017910
- Badrick, E., Bobak, M., Britton, A., Kirschbaum, C., Marmot, M., Kumari, M., 2008. The relationship between alcohol consumption and cortisol secretion in an aging cohort. *J. Clin. Endocrinol. Metab.* 93, 750–757. doi:10.1210/jc.2007-0737
- Badrick, E., Kirschbaum, C., Kumari, M., 2007. The relationship between smoking status and cortisol secretion. *J. Clin. Endocrinol. Metab.* 92, 819–824. doi:10.1210/jc.2006-2155
- Bäumler, D., Kirschbaum, C., Kliegel, M., Alexander, N., Stalder, T., 2013. The cortisol awakening response in toddlers and young children. *Psychoneuroendocrinology* 38, 2485–2492. doi:10.1016/j.psyneuen.2013.05.008
- Bäumler, D., Kliegel, M., Kirschbaum, C., Miller, R., Alexander, N., Stalder, T., 2014a. Effect of a naturalistic prospective memory-related task on the cortisol awakening response in young children. *Biol. Psychol.* 103, 24–26. doi:10.1016/j.biopsycho.2014.08.005
- Bäumler, D., Voigt, B., Miller, R., Stalder, T., Kirschbaum, C., Kliegel, M., 2014b. The relation of the cortisol awakening response and prospective memory functioning in young children. *Biol. Psychol.* 99, 41–46. doi:10.1016/j.biopsycho.2014.02.011
- Bhake, R.C., Leendertz, J.A., Linthorst, A.C.E., Lightman, S.L., 2013. Automated 24-hours sampling of subcutaneous tissue free cortisol in humans. *J. Med. Eng. Technol.* 37, 180–4. doi:10.3109/03091902.2013.773096
- Bouma, E.M.C., Riese, H., Ormel, J., Verhulst, F.C., Oldehinkel, A.J., 2009. Adolescents' cortisol responses to awakening and social stress; effects of gender, menstrual phase and oral contraceptives. The TRAILS study. *Psychoneuroendocrinology* 34, 884–893. doi:10.1016/j.psyneuen.2009.01.003
- Broderick, J.E., Arnold, D., Kudielka, B.M., Kirschbaum, C., 2004. Salivary cortisol sampling compliance: comparison of patients and healthy volunteers. *Psychoneuroendocrinology* 29, 636–650.
- Buchanan, T.W., Kern, S., Allen, J.S., Tranel, D., Kirschbaum, C., 2004. Circadian regulation of cortisol after hippocampal damage in humans. *Biol. Psychiatry* 56, 651–656. doi:10.1016/j.biopsycho.2004.08.014

- Buijs, R.M., van Eden, C.G., Goncharuk, V.D., Kalsbeek, A., 2003. The biological clock tunes the organs of the body: timing by hormones and the automatic nervous system. *J. Endocrinol.* 177, 17-26.
- Buss, C., Entringer, S., Reyes, J.F., Chicz-DeMet, A., Sandman, C.A., Waffarn, F., Wadhwa, P.D., 2009. The maternal cortisol awakening response in human pregnancy is associated with the length of gestation. *Am. J. Obstet. Gynecol.* 201. doi:10.1016/j.ajog.2009.06.063
- Chida, Y., Steptoe, A., 2009. Cortisol awakening response and psychosocial factors: A systematic review and meta-analysis. *Biol. Psychol.* 80, 265–278. doi:10.1016/j.biopsycho.2008.10.004
- Chrousos, G.P., 2009. Stress and disorders of the stress system. *Nat. Rev. Endocrinol.* 5, 374–381. doi:10.1038/nrendo.2009.106
- Clements, A.D., Richard Parker, C., 1998. The relationship between salivary cortisol concentrations in frozen versus mailed samples. *Psychoneuroendocrinology* 23, 613–616. doi:10.1016/S0306-4530(98)00031-6
- 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. doi:10.1016/j.neubiorev.2009.12.011
- Clow, A., Thorn, L., Evans, P., Hucklebridge, F., 2004. The awakening cortisol response: methodological issues and significance. *Stress* 7, 29–37. doi:10.1080/10253890410001667205
- Cole, R.J., Kripke, D.F., Gruen, W., Mullaney, D.J., Gillin, J.C., 1992. Automatic sleep wake identification from wrist activity. *Sleep* 15, 461–469.
- Coste, J., Strauch, G., Letrait, M., Bertagna, X., 1994. Reliability of hormonal levels for assessing the hypothalamic-pituitary-adrenocortical system in clinical-pharmacology. *Br. J. Clin. Pharmacol.* 38, 474–479.
