Psychoneuroendocrinology (2009) 34, 1423—1436

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/psyneuen

# Assessing salivary cortisol in large-scale, epidemiological research

Emma K. Adam [a,1], Meena Kumari [b,1,*]

[a] School of Education and Social Policy and Institute for Policy Research, Northwestern University, United States
[b] International Centre for Lifecourse Studies in Society and Health, Department of Epidemiology and Public Health, University College London, London WC1E 6BT, United Kingdom

Received 4 April 2008; received in revised form 1 June 2009; accepted 17 June 2009

**KEYWORDS**
Salivary cortisol;
Epidemiology;
Population-based;
Measurement;
Modeling

**Summary** Salivary cortisol measures are increasingly being incorporated into large-scale, population-based, or epidemiological research, in which participants are selected to be representative of particular communities or populations of interest, and sample sizes are in the order of hundreds to tens of thousands of participants. These approaches to studying salivary cortisol provide important advantages but pose a set of challenges. The representative nature of sampling, and large samples sizes associated with population-based research offer high generalizability and power, and the ability to examine cortisol functioning in relation to: (a) a wide range of social environments; (b) a diverse array individuals and groups; and (c) a broad set of pre-disease and disease outcomes. The greater importance of high response rates (to maintain generalizability) and higher costs associated with this type of large-scale research, however, requires special adaptations of existing ambulatory cortisol protocols. These include: using the most efficient sample collection protocol possible that still adequately address the specific cortisol-related questions at hand, and ensuring the highest possible response and compliance rates among those individuals invited to participate. Examples of choices made, response rates obtained, and examples of results obtained from existing epidemiological cortisol studies are offered, as are suggestions for the modeling and interpretation of salivary cortisol data obtained in large-scale epidemiological research.


## 1. Introduction

The addition of biological measures (''biomarkers'') to large-scale social science and epidemiological studies has recently been advocated by a number of funding bodies (Finch et al., 2001). In this interdisciplinary approach, extensive information on social and behavioural processes, health behaviours and self-reported health in existing large-scale, representative samples, is complemented by the addition of objective biological markers of physiological processes and pre-disease states.

Measures of the hypothalamic pituitary adrenal axis (HPA), and in particular levels of salivary cortisol are obvious candidates for inclusion in these studies, given the important role that the HPA axis plays in 'transducing' subjective social—environmental experience into physiological changes relevant

* Corresponding author.
*E-mail address:* M.Kumari@ucl.ac.uk (M. Kumari).
[1] The authors made equal contributions to this paper.

0306-4530/$ — see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.psyneuen.2009.06.011

to health. While salivary cortisol has been successfully gathered in a number of large-scale, population-based studies, the design, collection and interpretation of salivary cortisol data in naturalistic settings presents numerous challenges. Here we discuss these challenges, and related solutions, and review some existing results that have emerged from this new wave of research.

## 2. Background

The HPA axis serves as an important pathway by which social and psychological factors influence biology and health. Stressful stimuli serve to activate HPA function to cause an increase in peripheral cortisol. Cortisol can effect physiological changes that encompass most of the main organ systems, helping to provide the energetic resources needed to face the stressor at hand, and also helping to modulate and contain other components of the physiological stress response (Sapolsky et al., 2000). Although short-term activations of the HPA axis are adaptive and necessary for everyday functioning, extreme, frequent or chronic activation of this system are associated with negative health outcomes. Existing research has implicated the HPA axis in the development of a variety of sub-clinical and clinical conditions including metabolic syndrome (Phillips et al., 1998; Brunner et al., 2002;), depression (Herbert et al., 2006; McEwen, 2007; Belmaker and Agam, 2008), risk for cardiovascular disease (Smith et al., 2005) and cognitive decline (Seeman et al., 1997).

Many early large-scale studies of cortisol and health focused on ''average'' cortisol measures, particularly urinary measures of cortisol, in part because such methods provided levels that are pooled across multiple hours, and total exposure to cortisol across 12 or 24 h was the measure of interest (Seeman et al., 2002). Increasingly, however, HPA-axis researchers are focusing on the marked diurnal rhythm in the release of cortisol, with various elements of this rhythm viewed as essential indicators of HPA axis functioning. The diurnal cortisol rhythm is typically characterised by high levels upon waking, a substantial (50–60%) increase in cortisol concentration in the 30–45 min after waking (called the cortisol awakening response or CAR), and a subsequent decline over the remainder of the day, reaching a low point or nadir around midnight (Weitzman et al., 1971; Kirschbaum and Hellhammer, 1989; Pruessner et al., 1997). Healthy HPA-axis function is thought to require the presence of strong diurnal patterning, and deviations from the typical diurnal cycle of cortisol provide valuable information regarding environmental influences on the HPA axis and the role of the HPA axis in disease processes (Stone et al., 2001).

The major elements of the diurnal cortisol rhythm that are typically assessed, and examples of research focusing on these parameters are shown in Table 1. The exact interpretation of each of the elements of diurnal HPA axis functioning is still subject to debate, but recent reviews are beginning to close in on the meaning and relevance of the different aspects of the rhythm (Saxbe, 2008; Chida and Steptoe, 2009). The size of the CAR, for example, has been correlated with a variety of psychosocial processes and health outcomes (Pruessner et al., 1997, 2003b; Clow et al., 2004; Steptoe et al., 2004; Adam et al., 2006; Nater et al., 2007). Both the absence of a CAR, or an atypically large CAR have been found in past research to associated with negative health outcomes. Flattening of the diurnal cortisol slope, indicated by a slower rate of decline in cortisol across the day, has been related to both chronic and acute psychosocial stress (Adam et al., 2006), sub-clinical disease (Matthews et al., 2006) and increased mortality from breast cancer (Sephton et al., 2000). There is therefore increasing accord among investigators that diurnal changes in cortisol are important, and that the CAR and diurnal slope are key elements of diurnal cortisol activity to measure, whether in smaller scale projects or in the type of large-scale epidemiological research described here.

## 3. Epidemiological research: strengths and challenges

Epidemiological research is characterized by: (1) representative sampling, in which the sample is carefully selected and retained to ensure that participants precisely reflect the characteristics of the larger population of interest; (2) large sample sizes (in the order of hundreds to tens of thousands), typically of a broader socioeconomic and racial/ethnic range than in samples of convenience; and (3) extensive measurement of and statistical control for potentially confounding variables.

As a result, epidemiological studies incorporating HPA axis measures have the potential to offer assessments of typical functioning of the HPA in healthy populations, norms regarding the functioning of the HPA axis in different racial/ethnic groups and by age and gender, and information on HPA-axis functioning in relation to the full range of social and economic conditions and the full range of pre-disease and disease states present in the general population.

The large sample sizes employed provide the power to identify associations that may be missed in smaller samples of convenience due to low power. Studies examining diurnal cortisol secretion frequently have participant numbers in the tens (Steptoe et al., 2004; Eller et al., 2005) or hundreds (Kajantie et al., 2002), which may be sufficient to examine associations where there is a very high expected incidence rate (Sephton et al., 2000), as when participants are selected for the presence of particular disease states. To study naturally occurring disease states in the full population, however, such numbers are rarely sufficient (see Table 2).

