Title	Evaluation and update of the expert consensus guidelines for the assessment of the cortisol awakening response (CAR)
Authors	Stalder, Tobias;Lupien, Sonia J.;Kudielka, Brigitte M.;Adam, Emma K.;Pruessner, Jens C.;Wüst, Stefan;Dockray, Samantha;Smyth, Nina;Evans, Phil;Kirschbaum, Clemens;Miller, Robert;Wetherell, Mark A.;Finke, Johannes B.;Klucken, Tim;Clow, Angela
Publication date	2022-09-27
Original Citation	Stalder, T., Lupien, S. J., Kudielka, B. M., Adam, E. K., Pruessner, J. C., Wüst, S., Dockray, S., Smyth, N., Evans, P., Kirschbaum, C., Miller, R., Wetherell, M. A., Finke, J. B., Klucken, T. and Clow, A. (2022) 'Evaluation and update of the expert consensus guidelines for the assessment of the cortisol awakening response (CAR)', Psychoneuroendocrinology, 146. doi: 10.1016/j.psyneuen.2022.105946
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1016/j.psyneuen.2022.105946
Rights	© 2022, Elsevier Ltd. All rights reserved. . This manuscript version is made available under the CC BY-NC-ND 4.0 license. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2025-07-02 05:17:53
Item downloaded from	https://hdl.handle.net/10468/13805

Evaluation and update of the expert consensus guidelines for the assessment of the cortisol awakening response (CAR)

Tobias Stalder[a], Sonia J. Lupien[b], Brigitte M. Kudielka[c], Emma K. Adam[d], Jens C. Pruessner[e], Stefan Wüst[c], Samantha Dockray[f], Nina Smyth[g], Phil Evans[g], Clemens Kirschbaum[h], Robert Miller[h], Mark A. Wetherell[i], Johannes B. Finke[a], Tim Klucken[a], Angela Clow[g]

[a] Department of Psychology, University of Siegen, Siegen, Germany

[b] Department of Psychiatry, University of Montreal, Montreal, Quebec, Canada

[c] Department of Psychology, University of Regensburg, Regensburg, Germany

[d] School of Education and Social Policy, Northwestern University, Evanston, USA

[e] Department of Psychology, University of Konstanz, Konstanz, Germany

[f] School of Applied Psychology, University College Cork, Cork, Ireland

[g] Department of Psychology, University of Westminster, London, UK

[h] Department of Psychology, TU Dresden, Dresden, Germany

[i] Department of Psychology, Northumbria University, Newcastle, UK

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

Abstract

The cortisol awakening response (CAR) is frequently assessed in psychobiological (stress) research. Obtaining reliable CAR data, however, requires careful attention to methodological detail. To promote best practice, expert consensus guidelines on the assessment of the CAR were published (Stalder et al., 2016, *PNEC*). However, it is unclear whether these highly cited guidelines have resulted in actual methodological improvements. To explore this, the *PNEC* editorial board invited the present authors to conduct a critical evaluation and update of current CAR methodology, which is reported here. (i) A quantitative evaluation of methodological quality of CAR research published in *PNEC* before and after the guidelines (2013-2015 vs. 2018-2020) was conducted. Disappointingly, results reveal little improvement in the implementation of central recommendations (especially objective time verification) in recent research. (ii) To enable an update of guidelines, evidence on recent developments in CAR assessment is reviewed, which mostly confirms the accuracy of the majority of the original guidelines. Moreover, recent technological advances, particularly regarding methods for the verification of awakening and sampling times, have emerged and may help to reduce costs in future research. (iii) To aid researchers and increase accessibility, an updated and streamlined version of the CAR consensus guidelines is presented. (iv) Finally, a response of the *PNEC* editorial board to the present results is reported: potential authors of CAR research to be published in *PNEC* will be required to submit a methodological checklist (based on the current guidelines) alongside their article. This will increase transparency and enable reviewers to readily assess the quality of the respective CAR data. Combined, it is hoped that these steps will assist researchers and reviewers in assuring higher quality CAR assessments in future research, thus yielding more reliable and reproducible results and helping to further advance this field of study.

1. Introduction

The cortisol awakening response (CAR), i.e., the marked increase in cortisol secretion over the first 30-45 min following morning awakening (Pruessner et al., 1997), is frequently assessed in psychoneuroendocrinological research. High interest in the CAR arises from it being a unique aspect of hypothalamus-pituitary-adrenal (HPA) axis activity, combining aspects of both a reactivity index (response to awakening) and features of circadian regulation (being restricted to the morning sleep-wake transition). It can yield high ecological validity when captured by salivary sampling at participants' home settings. However, obtaining reliable CAR data is not trivial and requires careful attention to methodological detail. To promote best methodological practice in this area, in 2015, the International Society of Psychoneuroendocrinology (ISPNE) initiated an expert panel charged with establishing clear guidelines for the assessment of the CAR. The most active researchers in this field were invited and undertook an in-depth evaluation of relevant evidence, including several rounds of in-group consultations and an open discussion at the 2015 ISPNE conference in Edinburgh. This eventually led to the publication of best practice guidelines on the assessment of the CAR in this very journal (Stalder et al., 2016) as well as their presentation to the scientific community at the 2016 ISPNE conference in Miami. Based on this expert consensus, *Psychoneuroendocrinology* (*PNEC*) also revised their *guide for authors* stating that researchers seeking to report CAR data are strongly recommended to follow these guidelines and/or to describe and justify any deviations from them.

Six years have passed since the publication of the CAR consensus guidelines. While it appears that they have been well-received by the scientific community, e.g., as indexed by over 750 citations to date[1], it remains unclear to which extent this has indeed resulted in improvements in methodological quality. To gain insight into this important question, the *PNEC* editorial board invited the present authors to conduct a critical evaluation of the current state of the methodology in this research field. The present report summarizes the results of this undertaking. We first recapitulate the most central guidelines from the 2016 report (*Section 2*). In the interest of brevity and concision, only a short background (without full referencing) is provided. Readers interested in an in-depth and fully-referenced discussion of these factors are referred to the first guidelines report (Stalder et al., 2016). In *Section 3*, we report the results of a quantitative literature-based evaluation which examines the extent to which the most central guidelines outlined in Stalder et al. (2016) have been adhered to in recent research published in the journal *PNEC*. For this, changes in methodological details of CAR research published during a three-year-period after the guidelines (2018 – 2020) compared to a three-year-period before the guidelines (2013 – 2015) are examined. In *Section 4*, this is followed by a critical review which investigates whether the recommendations of the initial guidelines report are still in

[1] Based on scholar.google.com; accessed on August 05 2022.

agreement with current evidence or whether more recent findings call for adaptations. Finally, in *Section 5*, the present quantitative findings are discussed in the light of recent evidence and two strategies for improvement in future research are outlined. These strategies include (i) an updated and streamlined version of the CAR consensus guidelines as well as (ii) a revision of the CAR research submission procedures in *PNEC*, which features a compulsory guidelines checklist for future *PNEC* submissions.

2. Main consensus guidelines

2.1 Cortisol awakening response: background and main outcome measures

The CAR represents the sharp increase in cortisol levels across the first 30-45 min following morning awakening (e.g., Clow et al., 2004; Pruessner et al., 1997). It is seen as part of normal, healthy human physiology, although its magnitude and variability may be altered with pathology (e.g., Boggero et al., 2017). The CAR period is embedded within circadian cortisol rhythmicity, which is characterized by an increase in cortisol levels before awakening, followed by the marked CAR and a subsequent decline of mean cortisol levels over the remainder of the diurnal phase (e.g., Weitzman et al., 1971). Within circadian rhythmicity, the CAR represents a relatively distinct feature, which is not closely related to earlier or later phases as it is influenced by a unique set of state factors (e.g., well-described influences of light exposure and time of awakening; Edwards et al., 2001; Law et al., 2013; Petrowski et al., 2019; Thorn et al., 2004). Together, these features make the CAR a fascinating and highly researched object of study.

Concerning the overall conceptualization and measurement of the CAR, an important finding emerged from sleep-laboratory work showing that the CAR is not a mere continuation of the pre-awakening cortisol increase but is indeed a superimposed *response* to awakening (Wilhelm et al., 2007). Furthermore, differential processes are likely to affect the pre-awakening cortisol increase and the CAR (review: Clow et al., 2010). Together, this suggests that strategies for quantifying post-awakening cortisol secretion need to capture two main underlying components: First, the endpoint of the pre-awakening increase (starting point of the CAR), reflected by the first sample on awakening (S1) and, second, the actual CAR, i.e., the dynamic *response* to awakening. To report information on both distinct aspects components is important for advancing current understanding on underlying processes and main correlates.