- De Weerth, C., Buitelaar, J.K., 2005. Cortisol awakening response in pregnant women. *Psychoneuroendocrinology* 30, 902–907. doi:10.1016/j.psyneuen.2005.05.003
- DeSantis, A.S., Adam, E.K., Mendelsohn, K.A., Doane, L.D., 2010. Concordance between self-reported and objective wakeup times in ambulatory salivary cortisol research. *Int. J. Behav. Med.* 17, 74–78. doi:10.1007/s12529-009-9053-5
- DiMatteo, M.R., 2004. Social support and patient adherence to medical treatment: a meta-analysis. *Heal. Psychol.* 23, 207–218. doi:10.1037/0278-6133.23.2.207
- DiMatteo, M.R., Lepper, H.S., Croghan, T.W., 2000. Depression is a risk factor for noncompliance with medical treatment. *Arch. Intern. Med.* 160, 2101. doi:10.1001/archinte.160.14.2101
- Doane, L.D., Adam, E.K., 2010. Loneliness and cortisol: momentary, day-to-day, and trait associations. *Psychoneuroendocrinology* 35, 430–441. doi:10.1016/j.psyneuen.2009.08.005

- 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. doi:10.1016/j.psyneuen.2015.01.017
- Doane, L.D., Kremen, W.S., Eaves, L.J., Eisen, S.A., Hauger, R., Hellhammer, D., Levine, S., Lupien, S., Lyons, M.J., Mendoza, S., Prom-Wormley, E., Xian, H., York, T.P., Franz, C.E., Jacobson, K.C., 2010. Associations between jet lag and cortisol diurnal rhythms after domestic travel. *Health Psychol.* 29, 117–123. doi:10.1037/a0017865
- Dockray, S., Bhattacharyya, M.R., Molloy, G.J., Steptoe, A., 2008. The cortisol awakening response in relation to objective and subjective measures of waking in the morning. *Psychoneuroendocrinology* 33, 77–82.
- Edwards, S., Clow, A., Evans, P., Hucklebridge, F., 2001a. Exploration of the awakening cortisol response in relation to diurnal cortisol secretory activity. *Life Sci.* 68, 2093–2103.
- Edwards, S., Evans, P., Hucklebridge, F., Clow, A., 2001b. Association between time of awakening and diurnal cortisol secretory activity. *Psychoneuroendocrinology* 26, 613–622.
- Elder, G.J., Wetherell, M. A., Barclay, N.L., Ellis, J.G., 2014. The cortisol awakening response - applications and implications for sleep medicine. *Sleep Med. Rev.* 18, 195–204. doi:10.1016/j.smr.2013.05.001
- Evans, P., Hucklebridge, F., Loveday, C., Clow, A., 2012. The cortisol awakening response is related to executive function in older age. *Int. J. Psychophysiol.* 84, 201–204. doi:10.1016/j.ijpsycho.2012.02.008
- Evans, P.D., Fredhoi, C., Loveday, C., Hucklebridge, F., Aitchison, E., Forte, D., Clow, A., 2011. The diurnal cortisol cycle and cognitive performance in the healthy old. *Int. J. Psychophysiol.* 79, 371–377. doi:10.1016/j.ijpsycho.2010.12.006
- Federenko, I., Wüst, S., Hellhammer, D.H., Dechoux, R., Kumsta, R., Kirschbaum, C., 2004. Free cortisol awakening responses are influenced by awakening time. *Psychoneuroendocrinology* 29, 174–184. doi:10.1016/S0306-4530(03)00021-0
- Fekedulegn, D.B., Andrew, M.E., Burchfiel, C.M., Violanti, J.M., Hartley, T.A., Charles, L.E., Miller, D.B., 2007. Area under the curve and other summary indicators of repeated waking cortisol measurements. *Psychosom. Med.* 69, 651–659. doi:10.1097/PSY.0b013e31814c405c
- Figueiro, M.G., Hamner, R., Bierman, A., Rea, M.S., 2012. Comparisons of three practical field devices used to measure personal light exposures and activity levels. *Light. Res. Technol.* 45, 421–434. doi:10.1177/1477153512450453
- Figueiro, M.G., Rea, M.S., 2012. Short-wavelength light enhances cortisol awakening response in sleep-restricted adolescents. *Int. J. Endocrinol.* 2012. doi:10.1155/2012/301935
- Franz, C.E., Spoon, K., Thompson, W., Hauger, R.L., Hellhammer, D.H., Jacobson, K.C., Lupien, S., Lyons, M.J., McCaffery, J., McKenzie, R., Mendoza, S.P., Panizzon, M.S., Ramundo, A., Shahroudi, A., Kremen, W.S., 2013. Adult cognitive ability and socioeconomic status as mediators of the effects of childhood disadvantage on salivary

- cortisol in aging adults. *Psychoneuroendocrinology* 38, 2127–2139.
doi:10.1016/j.psyneuen.2013.04.001
- Fries, E., Dettenborn, L., Kirschbaum, C., 2009. The cortisol awakening response (CAR): facts and future directions. *Int. J. Psychophysiol.* 72, 67–73.