For example, a study of over 4000 participants is required to have sufficient power (90% at 5% level of significance) to detect a rate of development of disease that is 1.5 times greater in a high risk group compared to a low risk group for a disease with a prevalence of 5–10% (such as is the case with cardiovascular disease). These calculations assume a linear relationship between the exposure (e.g. a flatter diurnal cortisol rhythm at baseline) and increasing risk for disease. For more prevalent outcomes, such as depressive symptoms or hypertension (prevalence ~30%), the study would require some 460 participants. Thus, even for relatively prevalent disease states, many existing cortisol studies would not be

**Table 1** Diurnal cortisol measures frequently used in field-based research.

| Diurnal cortisol measure | Description and brief interpretation (see citations for more detail) | Selected citations (citations additional to those found in main text are listed in 'further reading') |
|---|---|---|
| Cortisol awakening response (CAR)[a] | Size of post-awakening surge in cortisol that occurs in the 30–45 min after waking. *Both heightened and blunted CARs have been related to psychosocial stress and poor health outcomes* | Pruessner et al. (1997); Chida and Steptoe (2009); Clow et al. (2004) |
| Diurnal cortisol slope[a] | Degree of change (typically decline) in cortisol levels from early morning to late evening. *Steeper decline typically associated with better psychosocial and physical health* | Adam and Gunnar (2001); Adam et al. (2006); Cohen et al. (2006) |
| Area under the daytime cortisol curve[a] | Area under all the cortisol data points measured across the waking day. *An estimate of average cortisol exposure, provides no indication of diurnal change* | Badrick et al. (2007); Nicolson (2004) |
| Waking cortisol | Level take as soon as possible after waking (immediately after opening eyes). *Low levels contribute to flatter cortisol slopes* | Kumari et al. (2009); Cohen et al. (2006); DeSantis et al. (2007) |
| Cortisol at specific time points across the waking day | Levels taken at specific clock times (e.g. 10 AM or 10 PM) or time points (e.g. afternoon). *Difficult to interpret unless also know waking and bedtime levels and time of waking that day, but elevated afternoon and evening levels thought to be problematic* | Powell et al. (2002); Lupien et al. (2001); Essex et al. (2002) |
| Bedtime cortisol | Level of cortisol at bedtime. *High levels contribute to flatter diurnal cortisol slopes* | Pendry and Adam (2007); Cohen et al. (2006) |
| Cortisol reactivity to momentary stressors | Elevations in cortisol above typical diurnal level for that time of day for that person. *Have been associated with concurrent or immediately prior negative mood states* | Adam (2006); van Eck et al. (1996); Smyth et al. (1998) |
| Cortisol reactivity to daily stressors | Changes in cortisol levels or rhythms from 1 day to the next. *Associated with changes in daily demands and events* | Adam et al. (2006); Rohleder et al. (2007); Schlotz et al. (2004) |

[a] These three measures are the most commonly employed, and the most robustly related to psychosocial phenomena and to health outcomes.

sufficiently powered to examine the issue, and large-scale, epidemiological studies stand to play an important role.

Large-scale epidemiological studies also have sufficient power and degrees of freedom available to employ multiple covariates for confounding influences that may account for, or obscure, associations between social exposures, HPA-axis variables and disease states. A list of covariates commonly employed in HPA axis research is shown in Table 3; the large samples sizes associated with epidemiological research allow

**Table 2** Total number of study participants required to have power of 90% to detect a significance difference at 5% level (two-sided).

| | Exposure prevalence (i.e. Top 50% vs. bottom 50%) | | | | | |
|---|---|---|---|---|---|---|
| | Rate ratio | | | | | |
| | 1.4 | 1.5 | 1.6 | 1.8 | 2.0 | 2.5 |
| Disease prevalence (in low risk group) | | | | | | |
| 5% | 6120 | 4094 | 2966 | 1804 | 1240 | 644 |
| 10% (overall prevalence) | 2868 (12%) | 1914 (12.5%) | 1382 (13%) | 836 (14%) | 572 (15%) | 292 (17.5%) |
| 20% (overall prevalence) | 1244 (24%) | 824 (25%) | 590 (26%) | 352 (28%) | 236 (30%) | 114 (35%) |
| 30% (overall prevalence) | 702 (36%) | 460 (37.5%) | 326 (39%) | 190 (42%) | 124 (45%) | ~100 |

*Note*. The sample size required to have a power of 90% to detect a significant difference at 5% (two-sided) is shown for varying disease prevalences and rate ratios. A rate ratio of 1.5 means that the expected rate of disease is 1.5 times greater in the at risk group compared to the low risk group.

**Table 3** Covariates that have been related to diurnal cortisol secretion[a].

| Covariates | Selected Citations (citations additional to those found in main text are listed in 'further reading') |
|---|---|
| Typically collected on day of sampling | |
| Schedule | |
| Whether weekend vs. weekday | Schlotz et al. (2004); Kunz-Ebrecht et al. (2004) |
| Time of waking | Kudielka and Kirschbaum (2003) |
| Lifestyle | |
| Smoking | Badrick et al. (2007) |
| Alcohol use | Badrick et al. (2008); Adam et al. (2006) |
| Caffeine | Lovallo et al. (2005) |
| Exercise/activity level | Kertes and Gunnar (2004) |
| Perceived stress and negative mood | Adam et al. (2006); Schlotz et al. (2006) |
| Recent meal | Follenius et al. (1982) |
| Not specific to day of sampling: | |
| Demographic | |
| Gender | Kirschbaum et al. (1999) |
| Socioeconomic Status | Cohen et al. (2006) |
| Race/ethnicity | DeSantis et al. (2007); Cohen et al. (2006) |
| Age | Ice (2005) |
| Pubertal stage | Adam (2006) |
| Medical/health | |
| Physical health diagnoses | Kudielka and Kirschbaum (2003) |
| Mental health diagnoses | Gillespie and Nemeroff (2005) |
| Body mass index/obesity | Therrien et al. (2007) |
| Medication use | Hibel et al. (2007) |
| Menstrual timing | Kirschbaum et al. (1999) |
| Typical exclusion criteria: | |
| Use of steroid-based medications | Masharani et al. (2005) |
| 3rd Trimester of pregnancy | Obel et al. (2005) |
| Illness on days of testing | Adam (2006) |
| Presence of endocrine disorder | Raff et al. (1998) |

[a] This list is not exhaustive, but highlights the frequently examined covariates.

investigators to account for a large number of these potential confounds.

## 4. Methodological priorities in epidemiological research: design implications

These defining features of epidemiological research, and their relative advantages, also require a shift in several priorities regarding the design of cortisol data collection protocols for large-scale research.

### 4.1. First, costs per participant matter more

The number of cortisol samples that can be collected per person is strongly limited by cost considerations. Collecting one extra saliva sample in a study of 100 people, assuming $5—10 per sample in collection and assay costs, costs an additional $500—1000. An extra sample in a study of 10,000 people, assuming the same unit costs, costs $50,000—100,000 (a sum that could fully fund several small lab-based studies). Thus, while typically, ''more is better'' when it comes to cortisol sampling protocols, in large-scale epidemiological research, the scientific gains associated with additional samples must outweigh the tremendous financial costs involved. The decision often comes down to administering a minimal or reduced-form salivary cortisol protocol, or not gathering salivary cortisol data at all.

### 4.2. Second, attrition and response rates matter more

Given the expense and effort required to recruit and retain epidemiological samples, the potential impact of adding salivary sampling on participant burden and attrition rates weigh heavily in investigators' decisions regarding salivary sampling protocols. Keeping protocols to a manageable size, so that they are likely to be completed by all or most participants, and ensuring participants have a pleasant experience are essential, in order to ensure high completion and retention rates that will maintain the representative nature of the sample.

## 5. Existing epidemiological cortisol research

While many salivary cortisol protocols have been implemented in convenience samples, we focus here on studies in which participants were purposefully sampled to accurately reflect the characteristics of a particular population of interest, such as samples chosen to be representative of a particular occupational status or age group (Ranjit et al., 2005; Rosmalen et al., 2005; Cohen et al., 2006; Badrick et al., 2007). The sample sizes, basic demographic characteristics, ages, data collection protocols employed, and response rates obtained by the population-based salivary cortisol studies we are aware of, and could obtain appropriate statistics for, are listed in Table 4. Population-based studies that have incorporated salivary sample collection into their protocols, have, for the most part, been successful, although very few have been able to achieve the types of response rates typically obtained for questionnaire or interview data, with some studies achieving rates that are clearly unsatisfactory by epidemiological standards (<50%), and others achieving excellent response rates (>90%).[2]

### 5.1. Saliva sampling protocol choices

The typical methodologies used in existing salivary cortisol studies range from a 'minimal protocol' including just three data points per person on a single day (Women's Employment Study), to medium intensity protocols involving six samples on a single day (Whitehall II, CARDIA) or three samples per day across 3 days (Chicago, Health, Aging and Social Relations Study), to moderately high intensity protocols involving multiple samples per day for multiple days (National Study of Daily Experiences; 4 samples per day for 4 days, for a total of 16 per person; Multi-ethnic study of atherosclerosis 6 samples per day for 3 days, for a total of 18 per person).

