ann. behav. med. (2016) 50:210–236
DOI 10.1007/s12160-015-9753-9

ORIGINAL ARTICLE

Use of Salivary Diurnal Cortisol as an Outcome Measure in Randomised Controlled Trials: a Systematic Review

Richella Ryan, Dr.[1,2] · Sara Booth, Dr.[1,2] · Anna Spathis, Dr.[1] · Sarah Mollart, Dr.[3] · Angela Clow, Prof.[4]

Published online: 23 March 2016

Abstract

Background Dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis is associated with diverse adverse health outcomes, making it an important therapeutic target. Measurement of the diurnal rhythm of cortisol secretion provides a window into this system. At present, no guidelines exist for the optimal use of this biomarker within randomised controlled trials (RCTs).

Purpose The aim of this study is to describe the ways in which salivary diurnal cortisol has been measured within RCTs of health or behavioural interventions in adults.

Methods Six electronic databases (up to May 21, 2015) were systematically searched for RCTs which used salivary diurnal cortisol as an outcome measure to evaluate health or behavioural interventions in adults. A narrative synthesis was undertaken of the findings in relation to salivary cortisol methodology and outcomes.

Results From 78 studies that fulfilled the inclusion criteria, 30 included healthy participants (38.5 %), 27 included patients with physical disease (34.6 %) and 21 included patients with psychiatric disease (26.9 %). Psychological therapies were most commonly evaluated (n=33, 42.3 %). There was substantial heterogeneity across studies in relation to saliva collection protocols and reported cortisol parameters. Only 39 studies (50 %) calculated a rhythm parameter such as the diurnal slope or the cortisol awakening response (CAR). Patterns of change in cortisol parameters were inconsistent both within and across studies and there was low agreement with clinical findings.

Conclusions Salivary diurnal cortisol is measured inconsistently across RCTs, which is limiting the interpretation of findings within and across studies. This indicates a need for more validation work, along with consensus guidelines.

Keywords Salivary cortisol · Diurnal rhythm · Randomized controlled trial · Systematic review · Intervention

Introduction

The hypothalamic-pituitary-adrenal (HPA) axis is known to be an important pathway in the regulation of the physiological stress response. HPA axis dysregulation has been shown to be

Electronic supplementary material The online version of this article (doi:10.1007/s12160-015-9753-9) contains supplementary material, which is available to authorized users.

✉ Richella Ryan
richella.ryan@nhs.net

Sara Booth
sb628@cam.ac.uk

Anna Spathis
anna.spathis@addenbrookes.nhs.uk

Sarah Mollart
sarahmollart@nhs.net

Angela Clow
clowa@westminster.ac.uk

[1] Palliative Care Department, Cambridge University Hospitals NHS Foundation Trust, Elsworth House, Box 63, Hill's Road, Cambridge CB2 0QQ, UK

[2] Department of Oncology, University of Cambridge, Hutchison/MRC Research Centre, Cambridge Biomedical Campus, Box 197, Cambridge CB2 0XZ, UK

[3] St. Nicholas Hospice Care, Hardwick Lane, Bury St. Edmunds, Suffolk IP33 2QY, UK

[4] Department of Psychology, University of Westminster, 101 New Cavendish Street, London W1W 6XH, UK

associated with important health outcomes including psychiatric illness [1], cardiovascular mortality [2], cancer prognosis [3, 4], and frailty and cognitive decline [5]. These associations are thought to be mediated by the deleterious effects of chronic stress on HPA axis function [6], with secondary effects on metabolic, immune and psychobiological systems [7]. The many associations between HPA axis dysregulation and markers of health status suggest that HPA axis modulation by therapeutic interventions may have a role in disease treatment and prevention. In order to demonstrate this, accurate and feasible measurement of HPA axis function within randomised controlled trials (RCTs) is necessary.

The use of salivary cortisol as a biomarker of stress and HPA axis function is a well-established practice in stress research, dating back to at least 20 years [8]. Due to marked diurnal variation in cortisol hormone secretion throughout the day [9], a variety of methods of salivary cortisol collection and analysis have been explored and utilised in an attempt to identify the most representative summary measure of HPA axis function. Essentially, two broad approaches have been taken [9]. The first approach is to measure HPA axis reactivity to a standardised acute stressor. Whilst this approach is useful, interpretation of the results is limited by the need to consider the time of day the stressor is administered, as well as the nature of the stressor. The second approach is to measure basal or unstimulated HPA axis function; thus avoiding the need to administer a stressor.

Measurement of basal HPA axis function has evolved considerably over the past two decades, as theoretical and empirical knowledge regarding the stress system, cortisol measurement and disease associations has increased. In the early days of salivary cortisol research, a single salivary cortisol measure, collected at a pre-specified time, was used to estimate basal HPA axis function, but this methodology proved to be unreliable, with large intra-individual and inter-individual variation [10]. Another common approach was to measure average or total cortisol exposure over a 12–24-h period [11]. Whilst this approach provides a summary measure, it does not accommodate the complex nature of HPA axis aberration, with both hypocortisolism and hypercortisolism now recognised to be linked to chronic stress [12]. Increasingly, therefore, there has been a move towards measuring the circadian rhythm of diurnal cortisol secretion, rather than focusing on absolute cortisol concentration [11].

Typically, under basal conditions, a healthy HPA axis is characterised by a distinctive circadian pattern of cortisol secretion, whereby cortisol rises to a peak within 30–45 min of waking and then falls to a nadir during sleep at approximately midnight [9, 11]. The major measurable parameters of this diurnal rhythm are (1) the cortisol awakening response (CAR), which is the rise in cortisol during the first 30–45 min following awakening [9], and (2) the diurnal cortisol slope, which is the rate of decline in cortisol levels across the day, from morning to evening [11]. This normal rhythm becomes disrupted when the HPA axis becomes dysregulated [9, 13], the pattern of disruption varying depending on the context or condition studied. In general, an abnormal cortisol awakening response, both abnormally large and small, or a flattened diurnal cortisol slope appear to be consistent markers of HPA axis dysfunction [11]. Importantly, there is evidence that these parameters are independently regulated, with the cortisol awakening response being mediated by an extra-pituitary pathway to the adrenal from the suprachiasmatic nucleus [14]. Thus, these parameters are believed to represent different aspects of HPA axis function [14–16].

Due to the many possible ways of measuring salivary cortisol as a biomarker in stress research, it is necessary to reach consensus regarding the most appropriate methodology, so that the results of different research studies can be compared and so as to avoid waste in the design, conduct and reporting of studies. Adam and Kumari [11] have reviewed the use of salivary diurnal cortisol in epidemiological studies and have, accordingly, published recommendations. They found that the cortisol awakening response, the cortisol slope and the area under the daytime cortisol curve (AUC) were most commonly measured within large epidemiological studies and have been most robustly linked with psychosocial phenomena and health outcomes, implying clinical relevance. They recommend that these parameters should each be assessed as separate indicators of HPA axis function and that the cortisol collection schedule be sufficient to measure, at minimum, the cortisol awakening response and the diurnal slope over more than 1 day.

There is no guidance available for the measurement of salivary diurnal cortisol within interventional studies, and little is known about how salivary cortisol has been employed, to date, as a biomarker within RCTs of health and behavioural interventions. The inherent complexity of salivary diurnal cortisol as a biomarker is likely to pose particular challenges within RCTs. Given that the diurnal profile is essentially a composite of at least three measurement parameters (the cortisol awakening response, diurnal slope and area under the curve), each reflecting different aspects of HPA axis function, it is possible that experimental interventions will have different effects on different parameters. This is likely to impact on a priori decisions about the primary measurement parameter of interest, on hypotheses about directions of change and on conclusions about efficacy, target engagement and mechanisms of action. As well as these challenges, there are concerns in the literature about the long-term stability of this biomarker over periods of greater than 1 month [17], as well as concerns about its reliability in the shorter term due to day-to-day state effects [18] and the effects of non-compliance [19]. Concerns have also been raised about the responsiveness of the biomarker and how different contexts and populations may impact on this [20].

To assess whether specific guidance is necessary, we systematically reviewed the literature with the aim of describing the RCTs of health and behavioural interventions which have used salivary diurnal cortisol as an outcome measure, particularly focusing on salivary diurnal cortisol methodology and findings. Specifically, we aimed to explore the following questions:

1. Which health and behavioural interventions have been evaluated using salivary diurnal cortisol?
2. What populations have been evaluated?
3. What collection protocols have been used to obtain a diurnal cortisol profile?
4. What parameters of the diurnal cortisol profile have been measured?
5. Where a change in a cortisol profile parameter is observed, when in the follow-up period does it occur?
6. How often is there consistency between the clinical and cortisol response to the intervention?

Methods

The protocol for this review is available in the Electronic Supplementary Material 1.

Study Inclusion and Exclusion Criteria

We restricted the sample to RCTs only. Though the review question is relevant to non-randomised uncontrolled longitudinal studies also, we chose to select RCTs only in order to reduce the scope of the search and the heterogeneity of the sample; this was deemed necessary given the broad review question with respect to interventions and sample populations. We also expected that RCTs would be of higher quality than other study designs, thus providing more reliable information. We defined 'diurnal cortisol profile measurement' as the collection of at least two samples of salivary cortisol over at least 1 day. This would enable the calculation of the diurnal slope or the cortisol awakening response, at minimum. Within the context of an RCT, there needed to be evidence that this measurement had been obtained on at least one occasion before the intervention and on at least one separate occasion (i.e. a separate day or period of days) after the intervention. Using these definitions, we adhered to the following inclusion and exclusion criteria:

Inclusion Criteria

1. Population: any adult (>18 years) population.
2. Study design: randomised controlled trials.
3. Interventions: any type of therapeutic intervention designed to improve an aspect of health or well-being, excluding exogenous corticosteroids.
4. Control or comparator: any type of control or comparator.
5. Outcome measures: studies that use salivary diurnal cortisol profile measurement as a primary or secondary outcome measure.

Exclusion Criteria

1. Non-RCT studies, including quasi-randomised controlled trials and trial protocol reports without results
2. Studies which use non-diurnal salivary sampling methods e.g. single salivary cortisol measures pre and post an intervention or salivary cortisol pre and post a stress-task
3. Studies which evaluate the effects of exogenous glucocorticoids (any type or route) on salivary cortisol
4. Studies which evaluate the cortisol response to stress-inducing interventions or conditions
5. Studies which measure diurnal cortisol under laboratory-induced conditions (e.g. light-wake conditions)
6. Studies in which the diurnal profile is not measured both before and after the intervention
7. Studies in which the diurnal profile is obtained on the same day as the intervention with a view to assessing its acute effects within the day
8. Studies in people with Cushing's disease
9. Animal studies
10. Abstract publication available only
11. Dissertation or non-journal publication available only
12. Non-English language publications

Search Methods for Identification of Studies

On 21 May 2015, we searched the following electronic databases using NHS Evidence Healthcare Databases Advanced Search tool: MEDLINE (1980 to May 2015), CINAHL (1980 to May 2015), PsychINFO (1806 to May 2015), AMED (1985 to May 2015), EMBASE (1974 to May 2015) and the Cochrane Central Register of Controlled Trials. We sought to identify a combination of keywords and MESH terms in the titles and abstracts of papers, adapting the search strategy, as appropriate, for each database. By way of example, the following keywords, MESH terms and publication types were searched in MEDLINE: [('cortisol' AND 'saliva*') OR (HYDROCORTISONE/ AND SALIVA/)] AND [('randomized controlled trial' OR 'controlled clinical trial' OR 'randomized' OR 'placebo' OR 'randomly' or 'trial' OR 'groups', excluding ANIMALS (exploded term)]. See Appendix A for

more detailed search strategies, including strategies used in the other databases.

Study Selection for Inclusion in the Review

All abstracts generated from the electronic searches were exported to Endnote X3 for removal of duplicates. One review author (RR) screened the titles and abstracts for the eligibility criteria, and where eligibility could not be determined, the full-text article was obtained. All full-text articles were reviewed for eligibility by RR, and a random selection of these articles (37 %) were reviewed independently by the other four authors (SB, AS, AC and SM), each reviewing a different selection, to ensure that the eligibility criteria were being correctly interpreted and adhered to. The inclusion criteria were applied in a hierarchical manner, first checking the population, then the study design, then the intervention and finally the cortisol methodology. Any disagreement was discussed in the first instance between the two authors in question. If consensus was not achieved between the two authors in question, a third party (one of the other authors) was consulted.

Data Collection and Extraction

All full-text articles which were deemed eligible were reviewed in more detail for data extraction. RR completed the data extraction for all eligible texts. In addition, another author independently completed data extraction on 10 % of the eligible articles to ensure that data were being extracted appropriately. A data extraction form containing the following fields was used to summarise the pertinent details of the study: trial ID, eligibility criteria checklist with decision outcome, study design, population, intervention/control, salivary cortisol collection protocol details, other outcome measures, salivary cortisol analysis details, salivary cortisol results and clinical outcome results.

Assessment of Quality and Relevance

The quality of each study and its relevance to the review aim were assessed using the Gough Weight of Evidence framework [21], which uses four domains of assessment (A, B, C and D), rating each domain as low, moderate or high. As the overarching aim of this review was to describe salivary cortisol methodology and findings in RCTs, the relevance of each study was assessed purely in relation to the degree to which it contributed information towards this aim.

Within the first domain of this framework, judgments are made in relation to the generic quality of execution of the study independent of the review aim (Weight of Evidence A). Within the second and third domains, judgements are made in relation to the specific aim of the review, including the appropriateness of the study design to the review aim (Weight of Evidence B) and the focus of the study, including its objectives and reporting, relative to the review aim (Weight of Evidence C). An overall judgement of quality and relevance (low, moderate or high) is formed by combining the assessments for these domains (Weight of Evidence D). This framework is weighted more heavily towards relevance than quality and was chosen as a means of highlighting those studies which provided the most relevant information towards the review aim. This was deemed to be the most appropriate approach to appraising the literature included in this review, given that we were not concerned with evaluating the efficacy of specific interventions.

Data Synthesis and Presentation

The selection process is presented using a PRISMA flow chart. A narrative synthesis of the scope, characteristics and findings of the selected studies is given and presented in tables. To enhance clarity and facilitate comparison, individual studies were organised into four categories according to the intervention being evaluated, and a separate table of studies was created for each category. Counts and percentages were used to describe data patterns across all studies and between study intervention categories. Medians and interquartile ranges were calculated to describe key features of the salivary collection protocol across studies and to describe the frequency and distribution of the follow-up time-points at which cortisol findings occurred. For a cortisol parameter, a significant finding was considered to be present if a study reported a statistically significant within- or between-group effect from baseline to follow-up, for either the intervention or the comparator group. If a significant finding was also reported for at least one clinical outcome measure in the same study, using the same statistical tests, the cortisol findings were considered to support the clinical outcome. In the same way, if there was no evidence of a statistically significant effect for both the cortisol and clinical outcomes, cortisol findings were considered to support clinical findings.