- Frost, C., Thompson, S.G., 2000. Correcting for regression dilution bias: comparison of methods for a single predictor variable. *J. R. Stat. Soc. Ser. A* 163, 173–189.
doi:10.1111/1467-985X.00164
- Gaab, J., Baumann, S., Budnoik, A., Gmünder, H., Hottinger, N., Ehlert, U., 2005. Reduced reactivity and enhanced negative feedback sensitivity of the hypothalamus-pituitary-adrenal axis in chronic whiplash-associated disorder. *Pain* 119, 219–224.
doi:10.1016/j.pain.2005.10.001
- Gao, W., Stalder, T., Kirschbaum, C., 2015. Quantitative analysis of estradiol and six other steroid hormones in human saliva using a high throughput liquid chromatography–tandem mass spectrometry assay. *Talanta* 143, 353–358.
- Garcia-Banda, G., Chellew, K., Fornes, J., Perez, G., Servera, M., Evans, P., 2014. Neuroticism and cortisol: pinning down an expected effect. *Int. J. Psychophysiol.* 91, 132–138. doi:10.1016/j.ijpsycho.2013.12.005
- Gibson, E.L., Checkley, S., Papadopoulos, A., Poon, L., Daley, S., Wardle, J., 1999. Increased salivary cortisol reliably induced by a protein-rich midday meal. *Psychosom. Med.* 61, 214–224.
- Golden, S.H., Sánchez, B.N., DeSantis, A.S., Wu, M., Castro, C., Seeman, T.E., Tadros, S., Shrager, S., Diez Roux, A. V., 2014. Salivary cortisol protocol adherence and reliability by socio-demographic features: The Multi-Ethnic Study of Atherosclerosis. *Psychoneuroendocrinology* 43, 30–40. doi:10.1016/j.psyneuen.2014.01.025
- Granger, D.A., Cicchetti, D., Rogosch, F.A., Hibel, L.C., Teisl, M., Flores, E., 2007. Blood contamination in children's saliva: prevalence, stability, and impact on the measurement of salivary cortisol, testosterone, and dehydroepiandrosterone. *Psychoneuroendocrinology* 32, 724–733. doi:10.1016/j.psyneuen.2007.05.003
- Granger, D.A., Hibel, L.C., Fortunato, C.K., Kapelewski, C.H., 2009. Medication effects on salivary cortisol: tactics and strategy to minimize impact in behavioral and developmental science. *Psychoneuroendocrinology* 34, 1437–1448.
doi:10.1016/j.psyneuen.2009.06.017
- 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. *Dev. Psychobiol.* 54, 412–422.
doi:10.1002/dev.20599
- Griefahn, B., Robens, S., 2010. The normalization of the cortisol awakening response and of the cortisol shift profile across consecutive night shifts - an experimental study. *Psychoneuroendocrinology* 35, 1501–1509. doi:10.1016/j.psyneuen.2010.05.002
- Griefahn, B., Robens, S., 2011. Cortisol awakening response - are sampling delays of 15 minutes acceptable? *Int. J. Psychophysiol.* 82, 202–205.
doi:10.1016/j.ijpsycho.2011.08.005

- Gröschl, M., Wagner, R., Rauh, M., Dörr, H.G., 2001. Stability of salivary steroids: the influences of storage, food and dental care. *Steroids* 66, 737–741.
- Hajat, A., Diez-Roux, A., Franklin, T.G., Seeman, T., Shrager, S., Ranjit, N., Castro, C., Watson, K., Sanchez, B., Kirschbaum, C., 2010. Socioeconomic and race/ethnic differences in daily salivary cortisol profiles: The Multi-Ethnic Study of Atherosclerosis. *Psychoneuroendocrinology* 35, 932–943. doi:10.1016/j.psyneuen.2009.12.009
- Harris, A., Ursin, H., Murison, R., Eriksen, H.R., 2007. Coffee, stress and cortisol in nursing staff. *Psychoneuroendocrinology* 32, 322–330. doi:10.1016/j.psyneuen.2007.01.003
- Harris, A., Waage, S., Ursin, H., Hansen, Å.M., Bjorvatn, B., Eriksen, H.R., 2010. Cortisol, reaction time test and health among offshore shift workers. *Psychoneuroendocrinology* 35, 1339–1347. doi:10.1016/j.psyneuen.2010.03.006
- Hellhammer, D.H., Wüst, S., Kudielka, B.M., 2009. Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology* 34, 163–171. doi:10.1016/j.psyneuen.2008.10.026
- Hellhammer, J., Fries, E., Schweisthal, O.W., Schlotz, W., Stone, A.A., Hagemann, D., 2007. Several daily measurements are necessary to reliably assess the cortisol rise after awakening: state- and trait components. *Psychoneuroendocrinology* 32, 80–86.