It was thus strongly recommended by the panel of experts that inferential results for *both* S1 and a measure of the CAR are reported (Stalder et al., 2016). While S1 comprises only a singular value, several measures are available for capturing endocrine responses, such as the CAR. These include the area under the curve with respect to increase (AUC_I; Pruessner et al., 2003), the mean Increase (MnInc), the min-max difference (Miller et al., 2018) or a specific component added to a multilevel

growth curve model (Adam et al., 2015). By contrast, measures that combine information on S1 and the CAR, particularly the area under the curve with respect to ground (AUC_G; Pruessner et al., 2003), are useful for capturing 'total post-awakening cortisol concentrations'. Such measures may be reported alongside S1 and CAR but should not be referred to as capturing the CAR (Stalder et al., 2016).

2.2 Objective verification of awakening and *sampling times*

The central challenge for obtaining reliable CAR data arises from the fact that the CAR is typically assessed by self-collection of saliva samples in participants' home settings, which implies a lack of oversight of the sampling procedure by researchers. This is problematic since the validity of CAR data critically relies on the temporal accuracy of sampling. A typical CAR sampling schedule commences with a first sample taken on awakening (S1) and continues with further samples taken at fixed intervals over the first hour post-awakening. Fig. 1a depicts an example of a common CAR sampling schedule, using four sampling points.

Importantly, abundant evidence shows that in uncontrolled domestic settings, non-adherence to a CAR sampling schedule (i.e., not taking saliva samples accurately at the pre-specified times) is highly prevalent and leads to measurement bias on CAR quantification (Dockray et al., 2008; Kudielka et al., 2003). Critically, such adverse effects may not only entail an inflated *unsystematic* measurement error but, since non-adherence is also likely to covary with central variables of interest (e.g., personality, pathological status, chronic stress), this may also induce *systematic bias* on CAR data and thus pose a major threat to scientific progress in this area. Maximizing protocol adherence by using objective verification of awakening *and* sampling times is thus the single most important guideline formulated in the consensus report (Stalder et al., 2016).

Several methods are available to objectively verify participants' self-reported awakening times, most commonly actigraphy-based strategies (either wrist or chest-worn), but also electrocardiography or polysomnography (see 4.1 for recent developments in this area). Each of these methods works sufficiently well for identifying the moment of awakening, with wrist actigraphy tending to be most cost-effective. Similarly, different methods have been used to objectively verify sampling times. A common strategy involves electronic monitoring systems, e.g., screw top bottles in which saliva sampling devices are placed and which record individual times of bottle openings. Another highly cost-effective approach, arising from modern technology, is to ask participants to use their smartphones to obtain time-stamped self-photographs ('selfies') of the sample taking.

When utilizing objective methods to verify awakening *and* sampling times, participants should also be informed about these verification strategies as doing so has been shown to improve sampling accuracy (Broderick et al., 2004; Kudielka et al., 2003). In addition, a self-report diary system (digital or paper-pencil) should be used to complement objective data. Once objective and self-report data on

awakening and sampling times have been obtained, information on potential deviations from the scheduled protocol can either be used as part of (i) data exclusion strategies or (ii) within statistical modeling approaches. Data exclusion strategies require setting a relatively strict accuracy margin (recommendation: ≤ 5 min discrepancy between scheduled and verified times) for acceptable data. However, as this may result in substantial data loss, it is important to keep the rate of classified inaccurate data as low as possible by employing a full range of measures to maximize adherence (a detailed description is provided in Stalder et al., 2016). The use of statistical modeling approaches, in which information on actual verified sampling times can be incorporated and utilized for the calculation of CAR estimates, provides a more economical strategy (Stalder et al., 2016).

In case research is not conducted in participants' homes but in a stationary setting (e.g., hospital, sleep laboratory), monitoring of saliva sampling by study personnel can be conducted. Furthermore, forced-awakening at a prespecified time may be used. However, this should also be accompanied by objective waketime verification to rule out that participants have woken up prior to the external wake-up time without the researcher's knowledge.

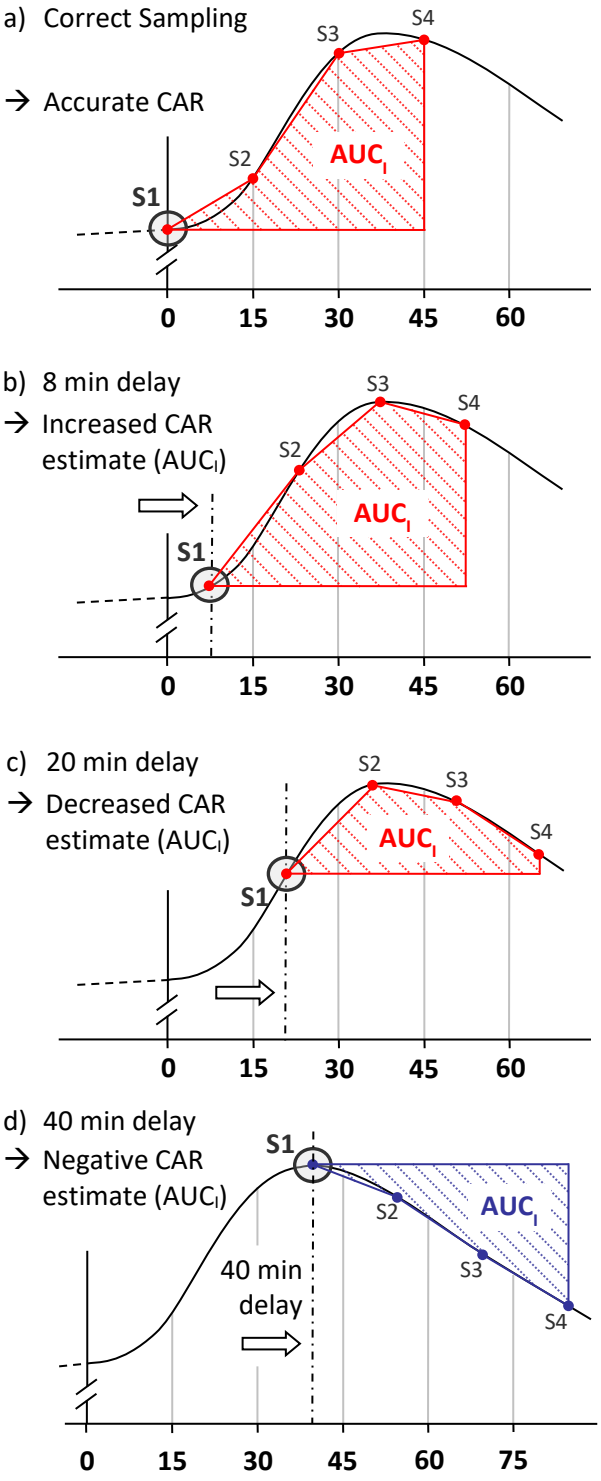

Fig. 1. Illustration of the impact of delayed sampling on two measures of post-awakening cortisol secretion, i.e. the first sample on awakening (S1) and the area under the curve with respect to increase (AUC_I; Pruessner et al., 2003), as a measure of the CAR. Shown are data based on (a) correct sampling and delays of (b) 8 min, (c) 20 min and (d) 40 min between awakening and the beginning of sampling. Later samples are shown to be taken as scheduled with 15 min intervals. Figure reproduced from Stalder et al. (2016) with permission.

2.3 Participant instructions and covariate control

Participants should be instructed to not engage in behavior known to influence cortisol secretion. In the context of CAR research, this mainly refers to (i) taking nil by mouth other than water, (ii) no smoking, and (iii) no exercising over the post-awakening sampling period (Stalder et al., 2016). Covariates that cannot be prevented by instruction and/or standardized under ambulatory conditions are usually dealt with through statistical adjustment or appropriate matching of study groups. Relevant trait-like covariates of the CAR include age, sex, ethnicity, socioeconomic status, habitual smoking, heavy drinking, body-mass-index or obesity, as well as oral contraceptive use (see Stalder et al., 2016).

While the above trait-like covariates are similar to other areas of psychoneuroendocrinological enquiry, a more specific situation arises from the fact that the expression of the CAR on a particular day is determined to a greater extent by state/situational factors than by stable trait-like influences (Almeida et al., 2009; Hellhammer et al., 2007; Stalder et al., 2010). Relevant state factors to consider include sleep-related variables, most importantly time of awakening, but also ambient light levels and season of sampling. A particularly important group of state covariates involves psychosocial factors surrounding the sampling day. For example, increased CARs have been observed on days which participants anticipate to be more demanding/challenging and which pose a higher prospective memory load (Bäumler et al., 2014; Wetherell et al., 2015). In line with this, larger CARs are also regularly observed on weekdays compared to weekend days (Kunz-Ebrecht et al., 2004; Schlotz et al., 2004). Potential influences of such state psychosocial factors on the CAR should thus be assessed and adjusted for (unless they are central variables of interest).