Results

Selection Process

The process of screening and reviewing articles for eligibility is summarised in Fig. 1. The database search identified 2374 articles. After removal of duplicates using Endnote X3, 1812 potentially relevant abstracts were identified. Screening of these abstracts led to the selection of 219 full-text articles for more detailed eligibility assessment. Of these full-text articles, 78 studies were selected for inclusion in the review after removal of ineligible studies and after exclusion of duplicate reports of the same study. The most common reason for

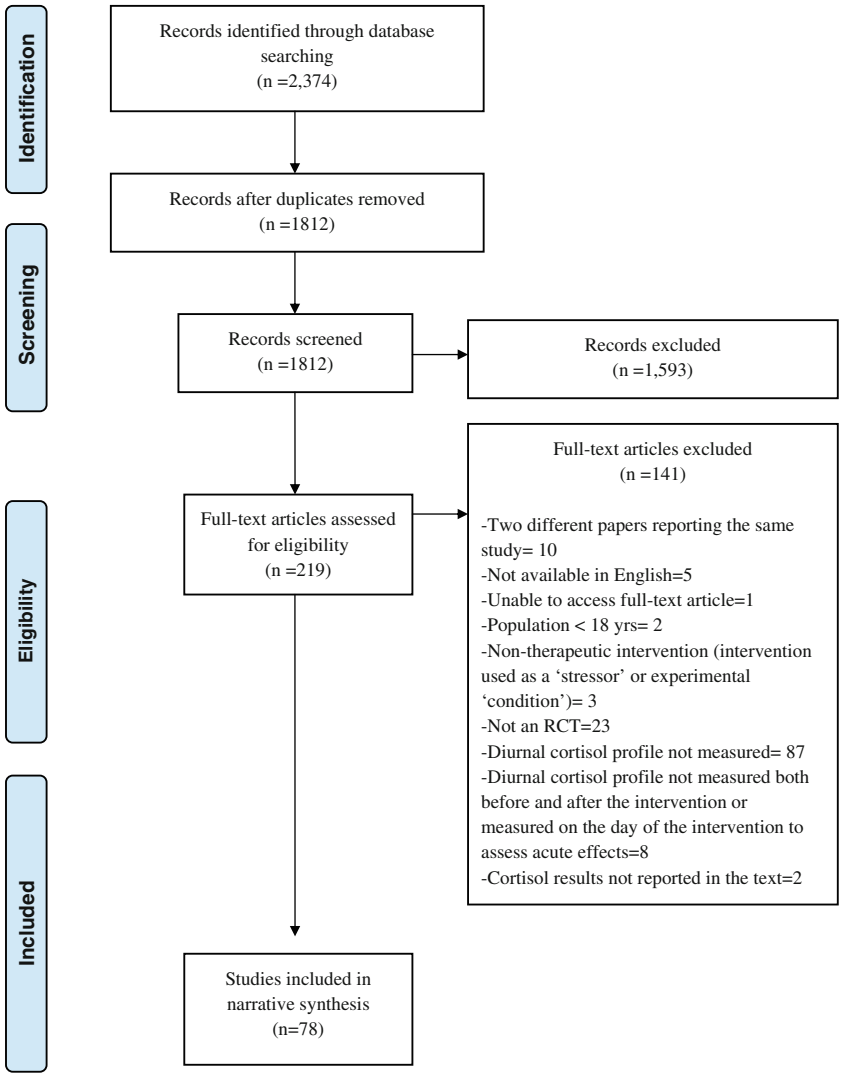

Fig. 1 PRISMA flow diagram illustrating the identification of studies

exclusion of articles related to salivary cortisol methodology. After excluding studies due to ineligible populations, interventions and designs, 87 of the remaining 175 RCTs (50 %) were excluded because the salivary cortisol measurements therein did not allow analysis of the diurnal rhythm. In most cases, this was due to the measurement of cortisol on a single occasion before and after the intervention being evaluated. A small number of RCTs (n=10), which did include diurnal cortisol measurements, were later excluded from the review because their cortisol findings were not adequately reported or because their measurements were conducted in a way that was not comparable with the other studies.

Characteristics of Included Studies

Included studies were published from 2003 to May 2015. There has been a notable increase in the number of published RCTs using salivary diurnal cortisol as an outcome measure in the past decade, with the yearly rate increasing from 2 studies per year in 2003 and 2004 to 14 studies in 2013 and 11 studies in 2014 (see Fig. 2). Indeed, over 50 % of the included studies have been published since 2012. Pertinent characteristics of individual studies are presented within Tables 1, 2, 3 and 4, with each table representing one of four study categories and each study being organised into one of such categories according to the intervention being evaluated: (1) RCTs evaluating psychosocial interventions, (2) RCTs evaluating pharmacological (including nutritional) interventions, (3) RCTs evaluating complementary therapies and (4) RCTs evaluating all other types of interventions.

Most commonly, studies evaluated psychosocial interventions (n=33, 42.3 %), such as cognitive behavioural therapy, mindfulness and psychotherapy (see Table 1). Pharmacological therapies, including nutritional therapies, comprised the second most common intervention category (n=22, 28.2 %), within which eight studies evaluated anti-depressant medications (see Table 2). Complementary therapies were evaluated in 14 studies (17.9 %), with six of these studies evaluating

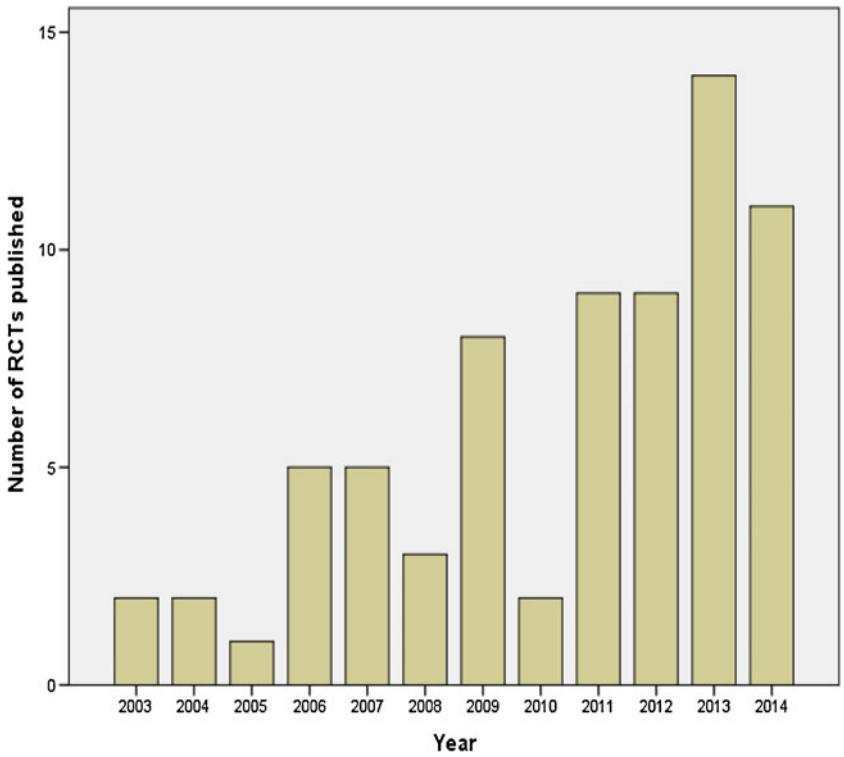

Fig. 2 Number of eligible RCTs published per year. The year 2015 was excluded from this graph as the complete results for this year are not yet available

yoga (see Table 3). Nine studies evaluated treatments which did not fall into the three major categories and were, therefore, classified as 'other' (see Table 4). These treatments included exercise (n=3), cranial electrostimulation (n=2), a lifestyle intervention (n=1), dietary restriction (n=1), prayer (n=1) and light treatment (n=1).

Interventions were evaluated in a wide variety of study populations. Overall, across all studies, these populations could be broadly classified as follows: people who were healthy or at risk of disease (n=30, 38.5 %), patients with physical or psychosomatic disease (n=27, 34.6 %) and patients with a psychiatric diagnosis (n=21, 26.9 %). Physical and psychosomatic disease categories included current or prior cancer (n=14; predominantly breast cancer), cardiovascular disease or metabolic syndrome (n=3), human immunodeficiency virus (n=3), dementia (n=3), Parkinson's disease (n=1), irritable bowel syndrome (n=1), tension headache (n=1) and fibromyalgia (n=1). Psychiatric pathology included depression (n=12), anxiety disorder (n=2), comorbid anxiety and depression (n=1), post-traumatic stress disorder (n=3), adjustment disorder (n=1), bipolar disorder (n=1) and alcoholism (n=1). Within study categories, psychosocial intervention studies and pharmacological studies most commonly evaluated people who were healthy or at risk of disease (42 and 54.5 % of psychosocial and pharmacological studies, respectively), whereas complementary therapy studies most commonly evaluated patients with physical or psychosomatic disease (86 % of complementary therapy studies).

Overall, across study categories, the median study sample size was 55 participants (IQR 34–77), with the smallest study including only 12 participants [76] and the largest including 379 participants [57]. Median sample size was similar between study categories: 59 (IQR 34–74) for psychosocial interventions, 51 (IQR 41–76) for pharmacological interventions, 49 (IQR 21–90) for complementary therapies and 59 (IQR 24–87) for other interventions. Eight studies were reported as pilot, feasibility or exploratory studies. Overall, the median length of follow-up from baseline was 10 weeks (IQR 4 to 21) and ranged from 1 to 72 weeks.

Quality and relevance (aggregate score) were rated as high for 20 studies and moderate for 58 studies. No study was given an aggregate score of low, reflecting the exclusion of studies of low relevance by the eligibility criteria. Of note, many studies were of low quality with respect to their RCT design but of high relevance with respect to the review aim, resulting in a high overall aggregate score using Gough's framework [21]. See Appendix B for the breakdown of scores per domain for each included study.

Salivary Cortisol Collection and Analysis Methodology

Pertinent details relating to the salivary cortisol collection protocols and parameters used in individual studies are presented in Tables 1, 2, 3 and 4.

In relation to saliva collection, the median number of days of saliva collection per time-point across studies was 1 day (IQR 1–2). As the median suggests, the majority of studies (n=57,

Table 1 Randomised controlled trials evaluating psychosocial interventions: study characteristics, salivary cortisol methodology and main findings

Study ID	Main study objective(s)	Study Population[c]	Intervention	Comparator or Control	Main outcome measures	Assessment period	Saliva collection protocol[a]	Diurnal cortisol parameters analysed[d]	Main cortisol findings	Main clinical findings	Quality/ Relevance[b]
Bergen-Cico et al. 2014 (22)	To assess the efficacy of a brief primary care mindfulness programme in military veterans with post-traumatic stress disorder (PTSD) by examining changes in cortisol as a biomarker.	40 military veterans	Primary-care-based brief mindfulness-based stress reduction programme (PCbMP) delivered in weekly 90 min sessions over 4 wks. Home practice encouraged.	Primary Care treatment as usual; they were given the option of receiving PCbMP after 12 wks.	1) Salivary cortisol 2) Psychometric measures for depression and PTSD.	4 wks	2 day 5 sample points (on awakening, 45 min after awakening and 3 other points)	1) Total AUC with respect to ground (AUCg) 2) AUC with respect to increase (AUCi) 3) CAR calculated as the AUC from awakening to 45 min	There was a significant group x time interaction effect on AUCi, with AUCi decreasing in the intervention group and increasing in the control group over time. Pre-post within group analysis indicated that the CAR reduced significantly in the intervention group but not in the control group.	There was a significant negative correlation between changes in AUCi and changes in depression score over time for the group as a whole but no correlation between the changes in cortisol and psychometric variables when each group was analysed separately.	High
Bormann et al. 2009 (23)	To determine the effect of a spiritually-based mantram intervention, in comparison with a control, on average daily salivary cortisol levels in HIV-infected adults.	71 HIV infected adults	Spiritually-based mantram intervention, delivered over 5 weeks in the form of face-to-face classes	Attention-based control intervention delivered in the format	1) Spiritual wellbeing scale, with faith/assurance subscale as a primary outcome 2) Salivary cortisol	10 wks	1 day 4 sample points (7am, 11am, 4 pm, and 9pm)	Mean diurnal cortisol level	There was no significant difference in mean daily cortisol level between groups at 5 and 10 weeks	Faith/assurance levels were significantly higher in the intervention group in comparison with the control group at 5 and 10 weeks	Moderate
Bougea et al. 2013 (24)	To evaluate whether emotional freedom technique affects the frequency and intensity of headaches in adults suffering from tension-type headaches	35 adults suffering from frequent tension-type headaches	Emotional freedom technique practised twice daily over 8 weeks	Standard care (continued use of medication)	1) Frequency and intensity of headaches 2) Perceived stress 3) Global health and health control measures 4) Sleep parameters 5) Salivary cortisol	8 wks	1 day 2 sample points (8am and 8pm)	Time-specific cortisol level	There were no significant differences in cortisol levels (8am or 8pm) between the intervention and control groups at 8 weeks	Headache frequency and intensity, and perceived stress, were significantly lower in the intervention group in comparison with the control at 8 weeks	Moderate
Carlson et al. 2013 (25)	To compare the efficacy of mindfulness-based cancer recovery (MBCR) and supportive-expressive group therapy (SET) in distressed survivors of breast cancer	271 distressed survivors of stage 1-III breast cancer	MBCR delivered in 90 min weekly group sessions over 8 wks plus a 6-hour workshop between weeks 6 and 7	SET delivered in 90 min weekly sessions over 12 wks OR Control condition: 1-day (6-hour) didactic stress management seminar (SMS)	1) Self-reported mood (primary) 2) Salivary cortisol (primary) 3) Self-reported stress 4) Quality of life 5) Self reported social support	Up to 12 wks	3 days 4 sample points (On awakening, noon, 5pm, bedtime)	1) Diurnal slope 2) Time-specific cortisol level	There was a significant group x time interaction effect for diurnal slope, with slopes becoming steeper after SET and MBCR, compared with the control (SMS). In the within-group pre-post analysis, there was no significant change in slope in either MBCR or SET, however, but the slope became significantly flatter in the control (SMS). There was a significant difference in the change in bedtime cortisol between the MBCR group and the control group, with an increase in the SMS group and a slight decrease in the MBCR group.	There were significant group x time interaction effects for mood and stress in the intention to treat analysis. Follow-up pair-wise comparisons for each group indicated no significant differences between the groups for mood disturbance, however. Stress symptoms, however, reduced significantly in the MBCR group in comparison with both the SET and the SMS groups.	High
Cash et al. 2015 (26)	To evaluate the effects of mindfulness-based stress reduction (MBSR) on symptoms and neuroendocrine function in fibromyalgia	91 females with fibromyalgia	MBSR delivered in weekly 2.5hour sessions over 8 wks	Waiting list group x 8 wks	1) Perceived stress 2) Pain 3) Sleep quality 4) Fatigue 5) Fibromyalgia impact scale 6) Salivary cortisol	16 wks	2 days 6 sample points (On awakening, 45 min after awakening, noon, 4pm, 8pm, bedtime)	1) Diurnal slope (excluding +45min sample) 2) Mean waking level 3) CAR-calculated as the mean percent increase from waking to +45 min 4) CAR slope	There were no significant differences between the two groups at 8 wks or 16 wks for any of the cortisol parameters, controlling for baseline values.	There were significant reductions in perceived stress, sleep problems, fatigue and symptom severity in the MBSR group relative to the control at 8 wks, but not in pain or physical functioning. The effects on fatigue were not maintained at 16 wks.	High
Chan et al. 2006 (27)	To investigate the psychophysiological effects of different psychosocial intervention methods for breast cancer patients	76 patients with breast cancer (stage I-III)	Body-mind-soul intervention OR Supportive expressive therapy OR Social support	Non-intervention control	1) Psychological measures including stress 2) Salivary cortisol	32 wks	2 days 5 sample points (On awakening, 45 min after awakening, 12pm, 5pm and 9pm)	1) AUC 2) Diurnal slope	There was a significant reduction in cortisol AUC in the body-mind-soul group, but not the other groups, at 8 months	There was a significant improvement in a range of psychological measures in the body-mind-soul group only.	Moderate
Delle Chiaie et al. 2012 (28)	To investigate if group psychoeducation normalizes HPA axis function in patients with bipolar disorder (preliminary results of an ongoing investigation)	20 patients with stabilized bipolar disorder	Group psycho-education delivered in 21 sessions.	Continuation of treatment as usual (TAU) along with 21 weekly group meetings in which no special instruction was given	1) Depression 2) Mania 3) Compliance with drug treatment 4) Salivary cortisol	20 wks	1 day 5 sample points (08.00am, 30 min and 60 min after, 1pm, 8pm)	The first 3 sample points (representing the CAR) and the last 2 sample points (diurnal decline) were analysed in separate ANOVA models at each timepoint, with sample time as the within group factor and treatment as the between group factor.	There were no significant differences between groups at baseline. After treatment, there was a significant difference between groups in the morning pattern of cortisol secretion (CAR), with an increase in the CAR in the psychoeducation group.	There were no significant changes in the clinical measures	Moderate
Feicht et al. 2013 (29)	To explore whether 'happiness training' improves individual happiness and satisfaction with life and relieves stress in an occupational setting.	147 adult volunteers working in an insurance company	Web-based happiness training for 7 wks	Waiting list group	1) Happiness and satisfaction with life on a visual analogue scale 2) Wellbeing 3) Stress warning signals 4) Mindfulness 5) Recovery experience 6) Flourishing 7) Salivary cortisol and alpha-amylase (subsample only-n=45) 8) Attention network test	11 wks	1 day 3 sample points (on awakening, 30 min later, 8pm)	1) CAR: 30 min sample minus awakening sample 2) 'Morning activity': mean of 30 min sample plus awakening sample 3) 'Evening activity': 8pm level 4) Circadian amplitude: evening sample minus awakening sample	There were no significant within or between group differences for any of the cortisol parameters	There were significant within and between group changes for a range of psychological measures, with evidence of increased happiness and satisfaction in life and reduced stress in the intervention group.	Moderate
Gaab et al. 2006 (30)	To evaluate the effects of cognitive behavioural stress management (CBSM) on psychological and somatic wellbeing and neuroendocrine responses in students preparing for an academic exam.	28 healthy economics students	Cognitive behavioural stress management programme over 4 weeks	Waiting list control	1) Psychological measures including anxiety, stress and depression 2) Somatic symptoms 3) Salivary cortisol	4 wks	1 day 9 sample points (On awakening, 15, 30, 45 and 60 min AND 8am, 11am, 3pm and 8pm)	1) CAR measured by AUC over the first hour 2) Total AUC over full day	There was a significant time x group interaction effect for the CAR, with a significant change in the CAR in the intervention group	There was a significant time x group interaction effect on anxiety and somatic symptoms, with less anxiety and somatic symptoms in the intervention group.	High