- Hill, E.E., Zacki, E., Battaglini, C., Viru, M., Viru, A., Hackney, A.C., 2008. Exercise and circulating cortisol levels: the intensity threshold effect. *J. Endocrinol. Invest.* 31, 587–591. doi:4732 [pii]
- Hruschka, D.J., Kohrt, B.A., Worthman, C.M., 2005. Estimating between- and within-individual variation in cortisol levels using multilevel models. *Psychoneuroendocrinology* 30, 698–714. doi:10.1016/j.psyneuen.2005.03.002
- Huber, T.J., Issa, K., Schik, G., Wolf, O.T., 2006. The cortisol awakening response is blunted in psychotherapy inpatients suffering from depression. *Psychoneuroendocrinology* 31, 900–904. doi:10.1016/j.psyneuen.2006.03.005
- Hucklebridge, F., Mellins, J., Evans, P., Clow, A., 2002. The awakening cortisol response - no evidence for an influence of body posture. *Life Sci.* 71, 639–646.
- Huikuri, H. V, Niemelä, M.J., Ojala, S., Rantala, A., Ikäheimo, M.J., Airaksinen, K.E., 1994. Circadian rhythms of frequency domain measures of heart rate variability in healthy subjects and patients with coronary artery disease. Effects of arousal and upright posture. *Circulation* 90, 121–126. doi:10.1161/01.cir.90.1.121
- Jönsson, B.A.G., Malmberg, B., Amilon, Å., Helene Garde, A., Ørbæk, P., 2003. Determination of cortisol in human saliva using liquid chromatography–electrospray tandem mass spectrometry. *J. Chromatogr. B.* doi:10.1016/S1570-0232(02)00753-5
- Kelly, J.M., Strecker, R.E., Bianchi, M.T., 2012. Recent developments in home sleep-monitoring devices. *ISRN Neurol.* doi:10.5402/2012/768794
- Khoury, J.E., Gonzales, A., Levitan, R.D., Pruessner, J.C., Chopra, K., Basile, V.S., Masellis, M., Goodwill, A., Atkinson, L., 2015. Summary cortisol reactivity indicators: interrelations and meaning. *Neurobiol. Stress* 2, 34–43.

- Kirschbaum, C., Hellhammer, D.H., 1989. Salivary cortisol in psychobiological research - an overview. *Neuropsychobiology* 22, 150–169. doi:10.1159/000118611
- Kirschbaum, C., Hellhammer, D.H., 1994. Salivary cortisol in psychoneuroendocrine research: recent developments and applications. *Psychoneuroendocrinology* 19, 313–333. doi:10.1016/0306-4530(94)90013-2
- Kivlighan, K.T., Granger, D.A., Schwartz, E.B., Nelson, V., Curran, M., Shirtcliff, E.A., 2004. Quantifying blood leakage into the oral mucosa and its effects on the measurement of cortisol, dehydroepiandrosterone, and testosterone in saliva. *Horm. Behav.* 39-46, 39–46.
- Kondratova, A.A., Kondratov, R. V., 2012. The circadian clock and pathology of the ageing brain. *Nat. Rev. Neurosci.* 13, 325–335. doi:10.1038/nrn3208
- Kudielka, B.M., Broderick, J.E., Kirschbaum, C., 2003. Compliance with saliva sampling protocols: electronic monitoring reveals invalid cortisol daytime profiles in noncompliant subjects. *Psychosom. Med.* 65, 313–319. doi:10.1097/01.PSY.0000058374.50240.BF
- Kudielka, B.M., Buchtal, J., Uhde, A., Wüst, S., 2007a. Circadian cortisol profiles and psychological self-reports in shift workers with and without recent change in the shift rotation system. *Biol. Psychol.* 74, 92–103. doi:10.1016/j.biopsycho.2006.08.008
- Kudielka, B.M., Gierens, A., Hellhammer, D.H., Wüst, S., Schlotz, W., 2012. Salivary cortisol in ambulatory assessment--some dos, some don'ts, and some open questions. *Psychosom. Med.* 74, 418–431. doi:10.1097/PSY.0b013e31825434c7
- Kudielka, B.M., Hawkey, L.C., Adam, E.K., Cacioppo, J.T., 2007b. Compliance with ambulatory saliva sampling in the Chicago health, aging, and social relations study and associations with social support. *Ann. Behav. Med.* 34, 209–216. doi:10.1007/BF02872675
- Kudielka, B.M., Hellhammer, D.H., Wüst, S., 2009. Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge. *Psychoneuroendocrinology* 34, 2–18.