Awareness of state psychosocial influences on the CAR is also important when designing cross-sectional research. For example, CAR profiles should not be compared between groups examined under different state circumstances, e.g., hospitalized patients vs. home-based controls or working vs. non-working groups. Under such circumstances, findings of group differences in CAR profiles can emerge due to differences in situational characteristics of assessment contexts but might be misinterpreted as arising from trait-like factors, e.g., participants' clinical status. This possibility of state-related confounding should thus always be considered when conducting CAR research and prevented by appropriate and considered design choices (for a more detailed discussion see Stalder et al., 2016).

In addition, participant exclusion is recommended under some circumstances, including e.g., brain damage (particularly to the hippocampus), use of glucocorticoid-based medication, presence of an endocrine disorder (e.g., Morbus Cushing, Morbus Addison) as well as current pregnancy. Likewise, sampling should be rescheduled to a later time in cases of current illness or acute circadian disruptions, such as shift work or jet lag (see Stalder et al., 2016).

2.4 Assessment of the CAR: number of samples and study days

CAR assessment protocols may differ in the number of samples used to capture the CAR. The decision on the appropriate number of samples involves a cost/accuracy trade-off, with more samples improving the accuracy of CAR estimation but also being associated with higher costs and greater participant burden. Minimal protocols relying on only two samples (usually 0 and 30 min) are often used in larger epidemiological research but run the risk of not accurately capturing peak levels of the CAR. Since the timing of the CAR peak may itself be related to variables of interest (Lopez-Duran et al., 2014), two-sample protocols cannot distinguish between associations with CAR magnitude and CAR peak timing. It is thus recommended that in adult populations, a protocol with at least three post-awakening samples (0, 30, 45 min) is used (Stalder et al., 2016).

For cross-sectional research, an important question is how many days of CAR assessment are necessary to obtain reliable estimates. Here, researchers need to be aware of the fact that expression of the CAR on a single day is determined to a larger extent by state/situational factors (61-82%) than by stable, trait-like components (15-37%; Hellhammer et al., 2007). Hence, in research seeking to relate CAR profiles to other trait-like measures, detectable effect sizes will likely be limited for single-day CAR data and will increase with each additional day of CAR assessment.

3. Evaluation of adherence to CAR consensus guidelines

The above summarized guidelines provide a methodological framework based on which reliable, high-quality CAR data can be obtained. However, in order for this to work, guidelines have to be followed by researchers and it is currently unknown to which extent this has been done in recent research. We thus conducted a quantitative literature-based evaluation of adherence to the most central guidelines outlined in Stalder et al. (2016). Methodological aspects are contrasted between CAR research published during a three-year-period before the guidelines (PRE, 2013 – 2015) vs. a three-year-period after the guidelines (POST, 2018 – 2020). Given that the guidelines report was published in *PNEC* and that a firm commitment to these guidelines was specific to this journal (e.g., through a note in the *guide for authors*), the present evaluation focuses on articles published in *PNEC* only.

3.1 Methodology

3.1.1 Search strategy and selection of studies

Potential studies were identified through a Web of Science database search up to 01 May 2021 using the following search strategy: (i) exact terms “cortisol awakening” or “awakening cortisol” or “cortisol morning” or “morning cortisol” or “cortisol response to awakening” in the title, abstract and/or keywords. The search was restricted to original articles in English language published in the years 2013-2015 and 2018-2020 in the journal *Psychoneuroendocrinology.* This initial search produced 143 records

($k_{2013\text{-}2015}$ = 76, $k_{2018\text{-}2020}$ = 67). Next, the search results were manually screened to remove articles without any data on the CAR (*k* = 19), animal work (*k* = 1), duplicate data (*k* = 1) and reviews or meta-analyses (*k* = 4). Further, as it was our goal to derive information on how the CAR is assessed in regular research, we excluded methodological work that investigated the impact of compliance on the CAR (*k* = 7). Finally, as the vast majority of CAR research is based on salivary assessments, work reporting CAR profiles from blood cortisol, which involves different methodological concerns, was excluded (*k* = 1). If an article reported data from more than one independent sample in which the CAR was assessed using identical methodological standards, this was treated as a single article/study and subsample sizes were added up. Conversely, if the methodological standards of CAR assessments differed between independent samples reported in one article, respective samples were treated as independent studies (*k* = 2). Combined, this led to a final set of 112 articles/studies that were included in our analyses.

3.1.2 Data extraction and preparation

The initial study selection and manual screening was conducted by TS and JBF. Next, detailed screening and data extraction (explained below) were conducted by three student helpers, who first extracted data independently and subsequently cross-checked each of their entries. Finally, another detailed check of individual entries was conducted by TS and TK. We first extracted basic information on bibliography and sample characteristics, which included author(s), year of publication and sample size. Concerning methodological aspects of CAR assessment, we extracted data concerning the main guidelines recommendations from the guidelines report. Specifically, we extracted information on (i) number of days of CAR assessment per timepoint, (ii) number of saliva sampling times per CAR assessment, and (iii) use of an objective method for verification of awakening time (levels: 0 = none, 1 = use of either actigraphy, polysomnography, electrocardiography/heart rate or mixed methods). Next, (iv) concerning the use of an objective method for verification of sampling times, we extracted data of any employed method and coded these data according to the guidelines report criteria (levels: 0 = none, 1 = use of either electronic monitoring systems, time-stamped photograph or sampling monitoring by personnel). Further, data were extracted on (v) the provision of clear instructions about unwanted morning behavior, i.e., nil by mouth other than water, no smoking, no exercising (levels: 0 = none provided, 1 = at least one provided), (vi) control for state covariates (levels: 0 = none considered, 1 = consideration of at least one relevant state covariate (i.e., time of awakening, sleep duration or quality, ambient light level, season, weekday vs. weekend, prior day experiences, anticipation of day ahead/prospective memory load); (vii) control for trait covariates (levels: 0 = none considered, 1 = consideration of at least one relevant trait covariate (i.e., age, sex, ethnicity, socioeconomic status, habitual smoking, heavy drinking, body-mass-index/obesity, oral contraceptive use); (viii) reporting of associations with the first sample on awakening (S1; levels: 0 = no, 1 = yes); (ix)

reporting of associations with the CAR, i.e., the dynamic of the increase (e.g., AUC_I, MnInc or delta; levels: 0 = no, 1 = yes); (x) reporting of associations with measures of post-awakening cortisol concentrations (e.g., AUC_G; levels: 0 = no, 1 = yes). Finally, (xi) we also extracted whether articles cited the CAR consensus report.

3.1.3 Analytical methods

The results of the present evaluation are descriptive in nature (i.e., the rate of following methodological guidelines in current research) and thus no formal inferential testing was conducted. In the following, we report descriptive data for the two assessment periods as well as information on PRE-to-POST changes and 95% confidence intervals of respective changes. Continuous variables are summarized as geometric means and PRE-to-POST changes were calculated as differences of geometric means. Confidences intervals were calculated by Normal approximation. Dichotomous variables summarized as proportions (percentage scaling) and PRE-to-POST changes were calculated as differences of proportions. Confidence intervals for respective differences of proportions were calculated using the Agresti-Caffo approach (Agresti and Caffo, 2000).

3.2 Results

The CAR guidelines report was cited by 66.7% of studies (36 out of 54) published in the POST-period. Table 2 provides descriptive data of methodological aspects of CAR research published in the three-year periods from 2013-2015 (PRE) vs. 2018-2020 (POST), i.e., before and after publication of the CAR guidelines report. In addition, information on PRE-to-POST changes are provided. The number of days over which the CAR was assessed per timepoint in cross-sectional research was found to decrease by 0.38 days, i.e., from 1.90 days in the PRE-period to 1.52 days in the POST-period. By contrast, the number of sampling points for each CAR assessment showed a mean PRE-to-POST increase of 0.38 samples, with an average of 2.97 samples being collected to assess the CAR in POST-research.

Table 2. Methodological aspects concerning the assessment of the CAR in research published before and after publication of the guidelines report.