Table 1 (continued)

Gex-Fabry et al. 2012 (31)	1) To confirm the efficacy of mindfulness-based cognitive therapy (MBCT) compared with Treatment as Usual (TAU) in reducing relapse risk in patients in remission from recurrent depression 2) To examine possible changes in diurnal salivary cortisol profiles over the follow-up period	60 adult patients in remission from recurrent (≥3 episodes) depression, off anti-depressant medication for at least 3 months	MBCT plus TAU MBCT was delivered in 8 weekly sessions over 2 months	TAU Any care deemed necessary	1) Rate of relapse of depression 2) Time to relapse of depression 3) Salivary cortisol	56 wks	1 day 7 sample points (On awakening, followed by 15, 30, 45 and 60 min post-awakening, and at 3pm and 8pm)	1) CAR: AUC above the minimum concentration in the first hour after awakening 2) Diurnal Slope: difference between wake-up and last evening values, divided by the time interval between these samples 3) Total AUC	There were no significant time x group effects for any cortisol parameter measured, indicating that there was no significant difference between groups at any follow-up point in relation to cortisol	There was no significant difference in depression relapse rate between the two groups over the follow-up period	High
Holt-Lunstad et al. 2008 (32)	To test the hypothesis that couple-based emotional support training that enhances warm physical contact between marital partners may induce increases in oxytocin activity and decreases in stress hormones and blood pressure.	34 healthy married couples	"Warm Touch" Support Enhancement Intervention-training provided over 4 weeks	Couples were advised to keep a diary of physical affection and mood over 4 weeks	1) Ambulatory blood pressure 2) Plasma and salivary oxytocin 3) Salivary cortisol 4) Salivary alpha-amylase	4 wks	1 day 5 sample points (On awakening, 7am, 12pm, 5pm, 10pm)	AUC	Mixed model analysis indicated that there was no main effect of the intervention on salivary cortisol AUC	The intervention had a significant effect on salivary oxytocin and alpha amylase levels but no significant effect on plasma oxytocin levels or ambulatory blood pressure.	Moderate
Hsiao et al. 2011 (33)	To examine the psychobiological effects of psychotherapy coupled with pharmacotherapy versus pharmacotherapy alone in patients with major depressive disorder	63 adults (18-65 years) with major depressive disorder attending an outpatient psychiatry clinic	Combined therapy (COMB): mind-body-spirit psychotherapy delivered over 8 weeks in addition to pharmacotherapy	Monotherapy (MT): pharmacotherapy alone	1) Depression scale 2) Anxiety scale 3) Salivary cortisol	32 wks	1 day 5 sample points (On waking, 45 min after waking, 12 noon, 5pm, 9pm)	1) Time-specific cortisol level 2) Diurnal pattern-modelled using mixed modelling techniques	Mixed model analysis indicated that there was no significant time x group interaction effect on second-order polynomial diurnal cortisol change pattern (planned analysis) but there was a significant time x group interaction effect on first-order polynomial diurnal cortisol change pattern (post-hoc analysis). Analysis of the change in cortisol slope from baseline to each follow-up timepoint indicated that the change in the cortisol slope over the 3 follow-up points differed significantly between COMB and MT, with the slope becoming steeper in the COMB group.	There were no time x group interaction effects or time main effects on depression scores or state anxiety scores. However, there was a group main effect in change in anxiety scores, with a greater change observed in the COMB group.	Moderate
Hsiao et al. 2012 (34)	To understand the effect of a body-mind-spirit intervention on depression state, sense of meaning in life, and diurnal cortisol patterns in breast cancer survivors	48 breast cancer survivors attending a surgical outpatient department	Body-mind-spirit group therapy delivered for 2 hours weekly over 2 months	Educational session delivered as 1 session	1) Depression state measure 2) Meaning in life measure 3) Salivary cortisol	32 wks	1 day 6 sample points (On awakening, 30 min, 45 min, 12 noon, 5pm, 9pm)	1) Diurnal decline: calculated based on regression of the 6 cortisol levels 2) Time-specific cortisol level	The control group developed a significantly higher 21.00 cortisol level and a flatter slope, in comparison with the intervention, over the 8 month follow-up period	There were no significant differences between groups over time in relation to the depression scores and the 'Meaning-in-life Questionnaire-Presence scores'. However, there was a significant difference between groups in the 'Meaning in Life Questionnaire- Search scores' over 5 months (but not 8 months)	Moderate
Hsiao et al. 2014 (35)	To examine the effects of psychotherapy on depressive and anxiety symptoms, suicidal ideation and diurnal cortisol patterns in patients with adjustment disorder and depressed mood.	71 patients with adjustment disorder associated with depressed mood, recruited from a psychiatric outpatient clinic	Body-mind-spirit (BMS) psychotherapy received in addition to treatment as usual. This was delivered in 8 wkly group sessions.	Control: similar to treatment as usual which includes medication and psycho-education	1) Depression 2) Anxiety 3) Occurrence of suicidal ideation 4) Salivary cortisol	56 wks	1 day 6 sample points (on awakening, 30 min after, 45 min after, 12.00, 17.00, 21.00)	Diurnal slope calculated by regressing cortisol values at all 6 sample points against time	There was a significant group x time interaction effect for diurnal slope, with a significantly steeper slope at 56 wks and a trend towards a steeper slope at 8 wks in the intervention (BMS) group compared to the control group.	There were no significant group x time interaction effects for depression and anxiety, with similar reductions in both groups. There was no group x time interaction effect for the occurrence of suicidal ideation but there was a trend toward a greater decrease in occurrence of suicidal ideation in the BMS group compared with the control.	Moderate
Jensen et al. 2012 (36)	To test the effects of mindfulness-based stress reduction (MBSR) on attention in healthy volunteers	48 healthy volunteers who were meditation novices, recruited from a university.	Mindfulness-based stress reduction (MBSR), delivered once weekly in 2.5 hour sessions, over 8 wks. It also included home assignments and an intensive retreat.	Non-mindfulness-based stress reduction (NMSR): resembled MBSR but did not include meditation practices nor training in a non-judgmental attitude OR Non-treatment inactive control	1) 5 attentional tasks 2) Salivary cortisol 3) Perceived stress 4) Mindfulness and awareness	8 wks	1 day 5 sample points (on awakening, 15min,30min, 45min and 60 min after awakening)	1) CAR AUCg: area under the curve with respect to ground, reflecting magnitude 2) CAR AUCi: area under the curve with respect to increase, reflecting pattern	There was a significant within-group reduction in AUCi in the MBSR group with no significant within-group changes in the other groups. There was a group x time interaction effect for AUCg when MBSR was compared to the control but not when MBSR was compared with NMSR.	There was a significant improvement in mindfulness and perceived stress in the MBSR group, with no significant change in the other groups. There was a group x time interaction effect for these measures when MBSR was compared to the control but not when MBSR was compared to NMSR. Attentional measures not relevant to this review.	Moderate
Klatt et al. 2009 (37)	To determine whether low-dose mindfulness-based stress reduction (MBSR-ld) significantly decreases symptoms of stress and to yield adherence rates similar to those for traditional MBSR interventions	48 healthy university staff members aged 18-60 years	MBSR-ld delivered over 6 weeks	Waiting list control over 6 weeks	1) Perceived stress 2) Sleep 3) Level of mindfulness 4) Salivary Cortisol	6 wks	2 days 3 sample points (20 min after waking, 1pm and 10pm)	Mean diurnal cortisol level	There was no significant change in cortisol levels over time in either group.	There was a significant improvement in mindfulness and perceived stress in the intervention group, but not in the control group.	Moderate
Krajewski et al. 2011 (36)	To explore the impact of different ways of spending lunch-time breaks on cortisol.	14 call centre employees	Progressive muscle relaxation (PMR) performed by the participant for 20 min during lunch-break in a 'silent room' over a 6-month period.	Small-talk (ST) for 20 min, with self-chosen colleagues in the staff room, over a period of 6 months.	Salivary cortisol	24 wks	1 day 5 sample points (On awakening, 30 min after, 11.55, 13.05, bedtime) 11.55 and 13.05 samples were done pre and post the intervention to assess its immediate effects	1) CAR (delta): 30 min sample minus awakening sample 2) CAR (mean): mean of awakening plus 30 min sample 3) Pre-post intervention cortisol effect: 13.05 sample minus 11.55 sample 4) Bedtime cortisol: to assess 'spillover' effects Cortisol assessment over 8 timepoints	There was a significant group x time interaction effect for CARdelta and bedtime cortisol, with a reduction in the PMR group relative to the ST group. For CAR, the change occurred towards the end of the 6 month period whereas there was a short-term and long-term effect on bedtime cortisol. Pre-posts acute changes not relevant to this review	N/A	Moderate

Table 1 (continued)

Letour-neau et al. 2011 (39)	To evaluate the effect of a home-based peer support intervention, that included maternal –infant interaction teaching, on mother-infant interactions and other secondary measures (Infant measures not reported in this review as not relevant)	60 mother-infant dyads in which the mother had a diagnosis of post-partum depression < 9 months after delivery	Peer support from a local volunteer, which included teaching on maternal-infant interaction techniques x 12 weeks	Waiting list group x 12 weeks (received peer support after 12 weeks)	1) Mother-infant interaction (primary) 2) Maternal depressive symptoms 3) Maternal perception of social support 4) Maternal salivary cortisol 5) Infant measures (not addressed here)	12 wks	1 day 4 sample points (On awakening, noon, mid-afternoon, before bed)	AUC	There were no significant time or treatment effects on maternal cortisol AUC.	There was a significant difference between groups in maternal-infant teaching interactions, favouring the control. Depressive symptoms improved significantly over time in both groups.	Moderate
Limm et al. 2011 (40)	To test the long-term effect of a stress-management intervention on self-perceived stress reactivity in employees	174 lower and middle management employees in an international plant	Stress management intervention delivered over 2 full days, with booster sessions over the following 8 months	Waiting-list group x 1 year	1) Perceived stress reactivity 2) Effort-reward imbalance measure 3) Anxiety and depression measures 4) Salivary cortisol 5) Salivary α-amylase	1 year	1 day 7 sample points (On awakening, 30 min, 60 min, 8am, 11am, 3pm, 8pm)	1) CAR measured in two ways: a) by subtracting the 'awakening level' from the 30min level and b) by calculating the AUC for the morning values 2) Diurnal decline slope measured by linear regression 3) AUC using all timepoints	There were no significant findings for the salivary cortisol parameters	There was a significant improvement in self-perceived stress reactivity in the intervention group compared with the control	High
Lindh-Astrand et al. 2013 (41)	To study the efficacy of group therapy with applied relaxation on vasomotor symptoms in postmenopausal women	60 healthy postmenopausal women	Applied relaxation (AR) therapy delivered in 10 group sessions over 12 wks	Non-treatment control	1) Average number of moderate and severe hot flashes per 24 hours (primary) 2) Quality of life 3) Salivary cortisol	24 wks	1 day 3 sample points (On awakening, 30 min after, bedtime)	Time-specific cortisol levels	'Morning cortisol' was significantly lower in the treatment (AR) group in comparison with the control at 24 wks but not at 12 wks	There was a significant decrease in number of hot flashes in the treatment group at 12 and 24 wks in the AR group, in comparison with the control group. There was also a significant improvement in a number of quality of life dimensions in the AR group relative to the control.	Moderate
Lipschitz et al. 2013 (42)	To explore whether salivary α-amylase (sAA) and salivary cortisol levels are positively modulated by sleep-focused mind-body interventions in cancer survivors with sleep disturbance (pilot) Non-biological outcomes reported in Nakamura et al. 2013 (58)	57 cancer survivors with self-reported sleep disturbance	Mind-Body Bridging (MBB) classes delivered in once wkly 2-hour sessions over 3 wks OR Mindfulness Meditation (MM) classes delivered in once wkly 2-hour sessions over 3 wks	Sleep Hygiene Education (SHE) classes delivered in once wkly 2-hour sessions over 3 wks	1) Salivary cortisol 2) Salivary sAA 3) Sleep problems 4) Perceived stress 5) Quality of life	3 wks	2 different schedules over 2 days (1 day of each schedule) 4 sample points on day 1 (30min post-awakening, noon, afternoon, evening) 1 sample point on day 2 (on awakening)	1) Waking cortisol 2) AUC: only the 4 measures from the first day included 2) Post-intervention diurnal profile modelled using mixed-effects ANCOVA: only the 4 measures from the first day included, with intervention and intervention x collection time as fixed factors	There were no significant group main effects or group x collection time effects, post- intervention, for any of the cortisol parameters.	Mean sleep problems score was significantly lower in the MBB group in comparison with the SHE group, while MBB and MM did not differ, post-intervention.	High
Lok et al. 2012 (44)	To investigate whether HPA axis activity in recurrent depressive disorder is influenced by cognitive therapy, along with other objectives relating to an embedded case-control study. Non-biological outcomes and RCT design reported in Bockting et al. 2005 (60)	187 highly recurrent major depressive disorder patients	Cognitive therapy delivered in once weekly 2-hour group sessions over 8 wks	Usual care	1) Depression recurrence 2) Salivary cortisol 3) Other secondary self-report outcomes	2 years	2 different schedules over 2 days: 2 sample points on Day 1: 8am and 8pm 1 sample point on day 2: 8am	Profile was modelled using mixed linear modelling with group, follow-up point and sample moment as independent variables	There was a borderline significant effect of steeper cortisol declines over the day throughout the follow-up period.	Cognitive therapy was reported to have had a significant protective effect on depression recurrence over the 2-year follow-up period in Bockting et al. 2005 (60).	High
Nickel C et al. 2007(46) & Nickel M.K. 2007 (47)	To determine the effectiveness of behavioural/psycho-educational group training in men suffering from chronic occupational stress	72 men (≤ 65 years) who self-identified themselves as suffering from chronic occupational stress	Behavioural psycho-educational group therapy delivered in twice weekly 90 min sessions over 8 weeks	90min group meetings twice weekly over 8 weeks, during which participants reported on work events	1) Systolic blood pressure 2) Salivary cortisol 3) Self-report measures of chronic stress, anger and health-related quality of life	8 wks	5 days 4 sample points (On awakening, 15, 30 and 60 min after awakening)	Time-specific cortisol level (averaged over 5 days for each collection time)	There was a significant decrease in cortisol level at all collection time points in the treatment group in comparison with the control group.	There was a significant improvement in systolic blood pressure and in most of the subscales measuring chronic stress, anger and health-related quality of life in the intervention group in comparison with the control group.	Moderate
Nunes et al. 2007 (48)	To examine the effects of relaxation and visualisation therapy (RVT) on psychological distress, cortsol levels, and immunological parameters of breast cancer patients undergoing radiotherapy	34 women with stage I or II breast cancer undergoing radiotherapy	RVT delivered daily in group sessions over 24 days, immediately after radiotherapy	Usual care	1) Psychological distress including self-report scales of stress, anxiety and depression 2) Salivary cortisol 3) Peripheral blood mononuclear cells	24 days	1 day 3 sample points (8am, 12 noon, 8pm)	1) Time-specific cortisol level 2) Total AUC	There was no significant change in cortisol parameters in either the intervention or control group	There was a significant reduction in stress, anxiety and depression scores in the intervention group	High
Oken et al. 2010 (49)	To evaluate whether a mindfulness meditation intervention might be effective in caregivers of close relatives with dementia and to help refine the protocol for future trials (pilot study)	31 healthy adults caring for a close relative with dementia	Mindfulness-based cognitive therapy (MBCT): delivered in once wkly 90 min sessions delivered over 6 wks	Dementia education classes: matched with the mindfulness meditation intervention in relation to time, social support, discussion time and home assignments OR Respite-only intervention: 3 hours wkly for 7 wks	1) Perceived stress (revised memory and behaviour problems checklist) 2) Salivary cortisol 3) Several exploratory secondary measures	7 wks	1 day 3 sample points (On awakening, 30 min after awakening, bedtime)	Time-specific cortisol levels: each collection point analysed separately	There were no significant differences between groups for any of the cortisol measures post-intervention	Perceived stress was significantly lower in the MBCT and education groups in comparison to the respite group post-intervention	Moderate
Pacella et al. 2014 (50)	To examine the impact of successful PTSD treatment on the cortisol awakening response (CAR)	29 adults with chronic PTST (subsample from larger RCT)	Psychotherapy delivered in once weekly 90-120 min sessions over 10 wks	Sertraline titrated upwards as indicated, monitored by a psychiatrist	1) PTSD symptoms 2) Depression 3) Salivary cortisol	10 wks	1 day 4 sample points (on awakening, 30, 45 and 60 min after awakening)	CAR: measured by AUCg and AUCi	Changes in CAR AUC did not differ between groups. CAR AUCi or AUCg were not predicted by clinical response to treatment in a regression analysis.	23 out of the 29 participants responded to either psychotherapy or sertraline in relation to PTSD symptoms.	Moderate