- Kudielka, B.M., Wüst, S., 2010. Human models in acute and chronic stress: assessing determinants of individual hypothalamus-pituitary-adrenal axis activity and reactivity. *Stress* 13, 1–14. doi:10.3109/10253890902874913
- Kudielka, M., Kirschbaum, C., 2003. Awakening cortisol responses are influenced by health status and awakening time but not by menstrual cycle phase. *Psychoneuroendocrinology* 28, 35–47.
- Kumari, M., Badrick, E., Ferrie, J., Perski, A., Marmot, M., Chandola, T., 2009. Self-reported sleep duration and sleep disturbance are independently associated with cortisol secretion in the Whitehall II study. *J. Clin. Endocrinol. Metab.* 94, 4801–4809. doi:10.1210/jc.2009-0555
- Kunz-Ebrecht, S.R., Kirschbaum, C., Marmot, M., Steptoe, A., 2004a. Differences in cortisol awakening response on work days and weekends in women and men from the Whitehall II cohort. *Psychoneuroendocrinology* 29, 516–528. doi:10.1016/S0306-4530(03)00072-6

- Kunz-Ebrecht, S.R., Kirschbaum, C., Steptoe, A., 2004b. Work stress, socioeconomic status and neuroendocrine activation over the working day. *Soc. Sci. Med.* 58, 1523–1530.
- Kupper, N., De Geus, E.J.C., Van Den Berg, M., Kirschbaum, C., Boomsma, D.I., Willemsen, G., 2005. Familial influences on basal salivary cortisol in an adult population. *Psychoneuroendocrinology* 30, 857–868. doi:10.1016/j.psyneuen.2005.04.003
- Lasikiewicz, N., Hendrickx, H., Talbot, D., Dye, L., 2008. Exploration of basal diurnal salivary cortisol profiles in middle-aged adults: Associations with sleep quality and metabolic parameters. *Psychoneuroendocrinology* 33, 143–151. doi:10.1016/j.psyneuen.2007.10.013
- Law, R., Hucklebridge, F., Thorn, L., Evans, P., Clow, A., 2013. State variation in the cortisol awakening response. *Stress* 16, 483–92. doi:10.3109/10253890.2013.817552
- Lichstein, K.L., Stone, K.C., Donaldson, J., Nau, S.D., Soeffing, J.P., Murray, D., Lester, K.W., Aguillard, R.N., 2006. Actigraphy validation with insomnia. *Sleep* 29, 232–239.
- Llabre, M.M.P., Spitzer, S.P., Siegel, S.M.S., Saab, P.G.P., Schneiderman, N.P., 2004. Applying latent growth curve modeling to the investigation of individual differences in cardiovascular recovery from stress. *Psychosom. Med.* 66, 29–41. doi:10.1097/01.PSY.0000107886.51781.9C
- Lopez-Duran, N.L., Mayer, S.E., Abelson, J.L., 2014. Modeling neuroendocrine stress reactivity in salivary cortisol: adjusting for peak latency variability. *Stress* 3890, 1–11. doi:10.3109/10253890.2014.915517
- Maina, G., Palmas, A., Bovenzi, M., Filon, F.L., 2009. Salivary cortisol and psychosocial hazards at work. *Am. J. Ind. Med.* 52, 251–260. doi:10.1002/ajim.20659
- Masharani, U., Shiboski, S., Eisner, M.D., Katz, P.P., Janson, S.L., Granger, D.A., Blanc, P.D., 2005. Impact of exogenous glucocorticoid use on salivary cortisol measurements among adults with asthma and rhinitis. *Psychoneuroendocrinology* 30, 744–752. doi:10.1016/j.psyneuen.2005.03.003
- Meltzer, L., Hiruma, L., Avis, K., Montgomery-Downs, H., Valentin, J., 2015. Comparison of a commercial accelerometer with polysomnography and actigraphy in children and adolescents. *Sleep* 38, 1323–1330.
- Menet, J.S., Rosbash, M., 2011. When brain clocks lose track of time: cause or consequence of neuropsychiatric disorders. *Curr. Opin. Neurobiol.* 21, 849–857. doi:10.1016/j.conb.2011.06.008
- Miller, R., Plessow, F., 2013. Transformation techniques for cross-sectional and longitudinal endocrine data: application to salivary cortisol concentrations. *Psychoneuroendocrinology* 38, 941–946. doi:10.1016/j.psyneuen.2012.09.013
- Miller, R., Plessow, F., Kirschbaum, C., Stalder, T., 2013a. Classification criteria for distinguishing cortisol responders from nonresponders to psychosocial stress: evaluation of salivary cortisol pulse detection in panel designs. *Psychosom. Med.* 75, 832–40. doi:10.1097/PSY.0000000000000002
- Miller, R., Plessow, F., Rauh, M., Gröschl, M., Kirschbaum, C., 2013b. Comparison of salivary cortisol as measured by different immunoassays and tandem mass

- spectrometry. *Psychoneuroendocrinology* 38, 50–57.