Importantly, most of these studies show a reduction in the total number of samples requested when compared to naturalistic cortisol studies of smaller convenience samples, which have collected 18 (DeSantis et al., 2007) and even up to 50 (van Eck et al., 1996) saliva samples per person. They also show a reduction in the number of samples used to capture specific elements of the diurnal cortisol curve such as the cortisol awakening response. The original protocol for the CAR, where the sample size under investigation was 152, involved sampling every 15 min for the first hour post-awakening (Pruessner et al., 1997), whereas the majority of these studies rely on two samples for this estimate (typically one samples at waking, and one sample 30—45 min after waking).

One study (1958 British Birth Cohort) used a protocol involving only two samples total (45 min after waking and 3 h later). Although a reasonable decision at the time, recent findings on the importance of the CAR makes this protocol inadvisable for future studies, as the CAR cannot be estimated without the presence of a wakeup sample.

For the most part the sampling time points for these studies have been carefully selected to retain the ability to obtain at least some estimate of the key diurnal cortisol parameters described earlier—the size of the awakening response (CAR), the degree of decline in cortisol across the day (Slope), and the total or average level of cortisol across the day (AUC). Fourteen out of the seventeen studies reviewed here utilized protocols that included a wakeup, post-awakening peak, and bedtime (or evening) sample, allowing them to estimate each of these diurnal cortisol measures.[3] Most studies employed 1-day protocols, but 4 of the 14 gathered wake (W), peak (P) and evening measures (CHASRS, National Study of Daily Experiences, MESA, CaPS) did so over more than one day.

Three studies utilized the "minimal protocol" involving only W, P, and B (or evening) measures (TRAILS, CHASRS, L.A. FANS), one study collected W and P only (AGES). The rest gathered additional data points beyond these three. In some cases, the additional samples were taken on another day during an in-home interview or laboratory visit (e.g.Women's Employment Study, Women's Health and Aging Study II). In other cases (e.g. Whitehall II, National Study of Daily Experiences, CARDIA, ELSA, NESDA, CaPs), additional samples are taken during the late morning and afternoon on the same day as the W, P, and B sampling, allowing for better definition of the shape of the diurnal cortisol curve.

The extent to which "stripped down" measures of cortisol parameters, such as those based on the three sample "minimal protocol", are a reasonable substitute for estimates resulting from more intensive sampling requires additional research attention. One study found that the AUC derived from 15 samples over 3 days correlates approximately 0.69 with the AUC derived from a 1-day minimal (W, P, B) protocol (Harville et al., 2007). Another study finds that the association between diurnal cortisol slopes based on two data points on 2 days (wakeup and bedtime) correlates 0.94 with cortisol slopes based on 6—7 samples per day over 2 days (Adam, in preparation).

Reducing a protocol to three samples on 1 day precludes the ability to examine the curvilinear nature of the decline in cortisol levels across the day, an aspect of diurnal decline which is typically modeled when more samples are available (Adam, 2006). This choice also precludes the ability to examine momentary cortisol reactions to naturalistic stressors throughout the day, which requires multiple measures of cortisol across the day gathered in relation to repeated dairy reports of events and emotions (van Eck et al., 1996; Adam, 2005, 2006; Steptoe et al., 2008). A related approach, which is more practicable in epidemiological research, is to collect experiential data not at the time of each sample, but once per day in a study in which multiple days of cortisol data are gathered. This allows investigators to examine how changing

[2] One additional large population based study that we are aware of the National Longitudinal Study of Adolescent Health (Add Health) piloted a three-sample salivary cortisol self-collection protocol in the 2007 Wave IV pretest (N = 193). Although 97% of the sample consented to sample provision, and 75% of those consenting actually mailed in samples, other data quality indicators were considered unsatisfactory. Saliva collection for cortisol measurement was therefore not maintained in the main data collection effort. Details on collection protocols and analysis of quality indicators are forthcoming from Add Health investigators (personal communication, Kathleen Mullan Harris, 3 March 2008).

[3] The CAR can be obtained from such data by subtracting the 30 min post-awakening sample from the wakeup sample, and the diurnal slope by subtracting the wakeup from the bedtime value, and dividing by the total time awake. A rough estimate of the AUC can be derived by taking the area under three data points.

**Table 4** Examples of study designs utilized and response rates obtained in large-scale epidemiological surveys.

| Study | Population characteristics | | | Response rate | Protocol | Timing of samples | #of full days | Reference/website (citations not found in main text are listed in 'further reading') |
|---|---|---|---|---|---|---|---|---|
| | Approx N | Age range | Socioeconomic status | % who returned samples | # of saliva samples collected | Waking (W), peak (P),[a] bedtime (B) | | |
| 1. Whitehall II | 4600 | 59–72 | Occupational cohort | 90 | 6 | W, P, B, 3× across day | 1 | Badrick et al. (2008) |
| 2. ELSA | 4744 | 50–80 | Full range | 80 | 4 | W, P, B plus 7 PM | 1 | www.ifs.org.uk/elsa |
| 3. Rotterdam study | 2000 | 67–100 | Full range | 90 | 4 | W, P, B plus 1 evening | 1 | Teimeier and Dekker (personal communication, March 6, 2008) |
| 4. 1958 British Birth Cohort | 6335 | 44–45 | Birth cohort, full range | 69 | 2 | P, 3 h after peak | 1 | Power et al. (2006) |
| 5. CARDIA | 718 | 40 | Full range | 84 | 6 | W, P, B, and 3× across day | 1 | Cohen et al. (2006) |
| 6. Gothenburg study | 222 | 9–11 | Rural | 25 | 5 | W, P, B, 9 AM, 11 AM | 1 | Osika et al. (2007) |
| 7. Women's Employment Study | 298 | 18–54 | Low income | 63 | 4 | W, P, B and at time of interview | 1 | Ranjit et al. (2005) |
| 8. TRAILS | 1768 | 10–12 | Full range | 79 | 3 | W, P, 8 PM | 1 | Rosmalen et al. (2005) |
| 9. CHASRS | 229 | 50–68 | Full range | 74 | 3 | W, P, B | 3 | Adam et al. (2006) |
| 10. L.A. FANS | 600–650 | 3–17 | Full range | In progress | 3 | W, P, B | 1 | Sastry et al. (2006), Narayan Sastry (personal communication, Feb. 22, 2008) http://www.lasurvey.rand.org/ |
| 11. National Study of Daily Experiences | 1200 | 33–85 | Full range | In progress, currently >85% | 4 | W, P, B and before lunch | 4 | Almeida et al. (2007) http://gerontology.ssri.psu.edu/nsde |
| 12. Cebu Longitudinal Health & Nutrition Survey | 1800 | 21–22 | Birth cohort, full range | 93 | 3 | W, P, B | 1 | Adair et al. (2001) www. cpc. unc. edu/proj ects/cebu/ |
| 13. Women's Health & Aging Study II | 238 | 80–90 | Full range | 90 | 7 | W[b], B, 5× during A.M. lab visit | 1 | Varadhan et al. (2008) |
| 14. MESA | 1000[c] | 45–84 | Full range | In progress >85% | 6 | W, P, B | 5 | http://www.mesa-nhlbi.org/ |
| 15. AGES-Reykjavik | 5764 | 66–98 | Full range | 89% | 2 | P, B | 1 | Harris et al. (2007) |
| 16. NESDA | 2981 | 18–65 | Full range | 74 | 7 | W, P (30 m), P (45 m), W+1 h, 10 PM, 11 PM, W after 0.5 mg dexamethasone | 1 | www.nesda.nl Penninx et al. (2008); Vreeburg et al. (2009) |
| 17. CaPS | 1100 | 64–80 | Full range | 86.4 | 4 | W P 1400, 2200 | 2 | Yoav Ben Shlomo (personal communication, Sept. 25th 2009) |

*Note*. ELSA = English Longitudinal Study of Ageing; CARDIA = Coronary Artery Risk Development in Young Adults Study; TRAILS = Tracking Adolescents' Individual Lives Survey; CHASRS = Chicago Health Aging and Social Relations study; L.A. FANS = The Los Angeles Family and Neighborhood Survey; MESA = Multi-Ethnic Study of Atherosclerosios; AGES-Reykjavik = The Age, Gene/Environment Susceptibility-Reykjavik Study.