Table 1 (continued)

Plag et al. 2014 (51)	To investigate the effect of cognitive behavioural therapy, in combination with physical activity program, on salivary cortisol and α-amylase levels. Psychometric outcomes reported in Gaudlitz et al. 2015 (66)	59 adults with panic disorder	Cognitive behavioural therapy (CBT) delivered in 90 min sessions over 1 month, followed by a booster session + High level endurance training on a treadmill: 30 min sessions, 3 times per week over 8 wks Both treatments ran parallel.	CBT delivered in the same schedule + Low level exercise: 30 min sessions, 3 times per week over 8 wks	1) Salivary cortisol 2) Salivary α-amylase 3) Anxiety 4) Clinical global impression	28 wks	1 day 5 sample points (08am, 12noon, 4pm, 8pm, 10pm)	AUC	There was a significant group x time interaction effect for AUC in an ANCOVA at 7 months, with a significantly higher AUC in the control (low-level exercise) group.	There was a significant group x time interaction effect for the Hamilton anxiety scale, with a significantly greater improvement in the high level endurance training group at 7 months.	Moderate
Richter et al. 2012 (53)	To examine the effects of a cognitive-behavioural group intervention on perceived stress and salivary cortisol levels in pregnant women	161 pregnant women (10-15 weeks gestation) with symptoms of stress, anxiety or depression	Cognitive behavioural group program delivered in 8 sessions (time period not stated)	Treatment as usual	1) Salivary cortisol 2) Psychological measures including perceived stress, pregnancy-specific distress, anxiety and depression	Up until 3 months post-partum	1 day 5 sample points (On awakening, 30 min after awakening, 11am, 5pm, 10pm)	1) CAR 2) Total AUC	There was a significant time x group interaction for CAR at the second time-point (post-intervention-approx. 8 weeks). Within group analysis demonstrated that there was a significant decrease in CAR in the intervention group but not in the control group.	There were no significant between-group differences for the psychological measures	Moderate
Sears et al. 2007 (54)	To investigate whether an ICD (implantable cardioverter defibrillator) shock and stress management program reduces psychological and physiological markers of anxiety and increases quality of life in patients with ICDs	30 adult patients with ICDs and a history of at least one shock in the previous year	Cognitive Behavioural therapy focused on shock and stress management delivered over 6 weeks	1-day psycho-educational workshop	1) Psychological measures of anxiety and depression 2) Quality of life 3) Device acceptance 4) Salivary cortisol 5) Inflammatory markers	4 wks for cortisol	1 day 6 sample points (times not reported)	Mean diurnal cortisol level	There was a significant reduction in mean cortisol in the overall sample (n=30). However, there was no time x group interaction effect in relation to mean diurnal cortisol level	There was a significant improvement in anxiety, depression and quality of life in the overall sample and improvements in anxiety and mental quality of life were significantly greater the CBT group	Moderate
Taylor et al. 2009 (55)	To determine the effects of improving depression in depressed older patients with elevated cardiovascular risk	48 older adults (≥55 years) with depression and elevated cardiac risk	Cognitive behavioural therapy (at least 10 sessions)	Waiting list control	1) Depression measures 2) Cardiovascular risk factor measurement 3) Salivary cortisol	24 wks	2 days 5 sample points (On awakening, 30 min after awakening, 12 noon, 5pm, 9pm)	1) Diurnal slope 2) Waking cortisol 3) CAR (referred to as the 'cortisol rise after waking'- no details on calculation)	There was no significant difference between groups in relation to cortisol parameters at 6 months	There was a significant improvement in mood in the intervention group in comparison with the control at 6 months	High
Urizar and Munoz, 2011(56)	To determine whether participation in a prenatal cognitive behavioural stress management programme would result in lower cortisol and self-reported stress, relative to the control, in women at high risk for depression during pregnancy Infant measures not reported here as not relevant to the review	57 pregnant (6-28 weeks gestation) women at high risk for depression	Cognitive behavioural stress management programme delivered over 12 weeks, with 4 booster sessions in the post-partum period	Usual care	1) Salivary cortisol 2) Perceived stress 3) Maternal mood	72 wks	1 day 2 sample points (Morning and evening)	1) Morning cortisol 2) Evening cortisol 3) Mean diurnal cortisol level 4) Diurnal decline: difference between morning and evening cortisol values	The mean cortisol level at 18 months was significantly lower in the intervention group in comparison with the control but there were no significant differences at 6 months	Women in the intervention group experienced significantly higher perceived stress than women in the control group at 6 months but no significant difference was found at 18 months	Moderate
Wilcox et al. 2014 (57)	Wilcox et al. performed a secondary statistical analysis of the salivary cortisol data collected in the Well Elderly 2 randomised controlled trial [Clark et al. (70)]. The primary aim of this RCT was to assess the effectiveness of a lifestyle intervention on mental and physical wellbeing and cognitive functioning in older people in a community setting.	460 elderly adults aged 60-95 years. Only 379 participants agreed to provide saliva samples, with 328 providing post-intervention data.	Lifestyle intervention: small group (2hours weekly) and individual sessions (1 hour x 10) led by an occupational therapist over 6 months.	Non-treatment control x 6 months (provided with the intervention upon completion of the study)	1) Physical and mental wellbeing (SF-36) 2) Depression 3) Cognitive outcome variables 4) Life satisfaction index	24wks	1 day 4 sample points (On awakening, 30 min after waking, before lunch, before dinner)	CAR	Two types of statistical analysis approaches were used. 1) Conventional approach: used mean cortisol and mean CAR values 2) Alternative approach: used median cortisol and median CAR values There was no significant change in CAR pre and post intervention using both approaches.	The intervention significantly improved 5 wellbeing subscale scores as well as depression and life satisfaction scores.	Moderate
Yang et al. 2009 (59)	To compare the effects of combination therapy (psychotherapy + antidepressant) and monotherapy (antidepressant alone) on salivary cortisol levels in outpatients with major depression	65 adults attending a psychiatric outpatient clinic with a diagnosis of major depressive disorder	Combination therapy (psychotherapy + antidepressant) Psychotherapy delivered in 8 sessions over 2 months and antidepressant prescribed for 4 months lunch-break in a 'silent room' over a 6-month period.	Monotherapy (antidepressant therapy) x 4 months over a period of 6 months.	1) Salivary cortisol 2) Depression measure	16 wks	1 day 5 sample points (On awakening, 30-45 min post awakening, 12 noon, 5pm, 9pm) 11.55, 13.05, bedtime) 11.55 and 13.05 samples were done pre and post the intervention to assess its immediate effects	1) CAR measured as the difference between morning and 30-45 min value 2) Diurnal slope measured in 2 ways: a) difference between morning and evening values b) difference between 30-45min and evening values 3) Time-specific cortisol levels plus 30 min sample 3) Pre-post intervention cortisol effect: 13.05 sample minus 11.55 sample 4) Bedtime cortisol: to assess 'spillover' effects Cortisol assessment over 8 timepoints	1) There was a significant decrease in the 9pm cortisol level in the combined therapy group in comparison with the monotherapy group over 4 months. 2) There was a significant difference in the diurnal slope between the two groups when measured using values from 30-45 min to evening, with a steeper slope in the combined therapy group. group relative to the ST group. For CAR, the change occurred towards the end of the 6 month period whereas there was a short-term and long-term effect on bedtime cortisol. Pre-posts acute changes not relevant to this review	Combined therapy and monotherapy both resulted in a significant improvement in depression scores, with no significant difference between groups	Moderate

Shaded rows represent studies with agreement between clinical and cortisol findings

[a]Describes the number of consecutive days of sampling and the number of samples per day per time-point

[b]Assessed using Gough's Framework

[c]Refers to the number and type of participants randomised

[d]Abbreviations used for cortisol summary measures: CAR, cortisol awakening response; AUC, area under the curve.

Table 2 Randomised controlled trials evaluating pharmacological (including dietary) interventions: study characteristics, salivary cortisol methodology and main findings