doi:10.1016/j.psyneuen.2012.04.019
- Nader, N., Chrousos, G.P., Kino, T., 2010. Interactions of the circadian CLOCK system and the HPA axis. *Trends Endocrinol. Metab.* 21, 277–286. doi:10.1016/j.tem.2009.12.011
- Nicolson, N.A., Van Diest, R., 2000. Salivary cortisol patterns in vital exhaustion. *J. Psychosom. Res.* 49, 335–342. doi:10.1016/S0022-3999(00)00166-5
- Okun, M.L., Krafty, R.T., Buysse, D.J., Monk, T.H., Reynolds, C.F., Begley, A., Hall, M., 2010. What constitutes too long of a delay? Determining the cortisol awakening response (CAR) using self-report and PSG-assessed wake time. *Psychoneuroendocrinology* 35, 460–468. doi:10.1016/j.psyneuen.2009.08.017.
- Oskis, A., Loveday, C., Hucklebridge, F., Thorn, L., Clow, A., 2009. Diurnal patterns of salivary cortisol across the adolescent period in healthy females. *Psychoneuroendocrinology* 34, 307–316. doi:10.1016/j.psyneuen.2008.09.009
- Perogamvros, I., Keevil, B.G., Ray, D.W., Trainer, P.J., 2010. Salivary cortisone is a potential biomarker for serum free cortisol. *J. Clin. Endocrinol. Metab.* 95, 4951–4958. doi:10.1210/jc.2010-1215
- Polk, D.E., Cohen, S., Doyle, W.J., Skoner, D.P., Kirschbaum, C., 2005. State and trait affect as predictors of salivary cortisol in healthy adults. *Psychoneuroendocrinology* 30, 261–272. doi:10.1016/j.psyneuen.2004.08.004
- Pruessner, J.C., Kirschbaum, C., Meinlschmid, G., Hellhammer, D.H., 2003. Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology* 28, 916–931.
- Pruessner, J.C., Wolf, O.T., Hellhammer, D.H., BuskeKirschbaum, A., von Auer, K., Jobst, S., Kaspers, F., Kirschbaum, C., 1997. Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. *Life Sci.* 61, 2539–2549.
- Raudenbush, S.W., Bryk, A.S., 2002. Hierarchical linear models: applications and data analysis methods. 2nd edition, SAGE Publications, Inc., London.
- Roa, S.L.R., Elias, P.C.L., Castro, M., Moreira, A.C., 2013. The cortisol awakening response is blunted in patients with active Cushing's disease. *Eur. J. Endocrinol.* 168, 657–664. doi:10.1530/EJE-12-0982
- Roberts, A.D.L., Wessely, S., Chalder, T., Papadopoulos, A., Cleare, A.J., 2004. Salivary cortisol response to awakening in chronic fatigue syndrome. *Br. J. Psychiatry* 184, 136–141. doi:10.1192/bjp.184.2.136
- Rohleder, N., Kirschbaum, C., 2007. Effects of nutrition on neuro-endocrine stress responses. *Curr. Opin. Clin. Nutr. Metab. Care* 10, 504–510. doi:10.1097/MCO.0b013e3281e38808
- Rosmond, R., Holm, G., Björntorp, P., 2000. Food-induced cortisol secretion in relation to anthropometric, metabolic and haemodynamic variables in men. *Int. J. Obes. Relat. Metab. Disord.* 24, 416–422. doi:10.1038/sj.ijo.0801173

- Schafer, J.L., Graham, J.W., 2002. Missing data: our view of the state of the art. *Psychol. Methods* 7, 147–177. doi:10.1037/1082-989X.7.2.147
- Scheer, F.A.J.L., Buijs, R.M., 1999. Light affects morning salivary cortisol in humans. *J. Clin. Endocrinol. Metab.* 84, 3395–3398. doi:10.1210/jcem.84.9.6102
- Schlotz, W., 2011. Ambulatory psychoneuroendocrinology: assessing salivary cortisol and other hormones in daily life, in: Mehl, M.R., Conner, T.S. (Eds.), *Handbook of research methods for studying daily life*. Guilford Press, New York, pp. 193–209.
- Schlotz, W., Hammerfald, K., Ehlert, U., Gaab, J., 2011. Individual differences in the cortisol response to stress in young healthy men: testing the roles of perceived stress reactivity and threat appraisal using multiphase latent growth curve modeling. *Biol. Psychol.* 87, 257–264. doi:10.1016/j.biopsycho.2011.03.005
- Schlotz, W., Hellhammer, J., Schulz, P., Stone, A.A., 2004. Perceived work overload and chronic worrying predict weekend-weekday differences in the cortisol awakening response. *Psychosom. Med.* 66, 207–214. doi:10.1097/01.psy.0000116715.78238.56
- Schmidt-Reinwald, A., Pruessner, J.C., Hellhammer, D.H., Federenko, I., Rohleder, N., Schurmeyer, T.H., Kirschbaum, C., 1999. The cortisol response to awakening in relation to different challenge tests and a 12-hour cortisol rhythm. *Life Sci.* 64, 1653–1660.