[a] Peak sample = a cortisol sample taken between 30 and 45 min post-awakening. Studies vary on the exact timing of this sample.

[b] The waking sample in this study was requested between 0 and 30 min of waking.

[c] This study is being conducted on a random subsample of 1000 participants selected from those available in the current wave of data collection in MESA.

experiences from day to day are related to daily changes in diurnal cortisol rhythms (CHASRS and NSDE use this approach).

Gathering data on only 1 day, rather than multiple days, has implications for the reliability of measurement of the cortisol measures. Cross-sectional studies suggest that it may be necessary to collect up to six consecutive days of samples to assess the CAR, as a single day collection biases the CAR to state rather than trait characteristics (Hellhammer et al., 2007). Where study response rates and participant burden cannot tolerate such a demanding design, researchers need to be aware of this potential bias.

### 5.2. Consenting and administering protocols

Most of the studies in Table 4 introduced and consented participants for their saliva collection as part of face-to-face interview and then left written instructions and the saliva sample collection kit to be completed soon after the interview and returned by regular mail or courier. The saliva sample collection kit typically contains instructions, the saliva collection device (Salivettes®, or straw and vial for passive drool) and a 'log book' in which participants can record sampling times and experiences on the day of sample collection. The protocol is therefore self-administered. One study (National Study of Daily Experiences) did not have any face-to-face contact with participants in explaining the procedures, instead sending the sampling kit to participants by mail and collecting daily diary data each evening by phone. This study, while still in progress, is obtaining excellent response rates despite having a moderately high intensity protocol. Investigators attribute this to their regular phone contact with participants during which rapport is built and sampling issues can be resolved (Dave Almeida, personal communication, 26 February 2008).

### 5.3. Response rates and explanations for variations in response rates

Response rates are not yet available for all the studies in Table 4, as several are still in progress. Where response rates are reported, they have ranged from an unusually low rate of 25% (Gotthenburg study) to 93% in the Cebu Longitudinal Survey (Adair et al., 2001). The average response rate for the completed (not still in progress) studies in Table 4 is 77%, if all studies are counted equally, or 81% if the average is weighted by sample size.

There are many variables that contribute to high response rates. All of the factors reported to predict non-response in survey studies remain true for saliva collection. Thus, better response rates are seen with the use of generous incentives, frequent and positive interactions, and use of short questionnaires/procedures that participants find interesting and user-friendly (Edwards et al., 2002). The factors most relevant to saliva collection are engaging the participant in the research and making the collection and return of saliva as user-friendly as possible (see Table 5 for some tips on how to accomplish these goals).

Beyond the tips described in Table 5, factors that contribute to improved compliance include a long history of positive interactions with participants, and a clear rationale for the inclusion of the additional measures. For example, the Cebu longitudinal survey actually used many of the *same* interviewers from year to year, resulting in high levels of trust in the interviewers and the study as a whole. In the Whitehall II study, participants were already part of a study they know as the ''stress and health study''. Therefore, the addition of a new biomarker to assess ''stress hormones'' was intuitive. In our experiences across multiple studies, participants have been known to refuse biological protocols due to everything ranging from reasonable concerns about drug testing to more unlikely concerns such as the possibility of cloning. These concerns need to be addressed directly, preferably in conversation with a sympathetic interviewer who they feel they can trust, and who is well prepared to respond to such questions.

### 5.4. Monitoring and maximizing compliance with sampling protocols

Beyond simply obtaining high response rates for sampling, appropriate compliance with sampling protocols is important for obtaining high quality data. Several studies now suggest that when morning samples are inaccurately timed in relation to either wakeup time (Dockray et al., 2008), or in relation to each other (Kudielka and Kirschbaum, 2003), this can have a significant (negative) impact on estimates of the CAR. As a result, smaller-scale research studies now typically employ objective monitors of sample timing compliance, such as MEMS® Track Caps. In very large-scale studies, the cost and practical implications of obtaining objective measures of waking may be prohibitive. Use of objective monitoring devices in a random subsample does not allow identification of who among the full sample are compliant or not, but provides at least some indication of the degree of noncompliance that is occurring in the sample as a whole. The MESA study has employed these devices into their saliva collection protocol. As a consequence, perhaps, saliva samples are not collected in the entire cohort but in 1000 participants randomly selected from the whole cohort.

Most investigators in large-scale studies rely on clearly and strongly emphasizing to participants the importance of accurate timing of the morning samples. Participants should also be told, however, that it is preferable to report their *actual* sampling times, even if deviations from the requested protocol occur. This allows investigators to be aware of, and statistically model the effects of, deviations from ideal sampling protocols (e.g. Adam et al., 2006). Thankfully, much evidence suggests that participants do collect morning samples accurately in relation to objectively determined waketimes (Kraemer et al., 2006; Dockray et al., 2008; DeSantis et al., in press). In addition, evidence suggests that participants who believe their sampling times may be monitored show dramatic improvements in the timing of their samples (Broderick et al., 2004). Monitoring the sample timings of a random subgroup of participants, while telling all participants that there is a chance they are being monitored, is perhaps the most ethical way to take advantage of this inexpensive approach to improving compliance. Table 5 summarizes tips for improving both response and compliance rates in large-scale salivary research.

**Table 5** Methods to optimise compliance with saliva collection protocols.

Objective compliance monitoring methods
1. Gold standard methods provide objective assessment of waking time (Actigraph) and sample collection time (MEMS Track Caps)
2. However, expensive to implement on a large scale

*Implementation choices suitable for population based research*
i. Use with random subsample of population to estimate level of incorrect collection in population
ii. 'Blind' incorporation into protocol advising participants that they may have received recording device

Practical support measures
i. Some evidence to suggest these methods improve compliance
ii. Require time commitment from researchers
iii. Data collection team/interviewers need to be informed about and comfortable with goals of saliva collection for these to be effective

*Implementation choices suitable for population based research*
i. Advice to participant when preparing for the data collection
  1. Engage the participant with the goals of the research
  2. Go through protocol with participant
  3. If possible, practice with participant
  4. Advise placing sample kit beside bed on the night before sample collection with a pen
  5. Call or e-mail the evening before sampling begins to highlight instructions and remind participant to start the next morning
ii. Clear, precise, easily understood instructions and kit
  1. Make collection kit user friendly and population appropriate; for example, large print in older populations
  2. Define 'waking' as precisely as possible
  3. Emphasise the importance of immediate collection of sample upon waking
  4. Emphasise the importance of an accurate record of time of sample collection
  5. Help the participant organise the collection; for example colour coded sampling devices matched with colour coded diary pages
  6 Make it easy for participants to return the kit; for example pre-paid packaging and advice of location of nearest drop point
iii. Provide a line of communication and follow-up for the participant
  1. Provide a free-phone support line
  2. Make reminder calls
  3. Consider in-person retrieval of materials from hard-to-reach participants
iv. Incentive structures should reinforce compliance
  1. Bonus for quick return of completed materials
  2. When expensive compliance monitoring methods are used, pay participants for return of equipment, regardless of study completion
  3. Payment can depend on degree of completion
  4. Incentive experiments would be helpful to determine optimal incentive scheme

## 5.5. Saliva collection methods

While we will not cover saliva sampling techniques in detail, a few points are worth making in relation to epidemiological research. First, given the expense of obtaining saliva samples in large-scale research, investigators may wish to think beyond cortisol and use a saliva sampling approach that would allow the investigator to analyze for the largest number of analytes. Where possible, when multiple analytes may be desired, a passive drool technique, without use of any stimulants for saliva flow, and without use of cotton absorbent material is recommended if a broader range of analytes is of interest, given research that salivary stimulants and cotton-based sampling devices may introduce bias into the measurement of some analytes (Shirtcliff et al., 2001). If only cortisol is to be tested, cotton-based sampling approaches such as Salivettes® are acceptable and commonly employed. Additional (safety, motor skill) considerations come into play when conducting salivary sampling with infants, young children or the very elderly, requiring a modification to sample collection techniques—these are beyond the scope of the current paper. If investigators do wish to test for multiple analytes, a reasonably broad consent, allowing storage of samples for a period of several years and permission to analyze for factors beyond salivary cortisol should be obtained, preferably at the time of initial sampling.