	2013-2015 (PRE, *k* = 58)	2018-2020 (POST, *k* = 54)	PRE-to-POST change[1]	95% CI of change
Days of CAR assessment per timepoint ($mean_{geom}$, SD_{geom})	1.90 (1.80)	1.52 (1.83)	-0.38	-0.85, 0.09
Number of CAR samples ($mean_{geom}$, SD_{geom})	2.59 (1.37)	2.97 (1.39)	0.38	0.02, 0.74
Objective verification of awakening times (%, *n*)	13.8 (8)	13.0 (7)	-0.8	-13.4, 11.8
Objective verification of sampling times (%, *n*)	17.2 (10)	31.5 (17)	14.3	-1.5, 30.0
Objective verification of awakening *and* sampling times (%, *n*)	6.9 (4)	9.3 (5)	2.4	-7.8, 12.5
Instructions on unwanted morning behaviors (at least one) (%, *n*)	72.4 (42)	90.7 (49)	18.3	4.5, 32.2
Control of state covariates (at least one) (%, *n*)	62.1 (36)	55.6 (30)	-6.5	-24.7, 11.7
Control of trait covariates (at least one)[2] (%, *n*)	90.2 (37)	89.5 (34)	-0.7	-18.7, 17.0
S1 reported[3] (%, *n*)	50.0 (29)	42.6 (23)	-7.4	-25.8, 11.0
CAR measure reported[3] (%, *n*)	94.8 (55)	89.1 (49)	-5.7	-13.7, 5.5
Total post-awakening cortisol measure reported[3] (%, *n*)	34.5 (20)	38.2 (21)	3.7	-13.4, 22.3
S1 *and* CAR reported[3] (%, *n*)	50.0 (29)	38.9 (21)	-11.1	-29.4, 7.2

[1] shown as changes in mean values (continuous measures) or changes in percentages (dichotomous measures); [2] only in cross-sectional research (ns PRE: 41, POST: 38), [3] refers to the reporting of associative results (i.e., use in inferential statistics) with these measures.

Figures 2a-c show descriptive data on the use of objective strategies to verify awakening and sampling times over the six examined years. This, together with data shown in Table 2, illustrates that there was negligible change in the use of methods to objectively verify awakening times. By contrast, there was a 14.3% PRE-to-POST increase in the use of objective methods to verify sampling times as classified in the guidelines report, with 31.5% of studies in the POST-period employing such methods. Conversely, there was hardly any change in research following the recommendation of objectively verifying *both* awakening and sampling times, with only 9.3% of studies conducted after publication of the CAR guidelines report applying both verification strategies.

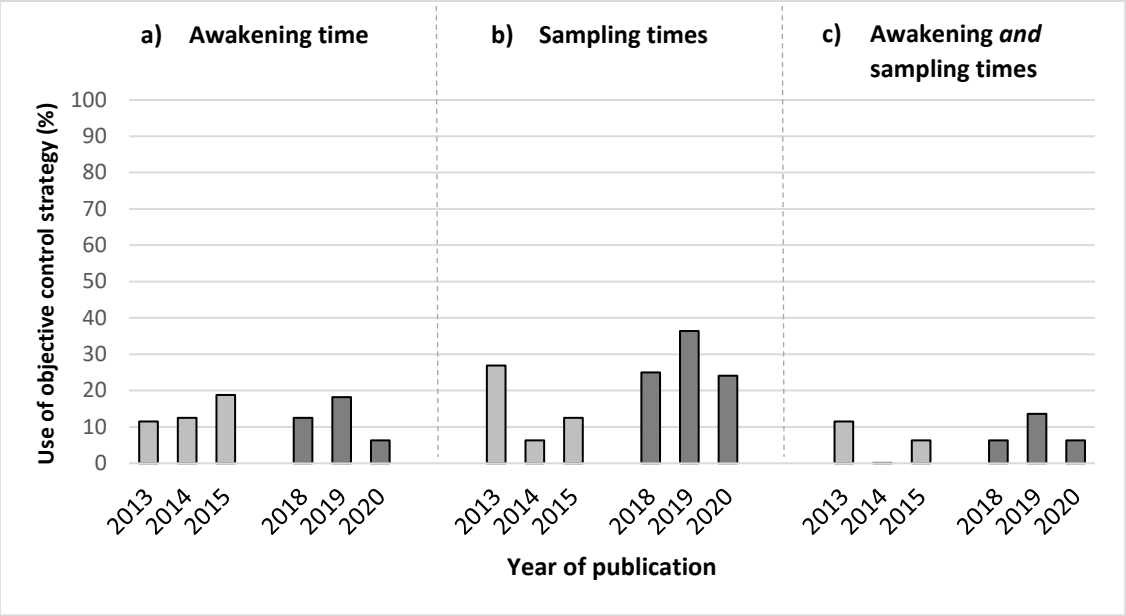

Figure 2. Use of objective strategies to verify (a) awakening time, (b) sampling times (alternative coding), and (c) both awakening and sampling times in articles published in individual years over the PRE- (light grey) and the POST-period (dark grey).

Beyond objective verification strategies, an 18.3% increase in providing instructions on unwanted morning behaviors was seen, with 90.7% of studies published after the guidelines report providing instructions on at least one unwanted morning behavior. Conversely, no substantial changes emerged with regard to the control of state and trait covariates, with numerical *decreases* being seen in both categories and particularly the control of state covariates remaining low, with only 55.6% of studies/articles in the POST-period controlling for at least one covariate. A similar picture emerged with regard to the reporting of post-awakening cortisol data, with no major changes in the reporting of associations with the first sample on awakening (S1), CAR measures or the combined reporting of S1 and the CAR. However, again, numerical decreases were seen for each named category, with the reporting of S1 *and* CAR (as recommended in the guidelines report) decreasing by 11.1% to only 38.9% of studies in the POST-period.

4. Current status: evaluation of guideline content

The results of the above quantitative evaluation suggest that adherence to central best-practice guidelines on the assessment of the CAR in current research published in *PNEC* is still low. Particularly, the essential recommendation that both awakening and sampling times should be verified objectively, is only followed by a small minority of studies. Before discussing potential implications of these findings, it is important to evaluate whether recommendations from these guidelines are still in agreement with current evidence or whether novel data, published after 2016, call for adaptations.

4.1 Objective verification strategies and dealing with verified inaccurate data

An interesting area of development concerns wearable devices for the objective verification of awakening times. Here, particularly the possibility that consumer grade devices might be usable for this purpose is of high interest as these tend to be cheaper and might thus increase feasibility in larger studies. Current evidence indicates rapid technological advancement in this area: A systematic review and meta-analysis on the accuracy of wristband fitbit models for assessing sleep parameters published in 2019 suggested that while early-generation devices performed poorly against gold-standard PSG, recent-generation models that collectively utilized heart rate variability and body movement data yielded results that did not differ significantly in estimates of 'total sleep time' or 'wake after sleep onset' from polysomnography (Haghayegh et al., 2019). Importantly, a recent study, which was conducted under unrestricted conditions in participants' domestic settings (as is the case in most CAR research), compared the performance of four commercial wearable devices and a research grade actigraphy watch against a high performing mobile EEG headband (Chinoy et al., 2022). Results indicated that most recent-generation commercial sleep tracking devices performed equally well or better than a research-grade actigraphy watch for assessing common sleep parameters (Chinoy et al., 2022). These are promising data, indicating that recent-generation consumer grade devices using both PPG sensors and actigraphy data for their sleep/wake classification algorithm (e.g., from Fitbit Inc., San Francisco, USA; Polar Electro, Helsinki, Finland or Garmin, Olathe, KS, USA) might provide feasible alternatives to more expensive research grade products in future research. Despite these advantages, negative aspects of commercial sleep-tracking devices are likely to relate to the quality and flexibility of the readouts that can be obtained, which for research grade devices are likely to be more researcher-friendly and to offer more control (e.g., by allowing researchers to download raw data and run their own wake detection algorithm). Furthermore, some consumer grade devices do not allow researchers to manually select the recording time window, which can lead to issues of premature stoppage of the sleep recordings, making it important to choose devices that allow manual time frame setting (e.g., Fitbit models) (Evan D. Chinoy, personal communication). Future research examining the utility of recent-generation consumer wearables for verifying awakening time in ambulatory CAR research will be important to further develop this important and promising possibility.

Objective wake-time verification needs to be complemented by objective verification of *sampling times* in order to yield high-quality CAR data. Two promising novel strategies to objectively verify sampling times have emerged. The first approach utilizes smartphones (or other handheld devices), which are set to prompt/remind participants of individual sampling times. Critically, with each prompt a unique code is presented for only a short amount of time, which participants have to record on the label of the saliva-sampling device used for the respective collection (e.g., Beddig et al., 2019; Powell et al., 2015; Powell & Schlotz, 2012). A second alternative approach involves participants being

sent a text message at individual sampling times, to which they have to reply within a short timeframe (1-2 minutes) confirming that they took the respective saliva sample (e.g., Rodriguez et al., 2018). Importantly, both approaches are only recommendable if participants themselves initiate the start of the prompting procedure (i.e., they provide the input about when they woke up), rather than relying on pre-set times of 'usual awakening'. This is important to exclude the possibility that participants woke up prior to their usual awakening times (e.g., due to being nervous about the sampling situation), which would invalidate the procedure. Besides these new approaches, it is also important to note that the previously outlined strategy of utilizing time-stamped photographs provides an excellent and low-cost option, which has now been applied successfully in recent CAR research (Zhu et al., 2019). Implementation of this approach has become easier since an increasing number of smartphone operating systems now feature automatic photo time-stamp functions. Further, taking photos only of the coded sampling devices (i.e., sparing participants' faces) alleviates potential privacy protection concerns, thus making this an excellent methodological choice for projects with a restricted budget.