Study ID	Main study objective(s)	Study Population[c]	Intervention	Comparator or Control	Main outcome measures	Assessment period	Saliva collection protocol[a]	Diurnal cortisol parameters analysed[d]	Main cortisol findings	Main clinical findings	Quality/ Relevance[b]
Barbadoro et al. 2013(60)	To evaluate the effects of fish-oil on perceived stress/anxiety and HPA axis activity in abstinent alcoholics	31 men with alcohol dependence attending 4-week inpatient detoxification and rehabilitation programme	1 capsule (1g) of fish oil per day x 3 wks, commenced after completing 1 week of medication-supported detoxification	1 placebo capsule per day x 3 wks, commenced after completing 1 week of medication-supported detoxification	1) Self-perceived stress 2) Basal salivary cortisol levels 3) Salivary cortisol in response to the Trier Social Stress Test (results not presented here)	3 wks	1 day 5 sample points (7.30am, 11.30am, 4pm, 8pm, midnight)	Time-specific cortisol levels	There was a significant group x time interaction effect, with significant decreases in 7.30am, 8pm and midnight cortisol levels over time in the intervention group but not in the control group.	There was a significant decrease in perceived stress in the intervention group but not in the control group.	Moderate
Camfield et al. 2013 (61)	To examine the relationship between chronic multivitamin supplementation, diurnal cortisol secretion and stress in healthy adults.	138 healthy adults	1 multivitamin tablet per day for 16 wks	1 placebo tablet per day for 16 wks	1) Salivary cortisol 2) Perceived stress 3) Plasma homocysteine, Vit B12, Vit B6 and folate	16 wks	1 day 4 sample points (On awakening, 15 min and 30 min after awakening, bedtime)	1) CAR, measured as the maximum value out of the 15 and 30 min samples minus the waking waking 2) Time-specific cortisol levels, in particular morning and evening values	There was a significant group x time interaction effect for CAR, with an increase in CAR in the intervention group and a decrease in CAR in the control group at 16 wks	Perceived stress increased significantly in both groups over 16 weeks, with no significant difference between the two groups	Moderate
Chaborski et al. 2015 (62)	To investigate the effects of a specific amino-acid mixture in combination with micronutrients in psychologically stressed adults attending a cardiology outpatient service. (pilot study)	34 psychologically stressed adults attending a cardiology outpatient service.	1 serving of a nutritional supplement per day (8.9 g powder dissolved in 200ml water) for 12 wks. The supplement consisted of specific amino acids and micronutrients to target stress physiology.	Placebo: 8.9g powder dissolved in 200ml water per day for 12 wks.	1) Psychological stress 2) Salivary cortisol 3) Blood pressure and heart rate 4) Plasma serotonin, CRP, lipids and glucose indices.	12 wks	1 day 2 sample points (morning and evening)	Time-specific cortisol levels	There were no significant within-group or between group changes in the cortisol parameters.	Psychological stress decreased significantly in both the intervention and placebo groups, with a significantly greater decrease in the placebo group.	Moderate
Deuschle et al. 2003 (63)	To study the effects of two different antidepressant classes on HPA axis activity in moderately to severely depressed inpatients	126 adult inpatients with moderate-severe depression	Paroxetine (selective serotonin-reuptake inhibitor) x 35 days (increased to 40mg after the first week)	Amitriptyline (tricyclic antidepressant) x 35 days (increased to 150mg per day after the first week)	1) Depression (Hamilton depression rating scale) 2) Salivary cortisol	5 wks	Participants were instructed to collect saliva daily for 35 days and samples were analysed in weekly blocks. Participants had to provide a minimum of 3 days of samples per week for inclusion in the analysis. Sample points: (8am, 4pm, 10pm)	Time-specific cortisol level: the mean cortisol level for each timepoint was calculated for each week	There was evidence of declining cortisol values over time in both groups, with no significant difference between groups without taking treatment response into account. However, there was a significant group x time x response interaction effect, with only clinical responders to Amitriptyline showing significantly lower levels of 8am and 4pm cortisol levels over the 5 wks.	There was a decrease in depression in both groups with no significant difference over time between groups.	Moderate
Eijsbouts et al. 2008 (64)	To investigate whether Naproxen has a direct or indirect effect on HPA axis activity	40 healthy volunteers	Naproxen capsules 500mg orally twice daily x 2 wks	Placebo capsules orally twice daily x 2 wks	1) Basal salivary cortisol 2) Basal urine cortisol 3) Basal plasma cortisol and ACTH 3) Plasma cortisol and ACTH response to exercise	2 wks	1 day 6 sample points over 24 hours (08am, noon, 4pm, 8pm, midnight, 4am)	AUC	There was no significant difference in AUC between the two groups before and after treatment.	N/A	Moderate
Garrison and Cambliss, 2006 (65) & Kalman et al. 2008 (66)	To determine the efficacy of a dietary supplement ingredient containing proprietary extracts of *Magnolia officinalis* and *Phellodendron amurense* in helping overweight premenopausal women manage their weight (pilot study)	42 overweight, otherwise healthy, premenopausal women who typically eat more in stressful situations and scored above the national mean for self-reported anxiety	250mg capsules (containing *Magnolia officinalis* and *Phellodendron amurense* extracts) three times per day for 6 weeks	Identical placebo capsules	1) Salivary cortisol 2) Weight 3) Psychological measures of stress and anxiety	6 wks	3 days 3 sample points (On awakening, 30 min post awakening, at bedtime)	1) Overall mean diurnal cortisol level: average of 9 measurements over 3 days 2) Average morning cortisol level: over 3 days 3) Average 30 min cortisol level 4) Average evening cortisol level	No statistically significant within-group change for any cortisol parameter, regardless of intervention, from baseline to 6 weeks	1) Statistically significant increase in weight in the placebo group with no significant change in the treatment group 2) Statistically significant improvement in psychological measures in both treatment and placebo groups	Moderate
Hellweg et al. 2008 (67)	To compare changes in serum levels of brain-derived neurotrophic factor (BDNF) and nerve growth factor (NGF) in depressed subjects after antidepressant treatment. A secondary aim was to examine the relationship between neurotrophin and salivary cortisol changes.	40 inpatients with a diagnosis of major depressive disorder	Amitriptyline 150mg once daily x 36 days	Paroxetine 40mg once daily x 36 days	1) Serum BDNF 2) Serum NGF 3) Depression measure 4) Salivary cortisol	5 wks	6 days 2 sample points (8am and 4pm)	Time-specific cortisol level	There were no significant changes between or within groups for salivary cortisol levels	Depression scores improved significantly in both groups with no significant difference between groups. There was a significant increase in BDNF in the Amitriptyline group in comparison with the Paroxetine group. There were no significant changes in NGF.	Moderate
Hinkelmann et al. 2012 (68)	To investigate if changes in cortisol secretion during antidepressant treatment are associated with improvement in cognition This study objective was a secondary objective within a separately reported proof-of-concept RCT [Otte et al.(81)] designed to investigate whether adjunctive treatment with a mineralocorticoid receptor (MR) agonist or antagonist accelerates the onset of action and improves the efficacy of Escitalopram.	52 adult patients with major depression	All patients were commenced on Escitalopram 10-20mg for 3 weeks. They were then randomised to receiving either Fludricortisone (n=19), Spironolactone (n=22) or placebo (n=11) add-on therapy The comparison of the 3 'add-on' treatments was not relevant to the study objective outlined in this paper.	Within-group effects of Escitalopram on cortisol, cognition and depression were compared to an age and sex-matched healthy control group.	1) Salivary cortisol 2) Neuropsychological measures of cognition 3) Depression scale	3 wks	1 day 4 sample points (8am, 12pm, 4pm, 10pm)	AUC	AUC decreased significantly after 3 weeks of Escitalopram	1) Significant improvement in depression after 3 weeks of Escitalopram (returned to healthy control level) 2) Significant improvement in some cognitive variables after 3 weeks of Escitalopram in comparison with the healthy control group	Moderate

Table 2 (continued)

Knorr et al. 2012(70)	To explore the specific pharmacodynamic effect of Escitalopram on hypothalamic-pituitary-adrenal (HPA) activity compared to placebo in healthy participants	80 healthy first-degree relatives of patients with depression (aged 18-60 years)	Escitalopram 10mg once daily x 4 weeks	Placebo tablets once daily x 4 weeks	1) Salivary cortisol 2) Perceived stress 3) Depression scale 4) Anxiety scale 5) Pain scale 6) Sleep scale 7) Quality of life scale 8) Aggression scale	4 wks	1 day 8 sample points (On awakening, followed by 15, 30, 45 and 60 min post-awakening, and at 12.00, 18.00 and 23.00)	1) CAR: measured by AUC for the first hour 2) Total AUC for the day	There was a significant decrease in the CAR AUC (P=0.04) and the all-day cortisol AUC (P=0.02) in the Escitalopram group compared to the placebo group. However, when a correction was made for multiple outcome measures, this finding did not reach statistical significance.	There was no significant change in clinical outcome measures between the two groups	High
Lenze et al. 2011 (71) & Lenze et al. 2012 (72)	To investigate the impact of Escitalopram versus placebo on the HPA axis in older adults with generalised anxiety disorder (GAD)	95 older adults (≥60 years) with GAD	Escitalopram 10mg once daliy x 12 weeks	Placebo x 12 weeks	1) Salivary cortisol 2) Anxiety measures 3) Cognitive measures	12 wks	2 days 6 sample points (On awakening, 30 min after awakening, noon, 4pm, 8pm, bedtime)	1) Peak cortisol level: defined as the cortisol level 30 min after awakening 2) AUC	There was a significant reduction in peak cortisol in the Escitalopram group in comparison with the placebo group. This was most marked for the subgroup with elevated cortisol levels at baseline.	The authors did not provide a report of the within or between group changes for the anxiety and cognitive measures. However, there was a significant association between improvement in memory and reduction in cortisol indices in the Escitalopram group. In addition, subgroup analysis indicated a significant association between change in cortisol indices and change in anxiety for those with high baseline cortisol levels.	Moderate
Lopresti et al. 2015(73)	To explore the effects of Curcumin for the treatment of major depression on blood, urinary and salivary biomarkers. Non-biological outcomes reported in Lopresti et al. (86)	50 adults with currently major depressive disorder	Curcumin capsule 500mg twice daily x 8 wks	Placebo capsule 500mg twice daily x 8 wks	1) Depression 2) Anxiety 3) Salivary cortisol 4) A range of other biomarkers in urine and plasma	8 wks	1 day 2 sample points (within 10 min of awakening and 30 min later)	1) CAR 2) Awakening cortisol 3) 30 min cortisol	There were no significant group x time interaction effects for the any of the cortisol parameters over 8 wks	There were no significant group x time interaction effects for anxiety or depression scores over the 8 wks, with a reduction in both groups. However, the reduction in symptoms was greater in the Curcumin group from 4 to 8wks.	Moderate
Mocking et al. 2014 (75)	To investigate the effect of varenicline on the HPA axis, in order to explore potential depressogenic effects at the biological level. (Varenicline is a partial nicotinergic acetylcholinergic receptor agonist)	41 healthy never-smoking volunteers	Varenicline 1 capsule per day on days 1-3 and day 7 and two capsules per day on days 4-6.	Placebo capsules in the same schedule.	Salivary cortisol	1 wk	1 day 3 sample points (On awakening, 15 min and 30 min after awakening)	CAR: sample moment was entered as a factor in the repeated measures ANOVA analysis	There was no significant group x treatment or group x treatment x 'sample moment' effect, indicating that there was no significant change in CAR between groups.	N/A	Moderate
Nonino-Borges et al. 2007 (76)	To determine the influence of meal time on salivary circadian cortisol rhythm and weight loss in obese women (Crossover design)	12 obese, otherwise healthy, women	Hypocaloric diet served between 9am and 11am OR Hypocaloric diet served between 6pm and 8pm	Five hypocaloric meals per day spread throughout the day from 9am to 8pm	1) Salivary cortisol 2) BMI and body composition measurements 3) Resting metabolic rate 4) Urinary nitrogen	18 days	1 day 6 sample points (8am, 12 noon, 5pm, 7pm, 9pm, and 8am the following morning)	1) Presence or absence of a circadian rhythm, defined as a 9pm value of <90% of the morning cortisol value 2) Mean diurnal cortisol concentration between groups at day 1 or day 18	The diurnal cortisol rhythm was preserved from day 1 to day 18 in all groups, with no significant change observed. There was no significant difference in mean cortisol concentration between groups at day 1 or day 18	There were significant reductions in most measures of body compositions from day 1 to day 18 in all groups but no significant difference between groups.	Moderate
Ruhe et al. 2015 (77)	To examine the longitudinal effects of the SSRI, paroxetine, on salivary cortisol in major depressive disorder. A randomised controlled sub-study was nested within this larger longitudinal study, commencing after 6 wks follow-up and continuing for another 6 wks. We describe the RCT component only.	42 patients with major depressive disorder who were identified as non-responders following treatment with paroxetine 20mg per day over 6 wks in the larger cohort study.	Paroxetine dose escalation every 5 days up to a maximum of 50mg per day	Placebo dose escalation. These patients continued on paroxetine 20mg per day and were given placebo capsules during the dose-escalation phase.	1) Clinical response-primarily depression measured by the Hamilton Depression Rating Scale 2) Salivary cortisol	6 wks	1 day 2 sample points (On awakening, 30 min after)	1) Awakening cortisol level (referred to as 'baseline cortisol level') 2) CAR: 30min value minus the awakening value 3) CAR AUC: awakening value + (1/2xCAR)	There were no significant group x time interaction effects for any of the cortisol parameters	There was no statistically significant difference in remission rate between the two groups with a remission rate of 14.3% in the dose-escalation group and 9.5% in the placebo-dose-escalation group.	Moderate
Scharn-holtz et al. 2010 (78)	To test the hypothesis that the HPA system-dampening effects of venlafaxine and mirtazapine differ. Crossover study	95 patients with a major depressive episode	Venlafaxine daily in flexible doses x 4 wks (minimum= 75mg per day)	Mirtazapine daily in flexible doses x 4 wks (minimum= 30mg per day)	1) Depression (Hamilton Depression Rating Scale) 2) Salivary cortisol	4 wks	1 day per wk, each week over 5 wks 2 sample points (8am, 4pm)	Time-specific cortisol level	There was a significant decrease in 4pm cortisol levels in the Mirtazapine group from baseline to week 4. There were no significant changes in cortisol levels at 8am or 4pm in the Venlafaxine group.	There was no significant difference between groups in relation to remission frequency over time. In addition, the decline in cortisol in the Mirtazapine group occurred in both 'remitters' and 'non-remitters' suggesting that the clinical effects and cortisol effects of Mirtazapine are unrelated.	Moderate
Schmidt et al. 2015 (79)	To explore the effects of two prebiotics on the secretion of cortisol and on emotional processing in healthy volunteers	48 healthy volunteers	One of two prebiotics: 1) Fructo-oligosaccharides (FOS) OR 2) Bimuno-galacto-oligosaccharides (B-GOS) 5.5g of powder with breakfast per day x 3 wks	Placebo: 5.5g powder per day with breakfast x 3 wks	1) Salivary cortisol 2) Processing of emotional stimuli 3) Perceived stress	3 wks	1 day 5 sample points (On awakening, every 15 min after awakening for 1 hour)	CAR: modelled in ANOVA with sample-moment as a within-subject factor and also calculated as AUC with respect to ground.	There was a significant decrease in CAR in the B-GOS group in comparison with placebo over 3 wks.	The B-GOS group showed significantly increased attentional vigilance to positive versus negative stimuli in comparison with placebo.	High
Schubert et al. 2011 (80)	To investigate the effect of high-content milk-based phospholipids on HPA axis activity and reactivity in chronically stressed men. (Details relevant to HPA axis activity only described here)	75 healthy adult non-smokers with elevated chronic stress scores at screening (screening questionnaire =Trier inventory for Chronic Stress)	Milk drink containing 0.5% phospholipid (1 carton per day over 42 days) OR Milk drink containing 1% phospholipid (1 carton per day over 42 days)	Placebo milk drink (1 carton per day over 42 days)	Salivary cortisol	6 wks	3 days 4 sample points (On awakening and 30 min, 45 min, 60 min after waking)	CAR measured by subtracting the 'awakening level' from the 60 min level	There were no significant findings for CAR	N/A	Moderate

Table 2 (continued)

Talbott et al. 2013a (81)	To study the effects of 'Tongkat Ali' supplementation (South East Asian "anti-aging" medicinal herb) in moderately stressed adults.	63 moderately stressed adults	Tongkat Ali (TA) herbal medicine 200mg per day x 4 wks	'Look-alike' placebo daily x 4 wks	1) Mood state (primary) 2) Salivary cortisol (primary) 3) Salivary testosterone (primary) 4) Liver enzymes 5) Body weight 6) Body fat percentage	4 wks	1 day 3 sample points (morning, afternoon, evening)	Mean diurnal cortisol level	Mean cortisol level was significantly lower in the treatment group compared to placebo at 4 wks. However, mean cortisol appeared to increase in both groups over the 4 wks.	There was a significant reduction in anger, tension, and confusion in the treatment group compared with the placebo group at 4 wks, but no significant change in depression, vigour, fatigue and overall wellbeing.	Moderate
Talbott et al. 2013b (82)	To study the effects of 'Relora' (combination of magnolia bark and phellodendron bark) on salivary cortisol and psychological well-being in healthy adults with stress.	60 healthy adults with stress	Relora 250mg per day, at breakfast and dinner time, x 4 wks	'Look-alike' placebo in the same schedule x 4 wks	1) Body fat percentage 2) Body weight 3) Stress 4) Mood state 5) Salivary cortisol	4 wks	1 day 3 sample points (on awakening- approx 6am, afternoon- approx 2pm, evening- approx 10pm)	Mean diurnal cortisol level ('cortisol exposure') Authors indicate that the 'circadian rhythm data' will be reported elsewhere	Salivary cortisol exposure was significantly lower in the treatment group at 4 wks. Change data not reported.	At 4 wks, global mood state and overall stress were significantly lower in the treatment group. No differences existed for body composition parameters.	Moderate
Tucker et al. 2004 (83)	To explore the effects of a selective serotonin-reuptake inhibitor (SSRI) versus placebo on cortisol and cytokines in patients with chronic PTSD	59 patients with chronic PTSD	Citalopram 20-50mg per day x 10 wks OR Sertraline 50-200mg per day x 10 wks	Placebo pill identical in appearance x 10 wks	1) Salivary cortisol 2) Serum interleukins (IL-1β, IL-2R) 3) PTSD symptoms 4) Depression	10 wks	1 day 2 sample points (8am, 4pm)	Time-specific cortisol levels	There were no significant changes in cortisol levels in either of the treatment groups. There was a significant increase in 4pm cortisol in the placebo group over 10 wks.	There was a significant improvement in PTSD and depression symptoms in all 3 groups, including the placebo group.	Moderate
Walsh et al. 2006 (84)	To evaluate the impact of enhanced slow-wave sleep on behavioural, psychological and physiological changes associated with sleep restriction	41 healthy adults subjected to 4 nights of sleep-restriction	Tiagabine 8mg x 4 nights, during sleep restriction	Placebo x 4 nights, during sleep restriction	1) Poly-somnography 2) Multiple Sleep Latency Test 3) Vigilance 4) Mood 5) Sleepiness 6) Learning, memory, perception & cognition tests 7) Heart rate variability 8) Salivary cortisol 9) Urinary catecholamines	7 days	1 day 8 sample points (hourly sampling from 2.20pm to 9.20 pm)	1) Mean overall cortisol 2) AUC	Mean diurnal cortisol and cortisol AUC increased significantly in the placebo group relative to the treatment group over the 7 days.	Tiagabine significantly increased slow-wave sleep as measured by polysomnography. This was accompanied by better scores on key executive function tests.	Moderate
Wit-bracht et al. 2013 (85)	To determine the influence of energy restriction and dairy food consumption on salivary cortisol in obese women	51 overweight (otherwise healthy) females who were habitually low-dairy consumers	Low amount of dairy food, with energy restrictions, for 15 weeks	Adequate amount of dairy food, with energy restrictions, for 15 weeks	1) Salivary cortisol 2) Weight loss 3) Fat loss	12 wks	1 day 6 basal sample points (On awakening, 15 min after awakening, prior to lunch, prior to dinner, at 7.45pm and prior to bed)	1) CAR (measured by subtracting 15 min sample from awakening sample) 2) Time-specific cortisol level 3) Mean diurnal cortisol level 4) Minimum cortisol 5) Maximum cortisol 6) Amplitude of cortisol profile 7) Variance of cortisol	In the overall sample (data from both groups combined), there was a significant increase in minimum cortisol and a significant decrease in amplitude of cortisol. There was no significant difference between groups for any of the basal cortisol parameters	Both groups lost weight and fat due to energy restriction but there was no significant difference between groups for these parameters	Moderate

Shaded rows represent studies with agreement between clinical and cortisol findings

[a]Describes the number of consecutive days of sampling and the number of samples per day per time-point

[b]Assessed using Gough's Framework

[c]Refers to the number and type of participants randomised

[d]Abbreviations used for cortisol summary measures: CAR, cortisol awakening response; AUC, area under the curve.