## 5.6. Measurement and modeling of covariates

Although response rates are best maintained by keeping additional measures to a minimum, information on several covariates needs to be collected in order to properly model the cortisol data obtained. As noted earlier, commonly employed covariates in HPA axis research are listed in Table 3; several of these are worthy of emphasis, including waketime on the days of sampling.

## 5.7. Waketime

It is essential that waketimes specific to the days of cortisol testing are assessed. Saliva collection for the assessment of diurnal cortisol profiles should be anchored to waking time,

rather than clock time, as diurnal rhythms are anchored primarily to person-specific sleep—wake schedules rather than dark—light cycles (Wilhelm et al., 2007). Sampling-day specific waketimes should then be used covariates in statistical models (e.g. Adam et al., 2006; Badrick et al., 2008). This necessitates participants recording the time they 'woke'. It is important to define "waking" in order to standardize the assessment across participants. Common definitions include 'as soon as you open your eyes and before your feet touch the ground' or 'when your eyes open and you are ready to get up' (Cohen et al., 2006). In older participants (or new parents), who can experience broken sleep (Bliwise, 1993), the definition can be modified to 'as soon as you are aware of being awake for the day and will not go back to sleep' after which they should be requested to sit up, remain in bed and collect the saliva sample.

### 5.8. Other covariates

Beyond time of waking on the days of testing, smoking (both whether the individual smokes, and number of cigarettes typically smoked) and presence of clinical depressionare among the most essential covariates to assess due to their relatively consistent associations with diurnal cortisol rhythms. Table 3 also lists additional variables that should be assessed on the days of cortisol testing, those that may be assessed in an interview or questionnaire measure gathered proximate to the days of testing, and those that typically serve as exclusion criteria for participation in salivary cortisol studies. Note that many of the variables that we are considering "covariates" in Table 3 could be considered key predictors of interest, depending on the topic of investigation.

## 6. Data analytic approaches and controversies

Approximately 1% of cortisol measures are found to be 3 standard deviations above the mean cortisol value (Whitehall II and Rotterdam Studies). It is unclear what these high cortisol values represent. In large-scale studies, the absolute numbers with these very high values can be substantial; it is not yet resolved whether it is best to remove these values altogether from analyses, or windsorize them in order to reduce their influence on the analysis. Even with such outliers removed, cortisol distributions typically exhibit a strong positive skew; it is typical to employ a natural logarithmic transformation to help normalize the cortisol distribution. This also has the effect of making the association between time of day and cortisol more linear in nature. Other transformations, such as log base 10 or squareroot, have also been used.

### 6.1. Modeling the cortisol awakening response

Most large-scale studies collect samples at waking followed by another sample between 30 and 45 min later. The CAR can be measured as the difference between these two cortisol measures (specifically, the value of the wakeup +30—45 min sample minus the value of the wakeup sample). Another approach uses the ratio of the CAR values to wakeup value (Cohen et al., 2006). Where multiple samples to be obtained within the hour after waking, and area under the curve approach is best employed (using the wakeup value as the baseline such that the CAR measure reflects an increase from the wakeup value) (Pruessner et al., 2003a).

### 6.2. Diurnal cortisol slope

Diurnal cortisol slope is best measured as the rate of decline in cortisol levels across the day, typically across the entire span of time from wakeup to bedtime. Many researchers choose not to include the CAR data points in the slope calculation (e.g. Adam, 2006; Cohen et al., 2006) because of suggestions that the CAR may regulated by different neurobiological mechanisms than the rest of the underlying diurnal cortisol curve (Clow et al., 2004). Using this approach, the slope is anchored on the first, wakeup sample of the day (see Adam et al., 2006; Cohen et al., 2006; Badrick et al., 2008 for examples). Other researchers, however, choose to measure the rate of decline in cortisol from the peak value of the day, which is typically the CAR sample, taken 30—40 min after waking. Which of these approaches is the most meaningful remains to be determined in further research.

Diurnal cortisol slopes have been calculated several different ways in large-scale studies. Where multiple (typically 5 or more) data points are available across the day or across multiple days, a line of best fit may be estimated separately through the data points of each individual using either linear regression or multilevel modeling approaches, and the slope of line is used as the cortisol diurnal slope estimate. The minimal protocol for estimating a diurnal cortisol slope includes two data points—one in the morning (either wakeup or 30—40 min post-awakening) with a slope calculated by subtracting bedtime from wakeup values, and dividing by the number of hours separating these two samples.[4] Resulting coefficients are negative, reflecting the declining slope of cortisol values. When multiple data points are gathered across the day, it is clear that the association between cortisol and time of day is not fully linear in nature. Currently, however, there is little to no information on whether the degree of curvilinearity of the diurnal cortisol rhythm is meaningfully related to either exposures or outcomes of interest.

### 6.3. Total cortisol/AUC

The area under the curve (AUC) can be calculated using data from all samples collected (including the CAR values) or only the samples collected for assessment of slope. AUC measures should be used to compliment other measures of diurnal cortisol secretion, such as slope and CAR, given that AUC measures discard information about diurnal variation. AUC measures do provide unique information, however, as they reflect the average level of cortisol across the day, which is not particularly strongly associated with the wakeup to bedtime diurnal cortisol slope (−0.427 in the Whitehall data, Kumari unpublished data).

[4] Some researchers take a simple difference between the wakeup and bedtime values as an indicator of slope, but this approach does not appropriately adjust for variations in the total time awake, and hence the total number of hours over which cortisol has the opportunity to decline.

### 6.4. Simultaneous modeling of wakeup levels, the CAR, and diurnal decline using multilevel modeling techniques

Although all the parameters above can be calculated using the simple mathematic methods described, increasingly, when sample sizes are large enough, investigators are using multilevel modeling techniques that allow investigators to estimate multiple parameters (e.g. elevation of curve at waking or midday, slope, and CAR) simultaneously, and to predict individual differences in these parameters from individual difference variables of interest as well as covariates. Further explanations and examples of this approach can be found in a growing number of published studies (van Eck et al., 1996; Hruschka et al., 2005; Adam, 2005, 2006). In addition to examining how differences between individuals are related to individual differences in diurnal cortisol rhythms, such approaches can also be used to systematically model how day-to-day variations in experience are related to day-to-day variations in cortisol diurnal rhythms, in studies that have at least 3 days of data available (Adam et al., 2006).

## 7. Some findings from epidemiological cortisol research

The investment that researchers have made thus far in gathering salivary cortisol in population-based research studies is beginning to pay off. Studies are now appearing in the published literature, examining problems and processes that would have been difficult to capture with cortisol research utilizing smaller sample sizes. One area of research that these studies are uniquely suited to examine is association between race/ethnicity and/or socioeconomic position and cortisol levels. In large-scale research, such associations can be examined while controlling for a wide variety of confounding variables (such as health behaviours) that may otherwise spuriously account for such associations. For example, Cohen et al. (2006), using the CARDIA data, found that lower socioeconomic status (measured by education and income), and race/ethnicity (being black) were both associated with higher evening levels of cortisol, resulting in flatter diurnal rhythms. The associations between SES and cortisol were mediated by worse health practices, including smoking, depressive symptoms, weaker social support networks, and feelings of helplessness. Similarly, in Whitehall II, lower occupational status measured by civil service employment grade is associated with 'flatter' profiles in cortisol secretion. In both the Whitehall II (Kumari et al., in preparation) and CARDIA studies (Cohen et al., 2006), no association is seen between measures of social position and the CAR. Interestingly, Whitehall II used a sampling protocol based on the CARDIA protocol and the similarity of these findings supports the importance of 'harmonising' saliva collection protocols when investigating similar issues. When variations in methodologies are observed, so are variable findings (e.g. Brandtstandter et al., 1991; Wright and Steptoe, 2005). To obviate some of these differences in future research we go on to make recommendations that we hope serve to help 'harmonise' saliva collection protocols in large-scale survey settings; these recommendation are found in Table 6.