Beyond novel objective verification approaches, recent findings also inform about the best strategy for dealing with data objectively verified as inaccurately sampled (e.g., due to a delayed initiation of sampling). Smyth et al. (2016) revealed that when using objectively verified sampling times, actual cortisol data accurately mapped onto a typical post-awakening cortisol growth curve, even if respective sampling times deviated from scheduled protocol times. These data are thus in favor of the incorporation of verified inaccurate data (up to a delay of 15 min) into a CAR estimation model, rather than excluding these data. Such incorporation strategies prevent against unnecessary data loss and are thus more economical. Still, if researchers decide to employ a data exclusion strategy, data from this study also corroborated the notion that a strict accuracy margin should be set (recommended: 5 min; Stalder et al., 2016), as even short delays of 4-6 min were found to lead to a misrepresentation of the CAR (Smyth et al., 2016).

4.2 State and trait covariates

Several recent findings corroborate guidelines concerning relevant covariates of the CAR. Concerning state covariates, the notion that anticipating greater demands/stress on the upcoming day leads to an elevated CAR was further supported (e.g., Elder et al., 2018; Kramer et al., 2019). In line with this, findings for elevated CARs on weekdays vs. weekend-days were seen (Skoluda et al., 2016). Together, this further consolidates the important notion that psychosocial factors surrounding the sampling day are an important source of state-specific influence to be considered when designing research featuring CAR assessments (see detailed discussion in Stalder et al., 2016). Besides psychosocial factors, recent sleep laboratory work also confirmed and extended earlier evidence of a stimulatory influence of post-awakening light exposure on the CAR, with particularly blue and green light leading to an increased

CAR (Petrowski et al., 2019). Similarly, in a well-controlled sleep laboratory study, participants woken in low intensity ultraviolet light were found to exhibit an attenuated CAR (Elder et al., 2016). These data thus further strengthen the notion that ambient light levels might be relevant state covariates in some research contexts.

Concerning more stable, trait-like influences, recent meta-analytic evidence confirmed an association between ethnicity and both S1 as well as the CAR (Boileau et al., 2019), thus marking ethnicity as a relevant covariate to consider, although it is important to note that increasing evidence suggests that ethnic differences in cortisol are driven by sociocultural experiences such as differential exposure to discrimination (Adam et al., 2015; Boileau et al., 2019). Another trait-like factor, which was not mentioned in the guidelines report, relates to individual differences in the timing of circadian activity preference, i.e., morningness vs. eveningness chronotype. Earlier research had indicated that greater morningness is particularly associated with an elevated S1 (Griefahn & Robens, 2008; Kudielka et al., 2006, 2007; Randler & Schaal, 2010, but see Dockray & Steptoe, 2011) and this finding was recently replicated under well-controlled sleep laboratory conditions (Petrowski et al., 2020). This suggests that chronotype is another relevant trait-like covariate of the CAR to consider and, depending on the study context, adjust for.

4.3 Unwanted behavior and participant exclusion

The previous guideline proposing that participants should be instructed to refrain from physical exercise during the post-awakening period was based on indirect evidence of a stimulatory influence of exercise above a certain intensity level on HPA axis activity (e.g., Kirschbaum & Hellhammer, 1994). Recent sleep-laboratory research confirmed that even short bursts of high-intensity exercise (30 sec cycling sprint) result in an increased CAR (Kovac et al., 2021), thus strengthening the respective guideline.

Concerning the need to exclude participants or postpone sampling, it was previously recommended that sampling during ovulation should be avoided. This was particularly based on findings by Wolfram et al. (2011) showing an increased CAR during the short period of ovulation. By contrast, another study reported an attenuated CAR during menses (Ozgocer et al., 2017). Importantly, recent studies could not replicate these findings, showing no influence of menstrual cycle phase on the CAR (Kayacan et al., 2021; Ozgocer et al., 2022). Combined, it appears that if effects of menstrual cycle phase exist, they are relatively small and that the respective guideline may thus be alleviated.

4.4 Measurement of the CAR

Some interesting recent contributions have also been made with regard to strategies for quantifying post-awakening cortisol secretion. As an addition to common measures of the CAR, the notion of CAR

'salience' was introduced, which includes the (negative) change from CAR peak to post-peak trough (Evans et al., 2019). Preliminary data suggest that CAR salience, including a measure of HPA axis feedback, might reflect trait-like influences more strongly than traditional measures of CAR rise. Importantly, salience can be determined from the same sampling regime used for calculation of CAR magnitude and reported alongside (Evans et al., 2019). Another relevant addition has emerged from research not directly investigating the CAR, which evaluated the utility of different composite measures, such as the AUC_I, AUC_G, Max-min, etc., for quantifying cortisol reactivity to the Trier Social Stress Test against a well-fitted pharmacokinetic model (Miller et al., 2018). Results indicate that the reactivity component of the stress response is best fitted by the minimum-maximum cortisol difference, which yielded better results than the commonly used AUC_I. Although not being directly based on CAR data, this suggests that the Max-min might also be a valuable measure to be explored (alongside other well-established measures, such as the AUC_I) in CAR research (Miller et al., 2018). Along the same lines, a recent paper by Benz et al. (2019) suggests that additional relationships with trait characteristics of the subjects might be obtained when extending sampling time to 120 min past awakening. In a sample e of 51 healthy young participants, cortisol was measured every thirty minutes for 120 minutes starting with awakening and observed that the duration of the entire first pulse after awakening (indicating rise from, and complete return to, baseline), lasted on average 108 minutes, with the duration of the pulse being associated with subjects' biological sex and menstrual cycle phase. Here, future studies have to determine whether this marker might also be linked to (psycho)pathological or demographic factors.

5. Overall discussion and updated guidelines

The results of the present quantitative evaluation revealed that central recommendations outlined in the CAR assessment guidelines report (Stalder et al., 2016) were only followed by a small minority of those articles/studies published in *PNEC* between 2018 and 2020. The fact that only 9.3% of these studies followed the central guideline of using objective methods to verify *both* awakening and sampling times weighs particularly heavy here. These are disappointing results given that a strong recommendation to follow these guidelines is provided in the *PNEC* guide for authors and given that, indeed, 66.7% of the examined articles cited the guidelines report. The low rate with which central guidelines were followed is even more surprising given that these guidelines are strongly grounded in current empirical evidence, a point also corroborated by the present review of original literature. Overall, our findings portray an unsatisfactory situation which threatens the integrity of CAR findings and likely limits further scientific progress in this area of research. It will thus be of particular importance to discuss potential reasons for these findings next to strategies for improvements in future research.

5.1 Use of objective control strategies

Within the present results, findings concerning the very low usage of objective strategies to verify awakening and sampling times stand out particularly. While there was a trend for an improvement in the use of objective verification of sampling times from the PRE to POST-period, the absolute rate of 31.5% of recent studies following this guideline is still low. Even more so, only 13.0% of studies objectively verified awakening time and only 9.3% provided full objective control, as proposed in the guidelines report. Importantly these findings are limited to research articles published in the journal *PNEC* and no inferences with regard to other journals can be made. However, given that the original guidelines were published in *PNEC* and that a firm commitment to these guidelines was specific to this journal, we consider it unlikely that a completely different, more positive situation applies to research published in other journals.

It is difficult to delineate potential reasons for these findings. One possibility is that the time period between the publication of the guidelines report in early 2016 and the commencement of the present data analysis in the year 2018 was too short for researchers to adapt, e.g., larger studies may have already been running at the time when the guideline report was published and thus researchers were unable to initiate changes in due time. However, our data also show that even during the later years of 2019 and 2020 the rates of compliance control were still very low, which makes it unlikely that this overall finding was merely a matter of insufficient time to adapt. This is also suggested by the fact that other guidelines, which could have been adopted at any time, such as the proposition to report S1 data alongside the CAR (38.9% in POST studies), were also not consistently followed.