- Schulz, P., Knabe, R., 1994. Biological uniqueness and the definition of normality. Part 2-- The endocrine 'fingerprint' of healthy adults. *Med. Hypotheses* 42, 63–68.
- Smith, V.C., Dougherty, L.R., 2014. Noisy spit: parental noncompliance with child salivary cortisol sampling. *Dev. Psychobiol.* 56, 647–656. doi:10.1002/dev.21133
- Smyth, N., Clow, A., Thorn, L., Hucklebridge, F., Evans, P., 2013. Delays of 5-15min between awakening and the start of saliva sampling matter in assessment of the cortisol awakening response. *Psychoneuroendocrinology* 38, 1476–1483. doi:10.1016/j.psyneuen.2012.12.013
- Smyth, N., Thorn, L., Hucklebridge, F., Evans, P., Clow, A., 2015. Detailed time course of the cortisol awakening response in healthy participants. *Psychoneuroendocrinology* <http://dx.doi.org/10.1016/j.psyneuen.2015.08.011>.
- Stalder, T., Bäuml, D., Miller, R., Alexander, N., Kliegel, M., Kirschbaum, C., 2013. The cortisol awakening response in infants: ontogeny and associations with development-related variables. *Psychoneuroendocrinology* 38, 552–559. doi:10.1016/j.psyneuen.2012.07.015
- Stalder, T., Evans, P., Hucklebridge, F., Clow, A., 2010a. Associations between psychosocial state variables and the cortisol awakening response in a single case study. *Psychoneuroendocrinology* 35, 209–214.
- Stalder, T., Evans, P., Hucklebridge, F., Clow, A., 2011. Associations between the cortisol awakening response and heart rate variability. *Psychoneuroendocrinology* 36, 454–462. doi:10.1016/j.psyneuen.2010.07.020
- Stalder, T., Evans, P.D., Hucklebridge, F., Clow, A., 2010b. State associations with the cortisol awakening response in healthy females. *Psychoneuroendocrinology* 35, 1245–1252. doi:10.1016/j.psyneuen.2010.02.014

- Stalder, T., Hucklebridge, F., Evans, P., Clow, A., 2009. Use of a single case study design to examine state variation in the cortisol awakening response: Relationship with time of awakening. *Psychoneuroendocrinology* 34, 607–614.
- Steptoe, A, Kunz-Ebrecht, S.R., Brydon, L., Wardle, J., 2004. Central adiposity and cortisol responses to waking in middle-aged men and women. *Int. J. Obes. Relat. Metab. Disord.* 28, 1168–1173. doi:10.1038/sj.ijo.0802715
- Steptoe, A., Ussher, M., 2006. Smoking, cortisol and nicotine. *Int. J. Psychophysiol.* 59, 228–235. doi:10.1016/j.ijpsycho.2005.10.011
- Tassi, P., Muzet, A., 2000. Sleep inertia. *Sleep Med. Rev.* 4, 341–353. doi:10.1053/smr.2000.0098
- Therrien, F., Drapeau, V., Lalonde, J., Lupien, S.J., Beaulieu, S., Tremblay, A., Richard, D., 2007. Awakening cortisol response in lean, obese, and reduced obese individuals: effect of gender and fat distribution. *Obesity* 15, 377–385. doi:10.1038/oby.2007.509
- Thorn, L., Evans, P., Cannon, A., Hucklebridge, F., Evans, P., Clow, A., 2011. Seasonal differences in the diurnal pattern of cortisol secretion in healthy participants and those with self-assessed seasonal affective disorder. *Psychoneuroendocrinology* 36, 816–823. doi:10.1016/j.psyneuen.2010.11.003
- Thorn, L., Hucklebridge, F., Esgate, A., Evans, P., Clow, A., 2004. The effect of dawn simulation on the cortisol response to awakening in healthy participants. *Psychoneuroendocrinology* 29, 925–930. doi:10.1016/j.psyneuen.2003.08.005
- Thorn, L., Hucklebridge, F., Evans, P., Clow, A., 2006. Suspected non-adherence and weekend versus week day differences in the awakening cortisol response. *Psychoneuroendocrinology* 31, 1009–1018.
- Trinder, J., Allen, N., Kleiman, J., Kravetski, V., Kleverlaan, D., Anson, K., Kim, Y., 2003. On the nature of cardiovascular activation at an arousal from sleep. *Sleep* 26, 543–551.