## 8. Future directions for epidemiological salivary cortisol research

### 8.1. Research guiding design choices

The addition of salivary cortisol protocols to large-scale studies heralds a new and exciting period for HPA axis research, that should lead to a better understanding of the role of daytime cortisol secretion, its behavioural and biological correlates and how it relates to the development of disease. More research is however needed to quantify the extent of unreliability introduced by minimal protocols, and the extent to which that unreliability compromises investigators' abilities to detect effects of interest, despite the additional power provided by large sample sizes. Research that is specifically designed to help guide difficult protocol choices, such as whether it is better to gather fewer samples over multiple days, or more samples over a single day, would be helpful. While evidence indicates that collection of cortisol data on a single day may contribute to lower reliability of cortisol measurement, and may bias results towards state rather than trait measures, most of the current large-scale collections are collected in this way. It remains to be seen whether, and the extent to which, this has compromised investigators' ability to observe prospective associations.

### 8.2. Prospective longitudinal research on change in cortisol diurnal rhythms

Although much research, such as that cited above, implicates stress-related *changes* in cortisol in the emergence of a wide range of disease processes, very few studies, whether large-scale or small, have actually observed changes in cortisol over time in relation to the emergence of disease processes, as this requires more than one wave of cortisol data collection for each individual. Lupien et al. (1998) provides an example of the type of interesting information that could emerge if HPA activity were to be examined across repeated waves of data collection in large-scale surveys. In this study, Lupien examined 24 h cortisol in 54 women annually for 5 years and found groups in which there were progressive increases in level ending with high values, progressive increases ending in moderate levels of cortisol and those with decreasing levels ending with moderate level. Participants who had increasing cortisol levels combined with high initial levels had greater memory impairments and increases in triglycerides (Lupien et al., 1998). That is, chronically high cortisol levels were important rather than high levels per se at the end of the study. Thus, trends in cortisol over a period of months or years may reveal important distinctions between groups that may have similar cortisol patterns at any one wave.

To our knowledge, there are no published epidemiological studies examining within-person change in diurnal cortisol secretion over long periods of time. In a short-term longitudinal approach using the population-based CHASRS data, Adam et al. (2006) modeled within-person change in diurnal cortisol over 3 days, finding that prior day psychosocial experience predicted next day cortisol, and that morning cortisol levels predicted fatigue for the rest of the day. In understanding the role of cortisol in disease outcomes, changes in diurnal cortisol patterns will need to be observed over a period

**Table 6** Design and analysis recommendations for epidemiological salivary cortisol research.

Study design

1. Participant selection strategy needs to be made a priori (for example, from a national database or from all schools in a specific area) in order for findings to be generalizable
2. The number of participants included needs to be sufficient to have power to examine the research question at hand

Sample collection

1. Sample collection protocols need to match the complexities of protocols and degree of participant burden to the characteristics of the population under consideration
2. Thorough and clear instructions to participants are necessary, including detailed descriptions of how to collect samples, where to place sampling materials, and how to return samples
3. Data collection kits should be made as user friendly as possible and age appropriate
4. Saliva collection should be designed at minimum to collect information on the CAR and the slope of diurnal decline in cortisol secretion. The minimal protocol currently employed in population-based research involves: one sample collected on waking, one at the peak of the cortisol awakening response (typically 30 min after waking) and one at bedtime
5. Employment of saliva collection protocols across more than 1 day improves the ability of the study to draw conclusions regarding trait, rather than day-specific state components of diurnal cortisol activity. For assessment of the CAR and diurnal slope, addition of more days of data with a minimal protocol provides more important information than addition of more data points within a single sampling day
6. All times of collection need to be recorded, with an emphasis placed on accurate reporting of sample times, regardless of whether samples were taken at the requested times
7. A minimum set of covariates must contain an assessment of waking time on day of sample collection, and information on health-related confounds known to relate to cortisol
8. Use electronic monitoring devices to measure timing of sample compliance for at least a subsample

Analyses

1. In sufficiently large studies, it can be possible to model the effects of noncompliance with the timing of sampling protocols rather than removing the data of less compliant participants
2. Values found outside of 3 standard deviations above mean cortisol should either be removed from analyses or windsorized to 3 standard deviations above the mean
3. The CAR, diurnal cortisol slope, and area under the curve across the day should each be modeled as separate indicators of diurnal cortisol functioning. Area under the curve measures provide an indication of total cortisol across the day, but current evidence suggests that the diurnal cortisol slope and CAR are more meaningfully related to stress exposure and disease outcomes
4. In rare instances where many samples per day, or multiple days of data are available in large-scale research projects, multilevel modeling techniques should be employed to model individual differences in diurnal cortisol rhythms

of months or years, in relation to changes in stress exposure and changes in disease symptoms and onsets. As the types of epidemiological studies reviewed here begin to gather multiple waves of cortisol data over the course of many years, the role of, and time scales over which changes in diurnal cortisol patterns contribute to the emergence of pre-disease and disease processes will be increasingly illuminated.

## Role of funding source

Whitehall II study has been supported by grants from the Medical Research Council; Economic and Social Research Council; British Heart Foundation; Health and Safety Executive; Department of Health; National Heart Lung and Blood Institute (HL36310), US, NIH: National Institute on Aging (AG13196), US, NIH; Agency for Health Care Policy Research (HS06516); and the John D and Catherine T MacArthur Foundation Research Networks on Successful Midlife Development and Socioeconomic Status and Health. The CHASRS study was supported by National Institute of Aging P01 AG1891 and the John Templeton Foundation. These funding bodies had no further role in the gathering of data, writing of the report or decision to submit the paper for publication. Dr. Adam's time on this manuscript was partially supported by a Faculty Fellowship from the Institute for Policy Research, Northwestern University, and by a Scholar Award from the William T. Grant Foundation. Dr Kumari's time on this manuscript was partially supported by the Economic and Social research council (RES-596-28-0001).

## Conflict of interest

The authors have no conflicts of interests to declare.

## Acknowledgements

We would like to thank Martin Shipley for providing the power calculations. We would like to thank all the principal investigators that provided information on the studies in which salivary samples have been incorporated.

## References

Adair, L.S., Kuzawa, C.W., Borja, J., 2001. Maternal energy stores and diet composition during pregnancy program adolescent blood pressure. Circulation 104, 1034—1039.

Adam, E.K., 2005. Momentary emotion and cortisol levels in the everyday lives of working parents. In: Schneider, B., Waite, L.

(Eds.), Being Together, Working Apart: Dual Career Families and the Work-Life Balance. Cambridge University Press, Cambridge, pp. 105—134.

Adam, E.K., 2006. Transactions among adolescent trait and state emotion and diurnal and momentary cortisol activity in naturalistic settings. Psychoneuroendocrinology 31, 664—679.

Adam, E.K., Hawkley, L.C., Kudielka, B.M., Cacioppo, J.T., 2006. Day-to-day dynamics of experience-cortisol associations in a population-based sample of older adults. Proc. Natl. Acad. Sci. U.S.A. 103, 17058—17063.

Badrick, E., Bobak, M., Kirschbaum, C., Britton, A., Marmot, M., Kumari, M., 2008. Alcohol consumption and cortisol secretion in an ageing cohort. J. Clin. Endocrinol. Metab. 93, 750—757.

Badrick, E., Kirschbaum, C., Kumari, M., 2007. The relationship between smoking status and cortisol secretion. J. Clin. Endocrinol. Metab. 92, 819—824.

Belmaker, R.H., Agam, G., 2008. Major depressive disorder. N. Engl. J. Med. 358, 55—68.

Bliwise, D.L., 1993. Sleep in normal aging and dementia. Sleep 16, 40—81.

Brandtstandter, J., Bakers-Gotz, B., Kirschbaum, C., Hellhammer, D., 1991. Developmental and personality correlates of adrenocortical activity as indexed by salivary cortisol: observations in the age range of 35 to 65 years. J. Psychosom. Res. 35, 173—185.

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.