Another factor that is likely to have contributed to the present findings is that objective verification of awakening and sampling times is costly and that particularly larger studies may have struggled to cover these additional costs. With regard to this important issue, it first needs to be highlighted that recent technological advances raise hopes that costs for objective verification strategies can be cut down considerably without compromising the quality of verification data (see 4.1). Specifically, concerning the objective verification of *awakening times*, evidence suggests that recent-generation consumer grade sleep tracking devices (ideally using actigraphy and PPG sensors) may allow for sleep-wake detection at equal or even higher quality than conventional research grade actigraphy watches (Chinoy et al., 2022). Some of these devices are available at considerably lower costs than previous research grade solutions and may thus increase the feasibility of objective awakening time verification in CAR research, even in larger studies. Equally, regarding the objective verification of *sampling times*, the use of time-stamped photographs provides an excellent low-cost solution, which has now been applied successfully in recent CAR research (Zhu et al., 2019). When using this approach, smartphones' automatic photo time-stamp functions can be used. By taking

photos of only the pre-coded sampling devices (i.e., sparing participants' faces) potential privacy protection concerns are alleviated. Combined, the named technological advances should significantly increase the feasibility of using objective verification strategies in future CAR research. In addition, objective sleep tracking devices are re-useable across multiple participants and multiple studies, a factor which further brings down their cost per usage. In addition, investigators can look into sharing or pooling these resources across laboratories for greater cost-efficiency. Such cooperation and collaboration would also serve to increase communication of both CAR and sleep measurement guidelines across laboratories.

Overall, it is critical to highlight again that the use of full objective verification strategies is absolutely essential in order to obtain trustable CAR results and should not be sacrificed under any circumstances. The clear benefits of using objective monitoring strategies (i.e., reliable data, with a lower measurement error and an increased chance of detecting genuine relationships) completely outweigh the associated costs. If larger epidemiological studies have insufficient funds for realizing full objective verification strategies in the complete sample, it is recommended that high-quality CAR data are obtained within a well-defined subsample instead of collecting low-quality and potentially confounded data in the complete sample. Moreover, the availability of published guidelines, which firmly demand the use of such objective verification strategies in CAR research, puts researchers in a strong position when applying for respective funding for future CAR research.

5.2 Other guidelines: assessment days, number of samples, covariate control and data reporting

Next to compliance control, findings regarding the adherence to other guidelines revealed a more mixed picture, including some areas of improvement. Indeed, the number of samples used to assess the CAR increased to a median of three samples in the POST-period, which corresponds to the minimal protocol proposed in the guidelines report. Similarly, there was a significant increase in the reported rate of informing participants about unwanted morning behaviors, with 90.7% of articles/studies having informed participants about at least one type of behavior. Conversely, however, the number of days across which the CAR was assessed decreased to a mean of approx. 1.5 days in cross-sectional research published during the POST-period. This is problematic since the expression of the CAR on a particular day is predominantly state-dependent (e.g., Hellhammer et al., 2007; Almeida et al., 2009; Stalder et al., 2010), thus limiting the chance of detecting associations between the CAR and relevant trait-like/long-term measures, e.g., personality or chronic stress exposure. Indeed, the notion that an insufficient number of CAR assessment days is an important factor contributing to unreliable findings in CAR research (alongside small samples sizes) was also suggested by recent results of a simulation study (Segerstrom and Boggero, 2020).

Related to the notion of high state-specificity of CAR expression, another important guideline proposed that state/situational covariates should be assessed and controlled for in CAR research. Here, again no improvement was seen, with only 55.6% of studies published between 2018 and 2020 controlling for at least one state covariate. As controlling for relevant state covariates of the CAR (e.g., time of awakening, weekday vs. weekend, psychosocial factors surrounding the study day) is likely to reduce unwanted error variance as well as limit the possibility of state-related confounding (see detailed description in Stalder et al., 2016), this is another important aspect of research practice that holds room for improvement in future work.

Finally, another aspect for which no advance in current research practice was seen concerns the way that CAR data are reported. Abundant evidence suggest that the CAR is distinct from the pre-awakening cortisol increase (e.g., Wilhelm et al., 2007; review: Clow et al., 2010) leading to the proposition that these two circadian aspects should be distinguished. Specifically, results of both the first sample on awakening (S1), i.e., the endpoint of the pre-awakening increase, and of the CAR should be reported alongside. However, our results revealed that only 89.1% of recent CAR studies reported an actual measure of the CAR, i.e., of the dynamic response to awakening, whereas the remaining articles reported results of a measure of total post-awakening cortisol levels (usually the AUC_G), which however was interpreted as representing the CAR. Even more so, results for S1 were only reported in 42.6% of articles. Future improvements in this area would be important in order to gain reliable information on the underlying mechanisms and specific correlates of these distinct aspects of circadian regulation.

5.3 Improvement strategies: Updated and streamlined guidelines

Results of the current quantitative evaluation showed unequivocally that the methodology of current CAR research needs further enhancement and that the 2016 guidelines report, while being highly cited, has not led to as much of a change in research practice as was hoped for. The present report aims to work on this situation in two ways: First, it is important to keep the guidelines up-to-date and thus we reviewed empirical research published over the past six years (see section 4) in order to formulate an updated, state-of-the-art version of the guidelines. This has largely led to a corroboration of central guidelines, but has also produced some interesting additions (e.g., technical advances regarding objective verification strategies) as well as some alleviations (e.g., recent evidence suggesting that influences of menstrual cycle phase on the CAR might be less important than initially conceived and that thus the respective guideline can be reduced). Second, beyond merely updating these guidelines, a potential weakness of the 2016 guidelines report was its overly comprehensive and non-weighted nature. The initial guidelines report sought to cover the respective ground as fully as possible, which led to a rather complex article structure with one main guidelines table, which cross-referenced four

other tables that provided more detailed, additional information. Furthermore, the 2016 report did not clearly distinguish between absolutely *essential* guidelines (i.e., necessary to obtain valid CAR data) and those that are important, but may be modifiable under specific circumstances. In order to counter these issues, an updated and streamlined version of the guidelines is provided in Table 2. An additional graphic representation of major guidelines in relation to pre- and post-awakening cortisol secretion is provided in Figure 3.

Table 2
Updated consensus guidelines for the assessment of the CAR in ambulatory settings and checklist for submissions to PNEC

Essential guidelines	Strategy applied:
Use an objective method to verify self-reported awakening times (consumer grade sleep-tracking device, research grade actigraphy, PSG, ECG/HR, or mixed)	☐
Use an objective method to verify self-reported sampling times (time-stamped photos, EMS/track caps, time-stamped codes, text-message-prompts with reply)	☐
Use strategies to increase participants' engagement with the study and to maximize adherence	☐
Incorporate information on verified inaccurate data into the CAR estimation model (preferred) or exclude inaccurate data (using a strict accuracy margin, $\Delta t = 0 \pm 5$ min)	☐
Provide clear instructions on unwanted morning behavior → nil by mouth other than water, no smoking, no exercising	☐
Postpone sampling in case of current illness, jet lag or shift work and exclude participants in case of glucocorticoid medication use, endocrine disorders or pregnancy (note: other medication use or disorders should be evaluated case-by-case)	☐
Assess and control for trait-like covariates (mainly: sex and age) and state influences (mainly: time of awakening, weekday-weekend, psychosocial states/anticipation of the upcoming day)	☐
Only refer to the *dynamic* increase of post-awakening cortisol as the CAR (e.g., AUC_I, Max-min, MnInc, 0-30 delta), but not to measures of total post-awakening cortisol (e.g., AUC_G)	☐
Other important recommendations	
Inform participants about the use of objective monitoring strategies	☐
Use a self-report diary system to complement objective data	☐
In adult research, a minimum of three post-awakening samples should be used (0, 30, 45 min)	☐
In cross-sectional research: The CAR should be assessed over as many days as possible (at least 2 days) to obtain more reliable trait estimates	☐
Acknowledge the distinctive nature of the CAR (i.e., it is not a general marker of HPA axis)	☐
Report associations with both S1 and the CAR	☐

AUC_G, area under the curve with respect to ground; AUC_I, area under the curve with respect to increase; CAR, cortisol awakening response; ECG, electrocardiography; EMS, electronic monitoring systems; HR, heart rate; MnInc, mean increase; PSG, polysomnography