73.1 %) collected saliva over 1 day only. Ten studies used protocols of 2 days, eight used protocols of 3 days and three used protocols of 4–6 days. The median number of samples collected per day was 4 (IQR 3–5), ranging from a minimum of two samples per day (n=12) to a maximum of nine samples per day (n=3). Seventeen studies (21.8 %) collected six or more samples per day. There was substantial variation in the timings of the samples per day. Importantly, of the 75 studies which reported sample times, only 49 (62.8 %) included an awakening sample, suggesting that many studies did not use waking time as a reference for subsequent diurnal sampling points, choosing clock times in preference.

A wide range of different salivary diurnal cortisol parameters were analysed both within and across studies. In relation to composite measures of diurnal cortisol, the cortisol awakening response was measured in 25 studies (32.1 %), the area under the curve from morning to evening was measured in 22 studies (28.2 %) and the diurnal decline/slope was measured in 18 studies (23.1 %). Five studies (6.4 %) modelled the diurnal profile using multi-level modelling techniques, obviating the need to calculate these composites separately. It was also common for studies to report on changes in mean cortisol levels across the day (18, 23.1 %) or changes in absolute cortisol levels as specific times during the day (34, 43.7 %). The majority of studies (42, 53.8 %) used just one of these methods of analysing and reporting diurnal cortisol, but 25 studies (32.1 %) used two methods and 11 studies (14.1 %) used three or more methods.

Overall, 39 studies (50 %) measured an indicator of circadian rhythm by calculating either a cortisol awakening response or a diurnal slope or by modelling the diurnal profile. Interestingly, this proportion differed between study intervention categories: psychosocial intervention studies (22/33, 66.7 %), pharmacological studies (8/22, 36.4 %) and complementary therapy studies (7/14, 50 %). Only nine studies (11.5 %) calculated both the cortisol awakening response and the diurnal slope. In addition, however, five studies modelled these simultaneously using multi-level modelling techniques, suggesting that 14 studies in total (17.9 %) analysed both indicators of HPA axis regulation. Only two studies (2.6 %) used all three diurnal profile parameters (cortisol awakening response, diurnal slope and area under the curve) as recommended by Adam et al. for epidemiological studies [11].

Table 3 Randomised controlled trials evaluating complementary therapies: study characteristics, salivary cortisol methodology and main findings

Study ID	Main study objective(s)	Study Population[c]	Intervention	Comparator or Control	Main outcome measures	Assessment period	Saliva collection protocol[a]	Diurnal cortisol parameters analysed[d]	Main cortisol findings	Main clinical findings	Quality/ Relevance [b]
Banasik et al. 2011 (86)	To examine the effect of regular Iyengar yoga practice on self-perceived psychosocial function and diurnal salivary cortisol in breast cancer survivors (pilot)	18 stage II-IV breast cancer survivors	Yoga practice delivered in 90 min sessions twice weekly for 8 wks	Waiting list x 8 wks	1) Salivary cortisol 2) Perceived quality of life (FACT B)	8 wks	2 days 4 sample points (morning, noon, 5pm, 10pm)	1) Time-specific cortisol levels 2) Diurnal slope	There were no significant group x time interaction effects for any of the cortisol parameters. Pre-post within- group analysis indicated that morning and 5pm cortisol levels decreased significantly at 8 wks in the yoga group, with no significant changes in the control group.	There was a significant group x time interaction effect for fatigue score, decreasing over time in the yoga group and increasing over time in the control group. Pre-post within group analysis indicated that emotional wellbeing and fatigue improved significantly at 8 wks in the yoga group, with no significant changes in the control group.	Moderate
Billhult et al. 2008 (87)	To study the effect of repeated effleurage massage on cellular immunity in patients with breast cancer. The study of its effect on cortisol was a secondary aim.	22 women who had recently undergone surgery for breast cancer and were about to commence adjuvant radiotherapy	Effleurage massage therapy delivered in 10 sessions over 3-4 weeks, each session taking place immediately after a scheduled radiation treatment.	Unstructured conversation with the massage therapist but no massage, following the same time schedule as the intervention group.	1) Biomarkers of immunity 2) Salivary cortisol 3) Plasma oxytocin 4) Anxiety, depression and quality of life.	4 wks	1 day 2 sample points (6am and 10pm)	No information provided but presumed to have used time-specific cortisol levels.	There were no significant changes in cortisol between groups	There were no significant changes in any of the non-cortisol outcome measures between groups	Moderate
Bower et al. 2014 (88)	To examine the effects of Iyengar yoga on genomic and circulating markers of inflammation in fatigued breast cancer survivors. This was a secondary objective of the RCT described. The primary objective of examining the effects of Iyengar yoga on fatigue and associated results are reported in Bower et al. (98).	31 stage 0-II breast cancer survivors, at least 6 months post-treatment, suffering from cancer-related fatigue	Iyengar yoga delivered in 90 min sessions twice weekly over 12 wks	Health education classes delivered in 120 min sessions once a week over 12 wks	1) Self-reported fatigue 2) Psychological: depression, sleep, stress 3) Disease-related and function-related measures 4) Gene expression profiling 5) Plasma inflammatory markers 6) Salivary cortisol	24 wks	2 days 4 sample points (on awakening, 30 min and 8 hrs after awakening, bedtime)	1) Diurnal slope 2) AUCg (excluding 30 min value)	There were no significant group x time interaction effects for diurnal slope or AUC and there were no significant changes in these parameters in the pre-post within group analysis.	There was a significant decrease in fatigue severity at 12 wks and 24 wks in the yoga group relative to the control group.	High
Campo et al. 2015 (90)	To assess the effect of tai chi chih (TCC) on blood pressure, salivary cortisol and inflammatory cytokines in senior female cancer survivors. This was a secondary objective of the RCT described. The primary objective was to assess the feasibility, acceptability and health-related quality of life outcomes of tai chi chih (TCC) in senior female cancer survivors; the results relating to this objective are reported in Campo et al. 2013 (100). (feasibility)	63 senior female cancer survivors, the majority surviving breast cancer	TCC delivered in 60 min sessions, 3 days per wk, over 12 wks	Health education delivered in 50 min sessions, 3 days per wk, over 12 wks	1) Feasibility and acceptability 2) Quality of life 3) Blood pressure 4) Salivary cortisol 5) Inflammatory cytokines	13 wks	1 day 5 sample points (On awakening, 30 min after awakening, noon, 5pm and 10pm)	1) CAR 2) Diurnal slope (from the 30 min sample to the 10pm sample) 3) AUC	The AUC was significantly lower in the intervention group relative to the comparator at 12 wks. There were no significant differences for the other parameters.	Quality of life scores did not differ significantly between groups at 12 wks. A large proportion of participants in both groups showed an improvement.	Moderate
Chandwani et al. 2014 (92)	To investigate the long-term effects of yoga in women with breast cancer undergoing radiotherapy	178 women with stage 0-III breast cancer being scheduled to undergo a 6-wk course of radiotherapy	Yoga classes lasting 60min up to 3 times weekly for 6 wks during radiotherapy	Stretching classes lasting 60 min up to 3 times weekly for 6 wks during radiotherapy OR Waiting list control group	1) Quality of life (physical and mental component scales) 2) Fatigue 3) Sleep disturbances 4) Depression 5) Salivary cortisol	30 wks	3 days 5 sample points (On awakening, 45 min, 8 hrs and 12 hrs after waking, bedtime)	1) Diurnal slope (without the waking sample) 2) Waking cortisol levels	The yoga group had a significantly steeper slope than the stretching and waiting list groups at 6 wks but this was not significant when missing data was imputed using a multiple imputation technique.	There was a significant increase in the physical component scale score in the yoga group relative to the waiting list group at 1 and 3 months after treatment completion (10wks and 18 wks from baseline). There was also a significant decrease in fatigue in both the yoga groups and stretching groups compared with the waiting list group at 6 wks.	High
Chen et al. 2013 (93)	To examine the efficacy of a qigong intervention on quality of life in women with breast cancer during and after radiotherapy.	96 women with breast cancer who were scheduled to received radiotherapy	Qigong delivered in once wkly 40 min classes during 5-6 wks of radiotherapy treatment.	Waiting list control group	1) Depression 2) Fatigue 3) Sleep disturbance 4) Overall quality of life 5) Salivary cortisol rhythm	18 wks	2 days 4 sample points (On awakening, 45 min and 8 hours after waking, bedtime)	1) Diurnal slope: cortisol level regressed against sample time 2) CAR: 45 min value minus awakening value	There were no significant group x time interaction effects for diurnal slope or CAR.	There was a significant group x time interaction effect for depressive symptoms, with a significant reduction in the treatment group relative to the waiting list group over 3 months.	High
Corey et al. 2014 (94)	To investigate the effect of restorative yoga on psychosocial and cortisol outcomes in adults with metabolic syndrome Metabolic outcomes for this RCT are reported in Kanaya et al. 2014 (103)	180 adults with metabolic syndrome	Restorative yoga classes: 90 min classes twice wkly for the first 12 wks, once wkly for the next 12 wks and then once monthly for another 6 months	Stretching classes: 90 min classes twice wkly for the first 12 wks, once wkly for the next 12 wks and then once monthly for another 6 months	1)Salivary cortisol 2) Depression 3) Self-reported social support 4) Positive affect 5) Self-reported stress	1 year overall but 24 wks for cortisol	3 days 4 sample points (On awakening, 30 min and 60 min after waking, bedtime)	1) Diurnal slope: waking cortisol minus evening cortisol (averaged over 3 days) 2) Cortisol percent change: evening cortisol minus waking cortisol divided by waking cortisol 3) Time-specific cortisol levels 4) Mean cortisol	There was a significant decrease in mean cortisol, waking cortisol and evening cortisol in the stretching group compared to the restorative yoga group. There were no other significant changes between groups at 6 months, controlling for baseline cortisol.	There were no significant between-group changes for any of the psychosocial variables at 6 months. Within-group pre-post analysis identified significant decreases in the severity of stressor and cognitive stress subscales in the stretch group at 6 wks, but not in the yoga group.	High
Danu-calov et al. 2013 (96)	To investigate whether a yoga and compassion mediation program (YCMP) alters stress, anxiety and depression levels in familial caregivers of patients with Alzheimer's disease	53 familial caregivers of patients with Alzheimer's disease	Yoga and compassion meditation programme (YCMP) delivered in 1.25 hr sessions, 3 times wkly over 8 wks	Non-treatment control group. These participants were offered the treatment upon study completion.	1) Self-reported stress 2) Depression 3) Anxiety 4) Salivary cortisol	8 wks	2 days 2 sample points (On awakening and 30min after)	1) Mean of the two samples on day 1 2) Mean of the two samples on day 2 3) Mean of day 1 and day 2 means	There was a significant reduction in mean morning cortisol (mean of day 1 and 2) in the intervention group in comparison with the control group.	There was a significant improvement in stress-related symptoms and a reduction in anxiety and depression in the YCMP group in comparison to the control group	Moderate

Table 3 (continued)

Hodgson and Lafferty, 2012 (97)	To investigate and compare the effects of reflexology and Swedish massage therapy on physiologic stress, pain and mood in older cancer survivors residing in a nursing home (crossover design) (pilot study)	18 nursing home residents, ≥ 75 years, diagnosed and treated for cancer within the past 5 years	Reflexology provided in 4 sessions over 4 weeks	Swedish Massage provided in 4 sessions over 4 weeks	1) Salivary cortisol 2) Pain 3) Mood	4 wks	1 day 4 sample points (7-7.30am, 11-11.30am, 1-1.30pm and 4-4.30pm)	Mean diurnal cortisol level across the day	There was a significant within-group reduction in mean cortisol level for both groups but no significant difference between groups.	There were significant changes in negative affect, positive affect and pain within both groups	Moderate
Huang et al. 2012 (98)	To explore the use the diurnal profile of salivary cortisol concentration as a marker of changes in stress following traditional Chinese acupuncture (TCA) and to evaluate whether TCA normalises the diurnal salivary cortisol profile in chronically stressed adults (exploratory study)	18 adult volunteers with high self-reported stress levels	Traditional Chinese Acupuncture (TCA) delivered over 5 weeks OR Attention only delivered over 5 weeks	Waiting list group x 5 weeks n=6	1) Salivary cortisol 2) Perceived stress	7 wks	1 day 4 sample points (On awakening, 30 min after waking, 3 hours after waking and 12 hours after waking)	1) CAR: awakening value subtracted from the 30 min value 2) Diurnal slope: 12 hour valued subtracted from awakening value 3) Mean daytime cortisol level: average of awakening, 3 hour and 12 hour values	There were no significant findings for cortisol parameters. However, there was a non-significant increase in CAR in both the acupuncture and attention groups.	There was no significant difference within or between groups for perceived stress score.	Moderate
Raghavendra et al. 2009 (99)	To determine the effects of a yoga program (versus comparator) on stress and HPA axis function in patients with non-metastatic breast cancer	88 women with stage II and stage III breast cancer undergoing adjuvant radiotherapy	Integrated yoga program delivered over 6 weeks	Brief supportive therapy delivered over 6 weeks	1) Psychological measures including anxiety, depression and perceived stress 2) Salivary cortisol	6 wks	3 days 3 sample points (6am, 9am, 9pm)	1) AUC 2) 'Pooled mean diurnal cortisol' level over the 3 days 3) Time-specific cortisol levels	There was a significant reduction in 6.00am cortisol and pooled mean diurnal cortisol in the intervention group in comparison with the control	There was a significant reduction in anxiety, depression and perceived stress measures in the intervention group in comparison with the control	Moderate
Schneider et al. 2007 (100)	To evaluate the effect of acupuncture on the neuroendocrine and autonomic nervous system in patients with irritable bowel syndrome. This was a secondary analysis of an RCT evaluating the efficacy of acupuncture and reported in Schneider et al. 2006 (108).	43 adults with irritable bowel syndrome	Acupuncture delivered twice per week over 5 weeks	Sham acupuncture delivered twice per week over 5 weeks	1) Salivary cortisol 2) Autonomic function 3) Quality of life	5 wks	1 day 4 sample points (7am, 12 noon, 5pm, 10pm)	Time-specific cortisol level	There was a significant decrease in salivary cortisol level at each collection time after the intervention but not after the control.	There was a significant increase in parasympathetic tone in the intervention group in comparison with the control but there was no significant difference between the groups in quality of life.	Moderate
Tornhage et al. 2013 (102)	To assess cortisol concentrations before, during and immediately after tactile massage in patients with Parkinson's Disease (short-term and long-term effects)	45 patients with Parkinson's Disease	Tactile massage: 2 sessions per wk for the first 3 wks, then 1 session per wk for 4 wks	Rest to music: same schedule and setting	Salivary cortisol 1) Diurnal (long-term effects) 2) Immediate pre-post measures before and after 1st and 8th intervention Details of collection schedule and parameters for diurnal measurements recorded here only	34 wks	1 day for diurnal cortisol 4 sample points over 24 hours: 8am, 1pm, 8pm and 8am the following morning	1) Total daytime secretion: AUC from 8am to 8pm 2) Total night-time secretion: AUC from 8pm to 8am 3) Median cortisol concentration 4) Time-specific cortisol levels	There no significant changes in cortisol levels or AUC within or between groups over the follow-up period.	N/A	Moderate
Woods et al. 2009 (103)	To examine the effect of therapeutic touch on behavioural symptoms in people with dementia	65 people with a diagnosis of dementia	Therapeutic touch for 5-7 min delivered twice daily for 3 days. After 5 days, this was treatment was repeated as a booster.	Placebo therapeutic touch: same touch process but without meditative approach, delivered in same schedule OR Routine care	1) Observed behavioural symptoms using a validated scale 2) Salivary cortisol	16 days	Samples were collected daily over the 20 day study period, which was divided into 5 time periods for analysis. This suggests that, on average, the sample period was over 4 days. 4 sample points (on awakening, 30 min, 6 hours and 12 hours after awakening)	Mean daily cortisol level: average of the 4 sample points	There was no significant group x time interaction effect.	There was no significant group x time interaction effect for the behavioural symptoms, though pairwise comparisons indicated a significant decrease in restlessness in the treatment group in comparison with the control, but not the placebo.	Moderate