Brunner, E.J., Hemingway, H., Walker, B.R., Page, M., Clarke, P., Juneja, M., Shipley, M.J., Kumari, M., Andrew, R., Seckl, J.R., Papadopoulos, A., Checkley, S., Rumley, A., Lowe, G.D.O., Stansfeld, S.A., Marmot, M.G., 2002. Adrenocortical, autonomic, and inflammatory causes of the metabolic syndrome. Circulation 106, 2659—2665.

Chida, Y., Steptoe, A., 2009. Cortisol awakening response and psychosocial factors: a systematic review and meta-analysis. Biol. Psychol. 265—278.

Clow, A., Thorn, L., Evans, P., Hucklebridge, F., 2004. The awakening cortisol response: methodological issues and significance. Stress 7, 29—37.

Cohen, S., Schwartz, J.E., Epel, E., Kirschbaum, C., Sidney, S., Seeman, T., 2006. Socioeconomic status, race, and diurnal cortisol decline in the Coronary Artery Risk Development in Young Adults (CARDIA) Study. Psychosom. Med. 68, 41—50.

DeSantis, A., Adam, E.K., Doane, L., Mineka, S., Zinbarg, R., Craske, M., 2007. Racial/ethnic differences in cortisol diurnal rhythms in a community sample of adolescents. J. Adolesc. Health 41, 3—13.

DeSantis, A., Adam, E.K., Mendelsohn, K., Doane, L.D., in press. Concordance between self-reported and actual wake-up times in ambulatory salivary cortisol research: implications for the cortisol response to awakening. Int. J. Behav. Med.

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, P., Roberts, I., Clarke, M., DiGuiseppi, C., Pratap, S., Wentz, R., Kwan, I., 2002. Increasing response rates to postal questionnaires: systematic review. Brit. Med. J. 324, 1—9.

Eller, N.H., Netterstrom, B., Allerup, P., 2005. Progression in intima media thickness-the significance of hormonal biomarkers of chronic stress. Psychoneuroendocrinology 30, 715—723.

Finch, C.E., Vaupel J., Kinsella K., 2001. Cells and Surveys: Should Biological Measures Be Included in Social Science Research? Commission on Behavioral and Social Sciences and Education National Academic Press. http://www.nap.edu/books/0309071992/html/.

Harville, E.W., Savitz, D.A., Dole, N., Herring, A.H., Thorp, J.M., Light, K.C., 2007. Patterns of salivary cortisol secretion in pregnancy and implications for assessment protocols. Biol. Psychol. 74, 85—91.

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.

Herbert, J., Goodyer, I.M., Grossman, A.B., Hastings, M.H., de Kloet, E.R., Lightman, S.L., Lupien, S.J., Roozendaal, B., Seckl, J.R., 2006. Do corticosteroids damage the brain? J. Neuroendocrinol. 18, 393—411.

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.

Kajantie, E., Phillips, D.I.W., Andersson, S., Barker, D.J.P., Dunkel, L., Forsen, T., Osmond, C., Tuominen, J., Wood, P.J., Eriksson, J., 2002. Size at birth, gestational age and cortisol secretion in adult life: foetal programming of both hyper- and hypocortisolism? Clin. Endocrinol. 57, 635—641.

Kirschbaum, C., Hellhammer, D., 1989. Salivary cortisol in psychobiological research: an overview. Neuropsychobiology 22, 150—169.

Kraemer, H.C., Giese-Davis, J., Yutsis, M., O'Hara, R., Neri, E., Gallagher-Thompson, D., Taylor, C.B., Spiegel, D., 2006. Design decisions to optimize reliability of daytime cortisol slopes in an older population. Am. J. Geriatr. Psychiatry 14, 325—333.

Kudielka, B.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., Chandola, T., Adler, N., Epel, E., Seeman, T., Kirschbaum, C., Marmot, M.G., in preparation. Measures of social position and cortisol secretion in an ageing population. Findings from the Whitehall II study.

Lupien, S.J., de leon, M., De Santi, S., Convit, A., Tarshish, C., Nair, N.P.V., McEwen, B.S., Hauger, R.L., Meaney, M.J., 1998. Cortisol levels during human aging predict hippocampal atrophy and memory deficits. Nat. Neurosci. 1, 69—73.

Matthews, K., Schwartz, J., Cohen, S., Seeman, T.E., 2006. Diurnal cortisol decline is related to coronary calcification: CARDIA study. Psychosom. Med. 68, 657—661.

McEwen, B.S., 2007. Physiology and neurobiology of stress and adaptation: central role of the brain. Physiol. Rev. 87, 873—904.

Nater, U.M., Maloney, E., Boneva, R.S., Gurbaxani, B.M., Lin, J.-M., Jones, J.F., Reeves, W.C., Heim, C., 2007. Attenuated morning salivary cortisol concentrations in a population-based study of persons with chronic fatigue syndrome and well controls. J. Clin. Endocrinol. Metab. 93, 703—709.

Phillips, D.I.W., Barker, D.J.P., Fall, C.H.D., Seckl, J.R., Whorwood, C.B., Wood, P.J., Walker, B.R., 1998. Elevated plasma cortisol concentrations: a link between low birth weight and insulin resistance syndrome? J. Clin. Endocrinol. Metab. 83, 757—760.

Pruessner, J.C., Kirschbaum, C., Meinlschmid, G., Hellhammer, D.H., 2003a. 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., Buske-Kirschbaum, A., von Auer, K., Jobst, S., et al., 1997. Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. Life Sci. 61, 2539—2549.

Pruessner, M., Hellhammer, D.H., Pruessner, J.C., Lupien, S.J., 2003b. Self-reported depressive symptoms and stress levels in healthy young men: associations with the cortisol response to awakening. Psychosom. Med. 65, 92—99.

Ranjit, N., Young, E.A., Raghunathan, T.E., Kaplan, G.A., 2005. Modeling cortisol rhythms in a population-based study. Psychoneuroendocrinology 30, 615—624.

Rosmalen, J.G.M., Oldehinkel, A.J., Ormel, J., de Winter, A.F., Buitelaar, J.K., Verhulst, F.C., 2005. Determinants of salivary cortisol levels in 10—12 year old children: a population-based study of individual differences. Psychoneuroendocrinology 30, 483—495.

Sapolsky, R.M., Romero, L.M., Munck, A.U., 2000. How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative action. Endocr. Rev. 21, 55—89.

Saxbe, D.E., 2008. A field (researcher's) guide to cortisol: tracking HPA axis functioning in everyday life. Health Psychol. Rev. 2, 163—190.

Seeman, T.E., McEwen, B.S., Singer, B.H., Albert, M.S., Rowe, J.W., 1997. Increase in urinary cortisol excretion and memory declines: MacArthur studies of successful aging. J. Clin. Endocrinol. Metab. 82, 2458—2465.

Seeman, T.E., Singer, B.H., Ryff, C.D., Dienberg Love, G., Levy-Storms, L., 2002. Social relationships, gender, and allostatic load across two age cohorts. Psychosom. Med. 64, 395—406.

Sephton, S.E.S.R.M., Kraemer, H.C., Spiegel, D., 2000. Diurnal cortisol rhythm as a predictor of breast cancer survival. J. Natl. Cancer Inst. 92, 994—1000.

Shirtcliff, E.A., Granger, D.A., Schwartz, E., Curran, M.J., 2001. Use of salivary biomarkers in biobehavioral research: cotton-based sample collection methods can interfere with salivary immunoassay results. Psychoneuroendocrinology 26, 165—173.

Smith, G.D., Ben-Shlomo, Y., Beswick, A., Yarnell, J., Lightman, S., Elwood, P., 2005. Cortisol, testosterone, and coronary heart disease: prospective evidence from the Caerphilly study. Circulation 112, 332—340.

Steptoe, A., O'Donnell, K., Badrick, E., Kumari, M., Marmot, M., 2008. Neuroendocrine and inflammatory factors associated with positive affect in healthy men and women: the Whitehall II study. Am. J. Epidemiol. 167, 96—102.

Steptoe, A., Siegrist, J., Kirschbaum, C., Marmot, M., 2004. Effort-reward imbalance, overcommitment, and measures of cortisol and blood pressure over the working day. Psychosom. Med. 66, 323—329.