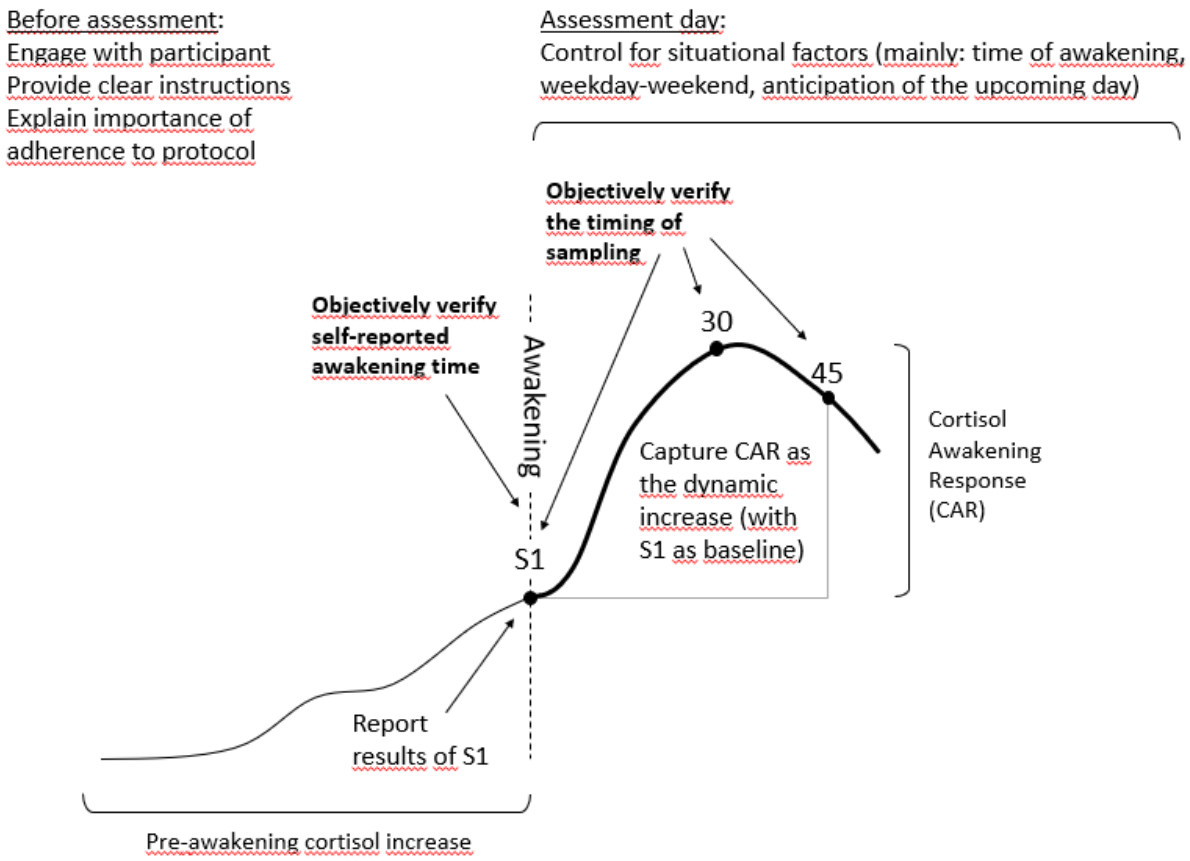

Figure 3. Graphic representation of major guidelines in relation to pre- and post-awakening cortisol secretion.

5.4 Improvement strategies: Compulsory guidelines checklist for future PNEC submissions

Beyond updating and streamlining the guidelines, correspondence with the *PNEC* editorial team about this matter led to the initiation of another strategy seeking to advance the quality of future CAR research published in this journal. This strategy targets the scientific reviewing process, which is deemed to play a critical role for ensuring a high level of methodological standards. Specifically, the goal is to maximize the transparency and accessibility of information on methodological quality of CAR assessments, so that reviewers are assisted in making an informed judgement on this matter. To this end, authors submitting an article to *PNEC* in the future will be required to fill in a checklist based on the present guidelines (Table 2), thus gathering transparent information on the utilized CAR methodology. This checklist will then be submitted alongside the article and assist reviewers in their decision making. This approach will ensure that a complete picture of methodological information is provided. It will also prevent against instances of 'cherry-picking', i.e., articles attesting that the CAR guidelines were followed by selectively reporting some commendable, but more minor aspects of CAR methodology (e.g., informing participants about unwanted morning behaviors) while not mentioning that other essential guidelines were not adhered to (chiefly, objective compliance verification). It is

hoped that this additional step will facilitate high quality reviewing practices and thus contribute to a long-term improvement in CAR assessment methodology.

Conclusion

The present manuscript evaluated and critically reviewed the current state of CAR research methodology. While our findings regarding the impact of the 2016 consensus guidelines on the quality of recently published CAR research generally paint a disappointing picture, recent developments also leave room for hope concerning the methodological quality of CAR assessments in the future. The preparedness of the *PNEC* editorial team to demand from potential authors the submission of a compulsory methodological checklist for future CAR research will increase transparency and enable reviewers to readily assess the quality of CAR data. Furthermore, it is hoped that the introduction of an updated and streamlined version of the CAR guidelines (incl. a graphical representation) will make it easier for researchers to gather key information. Beyond this, recent technological advances also make an important contribution by likely reducing the costs of using objective methods for awakening and sampling time verification in future research, thus rendering it feasible to implement this critical methodological step even in larger research with a restricted budget.

When it was first described some twenty-five years ago, the discovery that the CAR could easily be measured from self-collected saliva samples within the domestic setting was truly groundbreaking for psychoneuroendocrinological research. Assessment of the CAR in this way provides high ecological validity and opened the way for much fascinating neuroendocrine research probing a very diverse range of questions. Much has been achieved; however, it is necessary to take stock and balance the ease of use for this popular measure with measurement validity. Since its discovery we have learned the crucial importance of timing and state influences in CAR measurement. To ignore these issues and perpetuate the notion that CAR measurement is 'easy' will ultimately undermine its value as a meaningful construct. The methodological guidelines described here might appear onerous, but the authors unanimously agree that implementation of the presented guidelines is both essential and achievable. The dynamic CAR is an intriguing and important aspect of circadian function related to health and aging. It deserves rigorous investigation that will undoubtedly go on to enlighten much yet unknown.

References

Adam, E.K., Heissel, J.A., Zeiders, K.H., Richeson, J.A., Ross, E.C., Ehrlich, K.B., Levy, D.J., Kemeny, M., Brodish, A.B., Malanchuk, O., Peck, S.C., Fuller-Rowell, T.E., Eccles, J.S., 2015. Developmental histories of perceived racial discrimination and diurnal cortisol profiles in adulthood: A 20-year prospective study. Psychoneuroendocrinology 62. https://doi.org/10.1016/j.psyneuen.2015.08.018

Agresti, A., Caffo, B., 2000. Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. Am. Stat. 54. https://doi.org/10.1080/00031305.2000.10474560

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. https://doi.org/10.1037/a0017910

Bäumler, D., Kliegel, M., Kirschbaum, C., Miller, R., Alexander, N., Stalder, T., 2014. Effect of a naturalistic prospective memory-related task on the cortisol awakening response in young children. Biol. Psychol. 103. https://doi.org/10.1016/j.biopsycho.2014.08.005

Beddig, T., Reinhard, I., Kuehner, C., 2019. Stress, mood, and cortisol during daily life in women with Premenstrual Dysphoric Disorder (PMDD). Psychoneuroendocrinology 109. https://doi.org/10.1016/j.psyneuen.2019.104372

Benz, A., Meier, M., Mankin, M., Unternaehrer, E., Pruessner, J.C., 2019. The duration of the cortisol awakening pulse exceeds sixty minutes in a meaningful pattern. Psychoneuroendocrinology 105. https://doi.org/10.1016/j.psyneuen.2018.12.225

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

Boileau, K., Barbeau, K., Sharma, R., Bielajew, C., 2019. Ethnic differences in diurnal cortisol profiles in healthy adults: A meta-analysis. Br. J. Health Psychol. 24. https://doi.org/10.1111/bjhp.12380

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.

Chinoy, E.D., Cuellar, J.A., Jameson, J.T., Markwald, R.R., 2022. Performance of Four Commercial Wearable Sleep-Tracking Devices Tested Under Unrestricted Conditions at Home in Healthy Young Adults. Nat. Sci. Sleep 14. https://doi.org/10.2147/NSS.S348795

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

Clow, A., Thorn, L., Evans, P., Hucklebridge, F., 2004. The awakening cortisol response: methodological issues and significance. Stress 7, 29–37. https://doi.org/10.1080/10253890410001667205

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.

Dockray Samantha, S., Steptoe, A., 2011. Chronotype and diurnal cortisol profile in working women: Differences between work and leisure days. Psychoneuroendocrinology 36.

https://doi.org/10.1016/j.psyneuen.2010.09.008

Edwards, S., Evans, P., Hucklebridge, F., Clow, A., 2001. Association between time of awakening and diurnal cortisol secretory activity. Psychoneuroendocrinology 26, 613–622.