Shaded rows represent studies with agreement between clinical and cortisol findings

[a]Describes the number of consecutive days of sampling and the number of samples per day per time-point

[b]Assessed using Gough's Framework

[c]Refers to the number and type of participants randomised

[d]Abbreviations used for cortisol summary measures: CAR, cortisol awakening response; AUC, area under the curve.

Cortisol Findings: Overall Patterns Across Studies and Patterns Within Intervention Categories

Of the 78 included studies, 40 (51.2 %) reported a significant within- or between-group difference in at least one cortisol parameter in response to the experimental intervention or comparator. The significant effects for cortisol parameters were reported at a variety of different follow-up time-points from baseline, ranging from 1 week [38] to 72 weeks [56]. Fifty per cent occurred at a median of 6 weeks from baseline (IQR 4–12).

Seventy-four studies reported both cortisol and clinical findings, and these findings were in agreement in only 50 % of cases; these studies are shaded in Tables 1, 2, 3 and 4. The rate of agreement between clinical and cortisol findings differed between study intervention categories: psychosocial interventions (11/32, 34.4 %), pharmacological studies (11/20, 55 %) and complementary therapy studies (9/13, 69 %). In most cases of disagreement, significant effects for clinical outcome measures were not accompanied by significant effects for cortisol measures (25/37, 67.5 %). In some cases, significant effects were found for cortisol measures without significant effects for clinical measures (10/37, 27 %), and in two cases, significant effects occurred at different time-points for the two types of measures.

As expected, due to wide heterogeneity across studies in multiple domains (e.g. interventions, populations and cortisol parameters), it was not possible or meaningful to summarise the cortisol findings across all studies. Therefore, as an example of the type and range of findings reported, we chose to compare and contrast the cortisol outcomes for one study population, the breast cancer population, following similar interventions. This population was chosen as a number of larger RCTs have been conducted in this population in recent years,

Table 4 Randomised controlled trials evaluating a range of 'other' interventions: study characteristics, salivary cortisol methodology and main findings

Study ID	Main study objective(s)	Study Population[c]	Intervention	Comparator or Control	Main outcome measures	Assessment period	Saliva collection protocol[a]	Diurnal cortisol parameters analysed[d]	Main cortisol findings	Main clinical findings	Quality/ Relevance[b]
Boelens et al. 2009 (104)	To investigate the effect of direct person-to-person prayer on depression, anxiety, positive emotions and salivary cortisol levels (Parallel RCT with option of control group crossover into intervention group after 6 weeks)	63 adults with signs/symptoms of depression and/or anxiety (either self-reported or referred by doctor/nurse)	Prayer intervention delivered for 1 hour per week over 6 weeks	Non-treatment control	1) Depression and anxiety measures 2) Measures of life orientation and spiritual experiences 3) Salivary cortisol	10 wks	1 day 4 sample points (8am, 12 noon, 5pm, 9pm)	AUC	There were no significant differences within or between groups in relation to cortisol AUC	There was a significant improvement in depression, anxiety, spiritual experiences and life orientation scores in the intervention group in comparison with the control group at 6 weeks	Moderate
Dudgeon et al. 2012 (105)	To determine the effect of a low-volume, moderate-intensity resistance and aerobic exercise training programme in HIV-infected men	59 men infected with HIV	Low-volume, moderate intensity resistance and aerobic exercise training: 2 sessions per week over 6 weeks	Non-treatment control. Offered the treatment upon study completion.	1) Physical activity 2) Salivary cortisol 3) Blood cytokines 3) Body composition 4) Peak strength 5) Girth	6 wks	1 day 3 sample points (On awakening, 1 hour and 2 hours after waking)	Time-specific cortisol levels	There was a significant decrease in waking cortisol concentration in the exercise group.	There was a significant increase in lean tissue mass and upper and lower body strength in the exercise group.	Moderate
Emery et al. 2005 (106)	To evaluate the influence of exercise on wound healing and whether the stress system is a possible mechanism by which it acts	28 sedentary older adults (55-77 years) who agreed to a punch biopsy	1 hour exercise sessions for 3 days per week over 3 months	Non-exercise control	1) Wound healing 2) Exercise endurance 3) Salivary cortisol 4) Perceived stress	12 wks	2 days 4 sample points (no further details)	AUC	There was no significant difference between groups for salivary cortisol parameters	Wounds in the exercise group healed significantly faster than wounds in the control group	Moderate
Jaggers et al. 2014 (107)	To examine the effects of combined aerobic and resistance training on mood, stress and symptoms in HIV-infected adults	93 HIV-infected males and females Only 49 completed the study and a subsample of 20 provided saliva samples	Aerobic (30 min) and resistance (20 min) training sessions twice weekly over 6 wks	Attended the centre, following the same time schedule, and engaged in sedentary activities.	1) Mood 2) HIV symptom distress 3) Perceived stress 4) Salivary cortisol	6 wks	1 day 3 sample points (on awakening, 1 hour and 2 hours after waking)	Morning AUC	There was a significant decrease in morning AUC in the treatment group, which was not observed in the control.	There was a significant decrease in mood disturbance in the intervention group, which was not observed in the control. There were no significant changes in the other clinical measures.	Moderate
Lieverse et al. 2011 (108)	To determine the efficacy of bright light treatment (BLT) in elderly patients with major depressive disorder	89 outpatients ≥ 60 years with major depressive disorder	BLT (bright pale blue light) delivered for 1 hour each morning over 3 weeks	Placebo (dim red light) delivered each morning over 3 weeks	1) Depression scale (primary) 2) Urinary cortisol 3) Salivary cortisol 4) Salivary melatonin 5) Sleep actigraphy parameters	6 wks	1 day 8 sample points: 4 morning samples at 30 min intervals, starting 30 min after awakening 4 evening samples at hourly intervals starting 4 hours before bedtime	1) Morning cortisol AUC (9am-1pm) 2) Evening cortisol AUC (5 to 9pm)	There was a significant decrease in evening cortisol AUC in the BLT group in comparison with the placebo group at 6 weeks but not at 3 weeks.	There was a significant improvement in depression scores in the BLT group in comparison with the placebo group, at 3 and 6 weeks	High
Saxton et al. 2014 (109)	To investigate the effects of a pragmatic lifestyle intervention on indices of psychological health status, HPA axis regulation and immune function in overweight women recovering from early-stage breast cancer. Other outcomes reported in Scott et al. 2013 (117)	85 women treated for early-stage breast cancer 3-18 months previously, with a body mass index > 25kg/m^2	A pragmatic lifestyle intervention was implemented over 24 wks, which involved 3 supervised exercise sessions per week along with an individually tailored hypocaloric healthy eating programme.	Non-treatment control: offered exercise sessions and dietary advice upon completion of the study	1) Depression 2) Perceived stress 3) Salivary cortisol 4) Plasma cortisol 5) Plasma cytokines/ lymphocyte profiling	24 wks	3 days 4 sample points (8am, 12 noon, 5pm, 9pm)	AUC: calculated using trapezoidal rule	There was a significant increase in cortisol AUC in the intervention group in comparison with the control group at 24 wks, attributable to a higher morning cortisol in the intervention group post-intervention.	Depressive symptoms were significantly reduced in the intervention group in comparison with the control group at 24 wks. However, there was no significant reduction in perceived stress in the intervention group in comparison with the control.	High
Scherder et al. 2003 (111)	To investigate the effects of low-frequency cranial electrostimulation on rest-activity rhythm and salivary cortisol levels in patients with probable Alzheimer's Disease	16 older adults living in residential homes with a probable diagnosis of Alzheimer's Disease	Cranial electrical stimulation (low-frequency) administered for 30min per day, 5 days per week, over 6 weeks	Placebo: apparatus attached but no current administered; same schedule	1) Rest-activity rhythm (Actigraphy) 2) Salivary cortisol	12 wks	1 day 9 sample points (Measures obtained at irregular times from 8am to 10pm)	A multi-level model was constructed resulting in a mean cortisol curve for each group at each follow-up point	There was no significant time x group interaction effect for salivary cortisol outcomes	There was no significant time x group interaction for rest-activity outcomes	High
Scherder et al. 2006 (112)	To investigate the effects of high-frequency cranial electrostimulation on rest-activity rhythm and salivary cortisol levels in patients with probable Alzheimer's Disease	20 older adults living in residential homes with a probable diagnosis of Alzheimer's Disease	Cranial electrical stimulation (high-frequency) administered for 30 min per day, 5 days per week, over 6 weeks	Placebo: apparatus attached but no current administered	1) Rest activity rhythm (Actigraphy) 2) Salivary cortisol	12 wks	1 day 9 sample points (Measures obtained at irregular times from 7.28am to 11pm)	A multi-level model was constructed resulting in a mean cortisol curve for each group at each follow-up point	There was no significant time x group interaction effect for salivary cortisol outcomes	There was no significant time x group interaction for rest-activity outcomes	High
Tam et al. 2014 (113)	To test the hypothesis that moderate caloric restriction (CR) by diet or a combination of diet and exercise would alter morning and diurnal cortisol. Metabolic outcomes of RCT separately reported in Heilbronn et al. 2006 (121)	34 young overweight adults	Calorie restriction (25% reduction in energy intake) OR Calorie restriction group with exercise (12.5% reduction in energy intake + 12.5% increase in exercise)	Control: weight maintenance group	1) Body composition 2) Blood lipids, glucose, leptin, thyroxine, IGF-1 3) Insulin sensitivity and acute insulin response to glucose 4) Salivary cortisol	24 wks	1 day 8 sample points (8.00, 8.30, 11.00, 11.30, 12.30, 13.00, 16.00, 16.30)	1) Morning cortisol: average of 8.00 and 8.30 cortisol 2) Mean diurnal cortisol ('diurnal cortisol'): average of 8 samples across the day	There were no significant group x time interaction effects for morning cortisol or mean diurnal cortisol	Body weight, fat mass and visceral adipose tissue were significantly reduced from baseline in both intervention groups in comparison with the control. There were also significant changes in some of the biochemical measures in the intervention groups in comparison with the control.	Moderate

Shaded rows represent studies with agreement between clinical and cortisol findings

[a] Describes the number of consecutive days of sampling and the number of samples per day per time-point

[b] Assessed using Gough's Framework

[c] Refers to the number and type of participants randomised

[d] Abbreviations used for cortisol summary measures: CAR, cortisol awakening response; AUC, area under the curve.

with many scoring 'high' in the quality and relevance assessment. In addition, there is robust evidence that flatter diurnal cortisol slopes in this population are associated with shorter survival [3], pointing to the plausibility of the HPA axis as a potential therapeutic target.

Four studies evaluated the effects of different psychosocial interventions in patients with breast cancer. There was inconsistency between clinical and cortisol findings in two of these studies. In addition, the types of cortisol parameters measured, along with their patterns of change, were not uniform across studies. Two studies, one evaluating mindfulness-based cancer recovery and supportive expressive therapy [25], and the other evaluating mind-body-spirit therapy [34], found that the diurnal slope remained unchanged in the treatments groups but that it became significantly flatter in the control group, suggesting that these treatments had a buffering effect on the HPA axis. This finding corresponded with clinical findings in only one of the studies [25], however. In another study evaluating mind-body-spirit therapy, there was no change in the diurnal slope in either the treatment or the control groups, but the area under the curve decreased in the treatment group, mirroring a reduction in symptoms in this group [27]. A similar intervention (relaxation and visualisation therapy) had no effect on the area under the curve in another study, however, despite a reduction in symptoms [48].

Four studies evaluated yoga in patients with prior or current breast cancer. Banasik et al. [86] found no significant change in the diurnal slope, despite an improvement in symptoms; however, absolute cortisol levels (morning and evening) were found to be significantly reduced. Bower et al. [88] found that there was no change in the diurnal slope or the area under the curve, despite an improvement in symptoms. A further study found that, along with symptom improvement, the diurnal slope became significantly steeper over 6 weeks of treatment relative to comparator groups; this finding lost significance, however, after missing values were addressed using a multiple imputation technique [92]. Finally, Raghavendra et al. [99] found a reduction in symptoms, 6.00 am cortisol concentration and 'pooled mean diurnal cortisol' after yoga but no change in other time-specific cortisol levels or in the area under the curve. All four of these studies provide support for yoga in relation to symptom improvement, but there was no consistent pattern of change in cortisol parameters across studies. The problem of interpreting cortisol findings was further compounded by the use of a range of different cortisol parameters both within and between studies.

Discussion

This systematic review characterises the types of RCTs which have used salivary diurnal cortisol as an outcome measure for the evaluation of health and behavioural interventions and details the salivary diurnal cortisol methodology and findings therein. To the authors' knowledge, this is the first systematic review of this kind.