- Trinder, J., Padula, M., Berlowitz, D., Kleiman, J., Breen, S., Rochford, P., Worsnop, C., Thompson, B., Pierce, R., 2001. Cardiac and respiratory activity at arousal from sleep under controlled ventilation conditions. *J. Appl. Physiol.* 90, 1455–1463.
- Van De Water, A.T.M., Holmes, A., Hurley, D.A., 2011. Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography - a systematic review. *J. Sleep Res.* 20, 183–200. doi:10.1111/j.1365-2869.2009.00814.x
- Veldhuis, J.D., Iranmanesh, A., Lizarralde, G., Johnson, M.L., 1989. Amplitude modulation of a burstlike mode of cortisol secretion subserves the circadian glucocorticoid rhythm. *Am. J. Physiol.* 257, E6–E14.
- Vreeburg, S.A., Kruijtzter, B.P., van Pelt, J., van Dyck, R., DeRijk, R.H., Hoogendijk, W.J.G., Smit, J.H., Zitman, F.G., Penninx, B.W.J.H., 2009. Associations between sociodemographic, sampling and health factors and various salivary cortisol indicators in a large sample without psychopathology. *Psychoneuroendocrinology* 34, 1109–20. doi:10.1016/j.psyneuen.2009.04.024

- Weitzman, E.D., Fukushima, D., Nogueira, C., Roffwarg, H., Gallagher, T.F., Hellman, L., 1971. Twenty-four hour pattern of episodic secretion of cortisol in normal subjects. *J. Clin. Endocrinol. Metab.* 33, 14–22. doi:10.1210/jcem-33-1-14
- Wetherell, M.A., Lovell, B., Smith, M.A., 2015. The effects of an anticipated challenge on diurnal cortisol secretion. *Stress* 18, 42–48. doi:10.3109/10253890.2014.993967
- Whembolua, G.L.S., Granger, D.A., Singer, S., Kivlighan, K.T., Marguin, J.A., 2006. Bacteria in the oral mucosa and its effects on the measurement of cortisol, dehydroepiandrosterone, and testosterone in saliva. *Horm. Behav.* 49, 478–483. doi:10.1016/j.yhbeh.2005.10.005
- Wilder, J., 1962. Basimetric approach (law of initial value) to biological rhythms. *Ann. N. Y. Acad. Sci.* 98, 1211–1220. doi:10.1111/j.1749-6632.1962.tb30629.x. PMID 14000687.
- Wilhelm, I., Born, J., Kudielka, B.M., Schlotz, W., Wüst, S., 2007. Is the cortisol awakening rise a response to awakening? *Psychoneuroendocrinology* 32, 358–366. doi:10.1016/j.psyneuen.2007.01.008
- Williams, E., Magid, K., Steptoe, A., 2005. The impact of time of waking and concurrent subjective stress on the cortisol response to awakening. *Psychoneuroendocrinology* 30, 139–148. doi:10.1016/j.psyneuen.2004.06.006
- Willoughby, M., Vandergrift, N., Blair, C., Granger, D.A., 2007. A structural equation modeling approach for the analysis of cortisol data collected using pre-post-post designs. *Struct. Equ. Model. A Multidiscip. J.* 14, 125–145. doi:10.1207/s15328007sem1401_7
- Wolf, O.T., Fujiwara, E., Luwinski, G., Kirschbaum, C., Markowitsch, H.J., 2005. No morning cortisol response in patients with severe global amnesia. *Psychoneuroendocrinology* 30, 101–105. doi:10.1016/j.psyneuen.2004.05.001
- Wolfram, M., Bellingrath, S., Kudielka, B.M., 2011. The cortisol awakening response (CAR) across the female menstrual cycle. *Psychoneuroendocrinology* 36, 905–912. doi:10.1016/j.psyneuen.2010.12.006
- Wright, C.E., Steptoe, A., 2005. Subjective socioeconomic position, gender and cortisol responses to waking in an elderly population. *Psychoneuroendocrinology* 30, 582–590. doi:10.1016/j.psyneuen.2005.01.007
- Wüst, S., Federenko, I., Hellhammer, D.H., Kirschbaum, C., 2000a. Genetic factors, perceived chronic stress, and the free cortisol response to awakening. *Psychoneuroendocrinology* 25, 707–720. doi:10.1016/S0306-4530(00)00021-4
- Wüst, S., Wolf, J., Hellhammer, D.H.D., Federenko, I., Schommer, N., Kirschbaum, C., 2000b. The cortisol awakening response - normal values and confounds. *Noise Health* 2, 79–88.
- Zeiders, K.H., Doane, L.D., Adam, E.K., 2011. Reciprocal relations between objectively measured sleep patterns and diurnal cortisol rhythms in late adolescence. *J. Adolesc. Heal.* 48, 566–571. doi:10.1016/j.jadohealth.2010.08.012