Elder, G.J., Barclay, N.L., Wetherell, M.A., Ellis, J.G., 2018. Anticipated next-day demand affects the magnitude of the cortisol awakening response, but not subjective or objective sleep. J. Sleep Res. 27. https://doi.org/10.1111/jsr.12569

Elder, G.J., Ellis, J.G., Barclay, N.L., Wetherell, M.A., 2016. Assessing the daily stability of the cortisol awakening response in a controlled environment. BMC Psychol. 4. https://doi.org/10.1186/s40359-016-0107-6

Evans, P., Smyth, N., Thorn, L., Hucklebridge, F., Clow, A., 2019. Salience versus magnitude in the measurement of the cortisol awakening response. Psychoneuroendocrinology 103. https://doi.org/10.1016/j.psyneuen.2019.01.023

Griefahn, B., Robens, S., 2008. The cortisol awakening response: A pilot study on the effects of shift work, morningness and sleep duration. Psychoneuroendocrinology 33. https://doi.org/10.1016/j.psyneuen.2008.04.004

Haghayegh, S., Khoshnevis, S., Smolensky, M.H., Diller, K.R., Castriotta, R.J., 2019. Accuracy of wristband fitbit models in assessing sleep: Systematic review and meta-analysis. J. Med. Internet Res. https://doi.org/10.2196/16273

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.

Kayacan, Y., Makaracı, Y., Ozgocer, T., Ucar, C., Yıldız, S., 2021. Cortisol Awakening Response and Heart Rate Variability in the Menstrual Cycle of Sportswomen. Res. Q. Exerc. Sport 92. https://doi.org/10.1080/02701367.2020.1774486

Kirschbaum, C., Hellhammer, D.H., 1994. Salivary cortisol in psychoneuroendocrine research: recent developments and applications. Psychoneuroendocrinology 19, 313–333. https://doi.org/10.1016/0306-4530(94)90013-2

Kovac, K., Vincent, G.E., Paterson, J.L., Reynolds, A., Aisbett, B., Hilditch, C.J., Ferguson, S.A., 2021. The impact of a short burst of exercise on sleep inertia. Physiol. Behav. 242. https://doi.org/10.1016/j.physbeh.2021.113617

Kramer, A.C., Neubauer, A.B., Stoffel, M., Voss, A., Ditzen, B., 2019. Tomorrow's gonna suck: Today's stress anticipation predicts tomorrow's post-awakening cortisol increase. Psychoneuroendocrinology 106. https://doi.org/10.1016/j.psyneuen.2019.03.024

Kudielka, B.M., Bellingrath, S., Hellhammer, D.H., 2007. Further support for higher salivary cortisol levels in "morning" compared to "evening" persons. J. Psychosom. Res. 62. https://doi.org/10.1016/j.jpsychores.2007.03.016

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. https://doi.org/10.1097/01.PSY.0000058374.50240.BF

Kudielka, B.M., Federenko, I.S., Hellhammer, D.H., Wüst, S., 2006. Morningness and eveningness: The free cortisol rise after awakening in "early birds" and "night owls." Biol. Psychol. 72, 141–146. https://doi.org/10.1016/j.biopsycho.2005.08.003

Kunz-Ebrecht, S.R., Kirschbaum, C., Marmot, M., Steptoe, A., 2004. Differences in cortisol awakening

response on work days and weekends in women and men from the Whitehall II cohort. Psychoneuroendocrinology 29, 516–528. https://doi.org/10.1016/S0306-4530(03)00072-6

Law, R., Hucklebridge, F., Thorn, L., Evans, P., Clow, A., 2013. State variation in the cortisol awakening response. Stress 16, 483–92. https://doi.org/10.3109/10253890.2013.817552

Miller, R., Wojtyniak, J.G., Weckesser, L.J., Alexander, N.C., Engert, V., Lehr, T., 2018. How to disentangle psychobiological stress reactivity and recovery: A comparison of model-based and non-compartmental analyses of cortisol concentrations. Psychoneuroendocrinology 90. https://doi.org/10.1016/j.psyneuen.2017.12.019

Ozgocer, T., Ucar, C., Yildiz, S., 2022. Daily cortisol awakening response and menstrual symptoms in young females. Stress Heal. 38. https://doi.org/10.1002/smi.3074

Ozgocer, T., Ucar, C., Yildiz, S., 2017. Cortisol awakening response is blunted and pain perception is increased during menses in cyclic women. Psychoneuroendocrinology 77. https://doi.org/10.1016/j.psyneuen.2016.12.011

Petrowski, K., Schmalbach, B., Niedling, M., Stalder, T., 2019. The effects of post-awakening light exposure on the cortisol awakening response in healthy male individuals. Psychoneuroendocrinology 108. https://doi.org/10.1016/j.psyneuen.2019.05.016

Petrowski, K., Schmalbach, B., Stalder, T., 2020. Morning and evening type: The cortisol awakening response in a sleep laboratory. Psychoneuroendocrinology 112. https://doi.org/10.1016/j.psyneuen.2019.104519

Powell, D.J., Schlotz, W., 2012. Daily Life Stress and the Cortisol Awakening Response: Testing the Anticipation Hypothesis. PLoS One 7. https://doi.org/10.1371/journal.pone.0052067

Powell, D.J.H., Moss-Morris, R., Liossi, C., Schlotz, W., 2015. Circadian cortisol and fatigue severity in relapsing-remitting multiple sclerosis. Psychoneuroendocrinology 56. https://doi.org/10.1016/j.psyneuen.2015.03.010

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.

Randler, C., Schaal, S., 2010. Morningness-eveningness, habitual sleep-wake variables and cortisol level. Biol. Psychol. 85. https://doi.org/10.1016/j.biopsycho.2010.04.006

Rodriguez, K.E., Bryce, C.I., Granger, D.A., O'Haire, M.E., 2018. The effect of a service dog on salivary cortisol awakening response in a military population with posttraumatic stress disorder (PTSD). Psychoneuroendocrinology 98. https://doi.org/10.1016/j.psyneuen.2018.04.026

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. https://doi.org/10.1097/01.psy.0000116715.78238.56

Segerstrom, S.C., Boggero, I.A., 2020. Expected Estimation Errors in Studies of the Cortisol Awakening Response: A Simulation. Psychosom. Med. 82. https://doi.org/10.1097/PSY.0000000000000850

Skoluda, N., Linnemann, A., Nater, U.M., 2016. The role of week(end)-day and awakening time on cortisol and alpha-amylase awakening responses. Stress 19. https://doi.org/10.1080/10253890.2016.1174850

Smyth, N., Thorn, L., Hucklebridge, F., Clow, A., Evans, P., 2016. Assessment of the cortisol awakening response: Real-time analysis and curvilinear effects of sample timing inaccuracy. Psychoneuroendocrinology 74. https://doi.org/10.1016/j.psyneuen.2016.09.026

Stalder, T., Evans, P., Hucklebridge, F., Clow, A., 2010. State associations with the cortisol awakening response in healthy females. Psychoneuroendocrinology 35. https://doi.org/10.1016/j.psyneuen.2010.02.014

Stalder, T., Kirschbaum, C., Kudielka, B.M., Adam, E.K., Pruessner, J.C., Wüst, S., Dockray, S., Smyth, N., Evans, P., Hellhammer, D.H., Miller, R., Wetherell, M.A., Lupien, S.J., Clow, A., 2016. Assessment of the cortisol awakening reponse: Expert consensus guidelines. Psychoneuroendocrinology 63, 414–432. https://doi.org/10.1016/j.psyneuen.2015.10.010

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. https://doi.org/10.1016/j.psyneuen.2003.08.005

Weitzman, E.D., Fukushim.D, Nogeire, C., Roffwarg, H., Gallaghe.Tf, Hellman, L., 1971. Twenty-four Hour Pattern of Episodic Secretion of Cortisol in Normal Subjects. J. Clin. Endocrinol. Metab. 33, 14–22.

Wetherell, M.A., Lovell, B., Smith, M.A., 2015. The effects of an anticipated challenge on diurnal cortisol secretion. Stress 18, 42–48. https://doi.org/10.3109/10253890.2014.993967

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. https://doi.org/10.1016/j.psyneuen.2007.01.008

Wolfram, M., Bellingrath, S., Kudielka, B.M., 2011. The cortisol awakening response (CAR) across the female menstrual cycle. Psychoneuroendocrinology 36, 905–912. https://doi.org/10.1016/j.psyneuen.2010.12.006

Zhu, Y., Chen, X., Zhao, H., Chen, M., Tian, Y., Liu, C., Han, Z.R., Lin, X., Qiu, J., Xue, G., Shu, H., Qin, S., 2019. Socioeconomic status disparities affect children's anxiety and stress-sensitive cortisol awakening response through parental anxiety. Psychoneuroendocrinology 103. https://doi.org/10.1016/j.psyneuen.2019.01.008