Stone, A.A., Schwartz, J.E., Smyth, J., Kirschbaum, C., Cohen, S., Hellhammer, D., Grossman, S., 2001. Individual differences in the diurnal cycle of salivary free cortisol: a replication of flattened cycles for some individuals. Psychoneuroendocrinology 26, 295—306.

van Eck, M., Berkhof, H., Nicolson, N., Sulon, J., 1996. The effects of perceived stress, traits, mood states, and stressful daily events on salivary cortisol. Psychosom. Med. 58, 447—458.

Weitzman, E.D., Fukushima, D., Nogeire, C., Roffwarg, H., Gallagher, T.F., Hellman, L., 1971. Twenty-four hour pattern of the episodic secretion of cortisol in normal subjects. J. Clin. Endocrinol. Metab. 33, 14—22.

Wilhelm, I., Born, J., Kudielka, B.M., Wust, S., 2007. Is the cortisol awakening rise a response to awakening? Psychoneuroendocrinology 32, 358—366.

Wright, C.E., Steptoe, A., 2005. Subjective socioeconomic position, gender and cortisol responses to waking in an elderly population. Psychoneuroendocrinology 30, 582—590.

## Further reading

Adam, E.K., Gunnar, M.R., 2001. Relationship functioning and home and work demands predict individual differences in diurnal cortisol patterns in women. Psychoneuroendocrinology 26, 189—208.

Almeida, D.M., Savla, J., Stawski, R.S., Banks, S.R., 2007. Into the field and under the skin: measurement quality of salivary cortisol in the National Study of Daily Experiences. In: International Society of Psychoneuroendocrinology Annual Conference, Madison, WI.

Essex, M.J., Klein, M.H., Cho, E., Kalin, N.H., 2002. Maternal stress beginning in infancy may sensitize children to later stress exposure: effects on cortisol and behavior. Biol. Psychiatry 52, 776—784.

Follenius, M., Brandenberger, G., Hietter, B., Simeoni, M., Reinhardt, B., 1982. Diurnal cortisol peaks and their relationships to meals. J. Clin. Endocrinol. Metab. 55, 757—761.

Gillespie, C.F., Nemeroff, C.B., 2005. Hypercortisolemia and depression. Psychosom. Med. 67 (Suppl. 1), S26—S28.

Harris, T.B., Launer, L.J., Eiriksdottir, G., Kjartansson, O., Jonsson, P.V., Sigurdsson, G., Thorgeirsson, G., Aspelund, T., Garcia, M.E., Cotch, M.F., Hoffman, H.J., Gudnason, V., 2007. Age, gene/environment susceptibility-Reykjavik Study: multidisciplinary applied phenomics. Am. J. Epidemiol. 165, 1076—1087.

Hibel, L.C., Granger, D.A., Cicchetti, D., Rogosch, F., 2007. Salivary biomarker levels and diurnal variation: associations with medications prescribed to control children's problem behavior. Child Dev. 78, 927—937.

Ice, G.H., 2005. Factors influencing cortisol level and slope among community dwelling older adults in Minnesota. J. Cross Cult. Gerontol. 20, 91—108.

Kirschbaum, C., Kudielka, B., Gaab, J., Schommer, N.C., Helhammer, D.H., 1999. Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus—pituitary—adrenal axis. Psychosom. Med. 61, 154—162.

Kertes, D.A., Gunnar, M.R., 2004. Evening activities as potential confound in research on the adrenocortical system in children. Child Dev. 75, 193—204.

Kumari, M., Badrick, E., Chandola, T., Adam, E.K., Stafford, M., Marmot, M., Kirschbaum, C., Kivimaki, M., 2009. Cortisol secretion and fatigue: associations in a community based cohort. Psychoneuroendocrinology, doi:10.1016/j.psyneuen.2009.05.001.

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.

Lovallo, W.R., Whitsett, T.L., Al'Absi, M., Sung, B.H., Vincent, A.S., 2005. Caffeine stimulations of cortisol secretion across the waking hours in relations to caffeine intake levels. Psychosom. Med. 67, 734—749.

Lupien, S.J., King, S., Meaney, M.J., McEwen, B.S., 2001. Can poverty get under your skin? Basal cortisol levels and cognitive function in children from low and high socioeconomic status. Dev. Psychopathol. 13, 653—676.

Masharani, U., Shiboski, S., Eisner, M.D., Katz, P.P., Janson, S.L., Grainger, 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.

Nicolson, N., 2004. Childhood parental loss and cortisol levels in adult men. Psychoneuroendocrinology 29, 1012—1018.

Obel, C., Hedegaard, M., Henriksen, T.B., Secher, N.J., Olsen, J., Levine, S., 2005. Stress and salivary cortisol during pregnancy. Psychoneuroendocrinology 30, 647—656.

Osika, W., Friberg, P., Wahborg, P., 2007. A new self rating questionnaire to assess stress in children. Int. J. Behav. Med. 14, 1—9.

Pendry, P., Adam, E.K., 2007. Associations between parents' marital functioning, maternal parenting quality, maternal emotion and child cortisol levels. Int. J. Behav. Dev. 31, 218—231.

Penninx, B.W., Beekman, A.T., Smit, J.H., Zitman, F.G., Nolen, W.A., Spinhoven, P., Cuijpers, P., De Jong, P.J., Van Marwijk, H.W., Assendelft, W.J., Van Der Meer, K., Verhaak, P., Wensing, M., De Graaf, R., Hoogendijk, W.J., Ormel, J., Van Dyck, R., NESDA Research Consortium, 2008. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. Int J Methods Psychiatr Res. 17, 121—140.

Powell, L.H., Lovallo, W.R., Matthews, K.A., Meyer, P., Midgley, A.R., Baum, A., Stone, A.A., Underwood, L., McCann, J.J., Herro, K.J., Ory, M.G., 2002. Physiologic markers of chronic stress in premenopausal, middle-aged women. Psychosom. Med. 64, 502—. 509.

Power, C., Li, L., Hertzman, C., 2006. Associations of early growth and adult adiposity with patterns of salivary cortisol in adulthood. J. Clin. Endocrinol. Metab. 91, 4264—4270.

Raff, H., Raff, J.L., Findling, J.W., 1998. Late-night salivary cortisol as a screening test for Cushing's syndrome. J. Clin. Endocrinol. Metab. 83, 2681—2686.

Rohleder, N., Beulen, S.E., Chen, E., Wolf, J.M., Kirschbaum, C., 2007. Stress on the dance floor: the cortisol stress response to social-evaluative threat in competitive ballroom dancers. Personal. Soc. Psychol. 33, 69—84.

Sastry, N., Ghosh-Dastidar, B., Adams, J., Pebley, A.R., 2006. The design of a multilevel survey of children, families, and communities: The Los Angeles Family and Neighborhood Study. Soc. Sci. Res. 35, 1000—1024.

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.

Schlotz, W., Schulz, P., Hellhammer, J., Stone, A.A., Hellhammer, D.H., 2006. Trait anxiety moderates the impact of performance pressure on salivary cortisol in everyday life. Psychoneuroendocrinology 31, 459—472.

Smyth, J.M., Ockenfels, M.C., Porter, L., Kirschbaum, C., Hellhammer, D.H., Stone, A.A., 1998. Stressors and mood measured on a momentary basis are associated with salivary cortisol secretion. Psychoneuroendocrinology 23, 353—370.

Therrien, F., Drapeau, V., 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.

Varadhan, R., Walston, J., Cappola, A.R., Carlson, M.C., Wand, G.S., Fried, L.P., 2008. Higher levels and blunted diurnal variation of cortisol in frail older women. J. Gerontol. A: Biol. Sci. Med. Sci. 63, 190—195.

Vreeburg, S.A., Kruijtzer, B.P., van Pelt, J., van Dyck, R., DeRijk, R.H., Hoogendijk, W.J., Smit, J.H., Zitman, F.G., Penninx, B.W., 2009. Associations between sociodemographic, sampling and health factors and various salivary cortisol indicators in a large sample without psychopathology. Psychoneuroendocrinology 34, 1109—1120.