The review highlights the increasing use of salivary diurnal cortisol as an outcome measure in RCTs, particularly since 2012. The majority of these RCTs have evaluated psychosocial or complementary therapy interventions in a wide range of populations, ranging from healthy volunteers to patients with cancer. With regard to salivary diurnal cortisol methodology and outcomes, the review has identified the following findings: (1) many of the RCTs screened did not use diurnal measures of salivary cortisol, (2) the majority of RCTs measuring diurnal cortisol collected samples over 1 day only, (3) there is wide heterogeneity across studies in relation to sampling schedules, (4) there is wide heterogeneity in relation to the cortisol profile parameter chosen for analysis, with a large proportion of studies failing to analyse diurnal rhythm parameters, and (5) interpretation of cortisol findings within and between RCTs is challenging due to the use of different parameters in different studies, varying cortisol change patterns across studies and high levels of inconsistency between cortisol and clinical findings. These review findings are discussed below, and based on these findings, recommendations are made for the future incorporation of salivary diurnal cortisol into RCTs.

Many of the RCTs Screened Did Not Measure Diurnal Cortisol Profiles

During the selection process, after excluding articles for other reasons, 87 of the remaining 175 RCTs (50 %) were excluded because they did not measure diurnal profiles of salivary cortisol, despite including it as an outcome measure. In many cases, a single salivary cortisol sample was obtained before and after an intervention, either on the same day as the intervention was received or on a different day. It is long established that single measures of basal cortisol, even if collected at the same time each day, have very low reliability due to significant intra-individual variability [10]. For example, Coste et al. [115] demonstrated that when a single salivary cortisol sample was collected at 8 am at three time-points over 5 weeks the intra-class correlation coefficient (r) was as low as 0.18. In addition, single measures of basal cortisol have very low diagnostic utility, due to wide inter-individual variation, with normal ranges overlapping with abnormal ranges [10]. It is surprising that despite this knowledge, which dates back to 1994, many researchers are still using single measures of cortisol as biomarkers within their trials. This practice has the potential consequence of generating false positive results in response to interventions, particularly within small pilot studies. Apart from this, the use of unreliable measures within RCTs is a waste of limited financial resources.

The Majority of RCTs Collected Saliva Samples Over 1 Day Only

This review found that 57 out of the 78 included studies (73.1 %) collected diurnal samples over 1 day only. When the cortisol awakening response is measured on a single day, it has been shown to be highly influenced by situational or state factors, but reliable cortisol awakening response measurements have been obtained when the cortisol awakening response is averaged over at least 2 days [correlation coefficient (*r*) between 2-day pairs=0.7] [116]. Significant day-to-day variation has also been observed for the diurnal slope, where the frequencies of inconsistent diurnal patterns over 2–3 days were observed to be 31 % in one sample [117] and 43 % in another [118]. For these reasons, it is recommended that salivary cortisol is collected over more than 1 day in order to capture stable characteristics [11]. In fact, it has been suggested that it is better to add more consecutive days to the protocol than more samples per day in order to improve the reliability of diurnal rhythm assessment [11]. Measures of low reliability inevitably result in low validity. Therefore, the predominant lack of consecutive day sampling observed in this review necessitates that cortisol outcomes within the included RCTs be interpreted with caution. Indeed, the low level of agreement between cortisol and clinical findings across the RCTs (50 %) might well be explained by the low reliability of the diurnal profiles measured within these RCTs.

There Is Wide Heterogeneity Across Studies in Relation to Sampling Schedules

Within the included studies, the number of samples collected per day ranged from two samples per day to nine samples per day, the median being 4 (IQR 3–5) samples per day. The wide variation in protocols highlights the fact that there really is no consensus regarding the optimal frequency of sampling per day. Some of this is probably due to lack of knowledge in the field of stress research about the impact of different sampling schedules on diurnal profile validity. In their review, Adam and Kumari [11] referred to unpublished data of theirs which demonstrated that a 2-point diurnal slope (morning and evening) correlates extremely well with a 6–7 point slope (correlation coefficient=0.94), suggesting that delineating the curve more precisely does not significantly improve the accuracy of important summary measures such as the diurnal slope. Whilst this data suggests that a minimal protocol of 2 collection points per day can yield a meaningful diurnal slope, further validation studies are needed to confirm this and to investigate the maximum number of samples per day beyond which sampling would be wasteful and unnecessarily burdensome. Considering that 21.8 % of RCTs in this review used a schedule of 6 or more sampling points per day, this area of uncertainty needs to be addressed promptly.

In addition to variation in sample number per day, sampling times also differed between studies. For example, only 62.8 % of studies included an awakening sample. As a result, in many studies, the cortisol profile was anchored to clock time rather than waking time, which is suboptimal practice. Whilst it is preferable to calculate the diurnal slope using values outside of the awakening period, the cortisol profile from which it is derived should be anchored to waking time. The rationale for this is well documented, the practice being based upon the fact that waking up activates a burst of cortisol pulses which serve to 'synchronise' the circadian rhythm of the HPA axis [10]. Furthermore, it has been shown that diurnal cortisol rhythms are influenced primarily by personal sleep-wake cycles, predominantly wake time, rather than by dark-light cycles [11, 119].

There Is Wide Heterogeneity Across RCTs in Relation to the Cortisol Profile Parameters Analysed

Despite rhythm parameters being most robustly linked with health outcomes, it was surprising that only half of the RCTs included a marker of diurnal rhythm by measuring either the cortisol awakening response or the diurnal slope or by multilevel modelling techniques. Interestingly, these measures were most commonly used in studies measuring psychosocial interventions where the prevalence was 66.7 %. Lack of measurement of these parameters within RCTs suggests little awareness of the complexities of HPA axis regulation and function amongst clinical trialists and points to the need to better translate psychoneuroendocrinological knowledge into clinical trials research. Better collaboration between basic scientists, in the field of psychoneuroendocrinology, and clinical trialists, with an interest in salivary cortisol as a biomarker, may help ameliorate this problem. The higher prevalence of rhythm parameters in psychosocial intervention studies probably reflects the already well-established relationship between the disciplines of clinical psychology and psychoneuroendocrinology, owing to the natural proximity of the fields.

It was uncommon for RCTs using rhythm parameters to measure both the cortisol awakening response and the diurnal decline (17.9 %) and even more uncommon for RCTs to measure all three parameters recommended by Adam and Kumari [11] in epidemiological studies (the cortisol awakening response, diurnal decline and area under the curve) (2.6 %). In the context of RCTs, it would appear sensible to measure all three parameters in order to robustly assess HPA axis activity, particularly in the context of an exploratory study. In particular, it would make sense to measure both the cortisol awakening response and the diurnal slope given that they are believed to be regulated independently, representing different aspects of HPA axis function [14–16, 120]. Failure to measure all parameters within an RCT may result in false negative findings in relation to HPA axis function and may partially explain the low agreement between cortisol and clinical findings in

this review. On the other hand, where all three parameters are used, it would be important to guard against the practice of multiple testing and post hoc hypotheses. With this in mind, it would be wise for RCTs to state the primary HPA axis parameter of interest, including its hypothesised direction of change, in the protocol prior to commencing the study.

Studies which did not use rhythm parameters relied on 'area under the curve' measures, mean diurnal cortisol measures or absolute cortisol measures at specific times of the day to measure HPA axis activity. There are several disadvantages to these approaches. In relation to the area under the curve, whilst it is a useful measure of overall cortisol exposure, it is difficult to interpret its meaning without a co-measure of diurnal rhythm. This is because both hypocortisolism and hypercortisolism have been linked with chronic stress and its health implications [6], such that the amount of cortisol in the system has become a less discerning instrument for measuring clinically relevant stress. For the same reason, measurement of the mean cortisol level across the day has similar limitations. The measurement of absolute cortisol levels at specific times in the day and the reporting of within- or between-group pre-post changes for each specific time represented another approach. Due to the separate analysis for each sample point, however, this is no different, in many respects, to obtaining multiple single cortisol measures, with each cortisol measure having low reliability. In addition, with this approach, study findings are likely to become contaminated by false positive findings due to the inevitable consequences of multiple analyses. Thus, the RCTs which used this approach need to be interpreted with caution.

Interpretation of Cortisol Findings Within and Between RCTs Is Challenging

Analysis of the cortisol findings for psychosocial intervention studies and complementary therapy studies in the breast cancer population demonstrated the challenge of interpreting cortisol findings both within and between RCTs. This population, as a whole, is believed to have a flatter diurnal slope than a healthy population, and assuming this relates to chronic stress, one would expect a stress-relieving intervention to result in a steeper slope. No study was able to robustly demonstrate this, however. Instead, the findings of two studies [25, 34] suggested that the diurnal slopes would have become progressively flatter without intervention, due to a pattern of progressively flattening slopes in the control groups. In the absence of longitudinal studies of HPA axis regulation over weeks, months and years, it is not possible to firmly draw this conclusion, however. The findings of these studies illustrate the importance of understanding the natural history of HPA axis regulation within the target population before evaluating the effects of interventions in RCTs. Without understanding this, it is not possible to form a priori hypotheses regarding the direction of change in a cortisol parameter in response to an intervention. It may well be that stress-relieving interventions serve to 'stabilise' the HPA axis and protect it from further dysregulation, but this can only occur in a population within which unstable function or progressive HPA axis dysregulation exists.

Within studies, there was a high rate of inconsistency between clinical and cortisol findings, with cortisol findings supporting clinical findings in only 50 % of studies. In many cases, there was a significant clinical response to the intervention but no cortisol response. This may have occurred for a wide variety of reasons. The lack of cortisol response most likely reflects flaws in the cortisol measurement methodology, as discussed above. Lack of engagement of the HPA axis by the intervention is also a possibility, indicating that the intervention works by an alternative mechanism. Alternatively, it is possible that 'target engagement' did occur but that the impact on cortisol was obscured by the effects of other pathways and systems. Finally, another reason for lack of effect may be the absence of HPA axis dysregulation at baseline in the sample population receiving the intervention. For many of the studies, the prevalence or degree of HPA axis dysregulation in the population at baseline was not clear; this would need to be high in order to observe an improvement in HPA axis function after a therapeutic intervention, particularly in the presence of many confounders, as would be common in a patient population.

In a minority of cases of disagreement between findings, positive cortisol findings occurred in the absence of clinical findings. This may represent a time lag between HPA axis restoration and clinical improvement, with HPA axis restoration temporarily preceding clinical improvement. It may also result from the use of inappropriate clinical outcome measures, resulting in false negative clinical findings. Alternatively, however, this disagreement may reflect lack of reliability in the cortisol measure, resulting in false positive cortisol findings. Low reliability is highly likely for the studies included in this review, given the high prevalence of 1-day saliva collection protocols. Along with short-term reliability issues, the long-term stability of diurnal cortisol measures is also likely to impact on results, and there is a growing literature to suggest that this is low [17]. For example, Ross et al. [121] analysed visit-to-visit cortisol stability for the diurnal cortisol profile in a population of 46 healthy adults, providing 3-day cortisol profile samples at 2.5 monthly visits over 8 months and found only low-modest intra-class correlation coefficients (ICC) for the cortisol awakening response (ICC 0.219), the diurnal slope (ICC 0.473) and the area under the curve (ICC 0.556), with even lower stability at the individual level.

Due to heterogeneity across studies in relation to the HPA axis parameters measured, it was difficult to explore the timeframe over which a given parameter might be expected to change following an intervention, which was an important review aim. Nevertheless, the review has shed some light on this area of uncertainty by identifying that changes in parameters occurred at a median of 6 weeks from baseline (IQR 4–12). Though this finding needs to be interpreted with caution,

given the wide heterogeneity across studies in relation to parameters used, intervention duration and follow-up schedule, it at least provides a guide for the design of future RCTs in relation to the optimal timing of the primary endpoint and the length of the follow-up period.

Recommendations for the Future

In view of the increasing use of salivary diurnal cortisol as a biomarker within RCTs and the marked heterogeneity in practices and findings across studies, there is a clear need for guidance on how best to incorporate this biomarker into RCTs, in order to prevent unnecessary research costs and participant burden. The high level of inconsistency between clinical and cortisol findings and the difficulty in interpreting cortisol change patterns suggests a need for further validation studies. There is also a need for greater precision in diurnal cortisol measurement. Furthermore, there is a need for greater uniformity in the collection and analysis of cortisol, to allow findings to be compared across studies. We have summarised recommendations towards the achievement of these goals in box 1.

Box 1. Recommendations for the use of salivary diurnal cortisol as a biomarker within randomised controlled trials.

A. Decide whether or not it will be a useful biomarker:

- Establish the prevalence and pattern of HPA axis dysregulation in the target population
- Establish the longitudinal change in the pattern of HPA axis activity over the planned time-frame for the RCT
- Establish the construct validity of HPA axis parameters against relevant clinical measures
- Be able to form an a priori hypothesis regarding the expected direction of change in at least one HPA axis parameter in response to the experimental intervention

B. Optimise the reliability and validity of the cortisol measure:

- Collect salivary cortisol over at least 2 days both before and at least once after the intervention
- Collect all samples with reference to awakening time rather than a clock time
- Ideally, include enough sample points in the day to analyse all three parameters (the cortisol awakening response, the diurnal slope and the area under the curve), to provide a full picture of HPA axis activity, unless there are valid reasons to exclude some components (e.g. expected high non-compliance rates for the cortisol awakening response)

C. Optimise the ability to interpret and compare clinical trial findings:

- Choose one cortisol parameter as the primary cortisol outcome measure (e.g. cortisol awakening response **or** area under the curve **or** diurnal slope) in advance of the study, linking this with the a priori hypothesis; this should be identified as the primary parameter in the protocol and the published report.
- Include all other cortisol parameters as secondary outcome measures

Limitations

A number of methodological limitations need to be borne in mind when interpreting the findings of this review. Firstly, though we searched six electronic databases using sensitive search terms for RCTs and salivary cortisol, we excluded animal studies from three databases (MEDLINE, EMBASE and AMED) using the exploded term, which, we realised in retrospect, may have inadvertently eliminated some human studies. Having assessed the impact of this on the MEDLINE results, however, we are confident that this has not had a significant impact on the overall yield of eligible studies due to the substantial overlap of these databases with each other and with both the Cochrane Central Register of Controlled Trials and PsychINFO. Secondly, we did not perform a supplementary manual literature search. Whilst this strategy may have improved our yield of RCTs, given the very broad search criteria in relation to type of intervention and population, it was not feasible to devise a comprehensive manual search strategy without biasing the study selection process.

Conclusions

This review systematically maps the literature which reports on the use of salivary diurnal cortisol as an outcome measure within RCTs. It demonstrates that there is wide heterogeneity across RCTs in the methodology of salivary cortisol collection, and in the profile parameters analysed. Furthermore, it has demonstrated that such methodological heterogeneity has consequences for both the internal validity of individual trials and the ability to compare and synthesise results across trials of similar interventions. As such, it highlights a need for better validation of this measure, more reliable approaches to measurement and the need for greater collaboration between the disciplines of psychoneuroendocrinology and applied science disciplines such as medicine, psychology and nursing, with a view to better and more prompt translation of basic science knowledge about HPA axis measurement into clinical trials research.

Acknowledgments Dr. Ryan is funded by the National Institute for Health Research (NIHR) Doctoral Research Fellowship Programme Award in the UK. The views expressed are those of the authors and not necessarily those of the National Health Service (NHS), the NIHR or the Department of Health.

Compliance with Ethical Standards

Authors' Statement of Conflict of Interest and Adherence to Ethical Standards Authors Richella Ryan, Sara Booth, Anna Spathis, Sarah Mollart and Angela Clow declare that they have no conflict of interest. All procedures, including the informed consent process, were conducted in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000.

Appendixes

Appendix A Search strategy for systematic review.

MEDLINE (1980 to 21 May 2015):

1. cortisol.ti,ab;
2. saliva*.af;
3.1 AND 2;
4. HYDROCORTISONE/;
5. SALIVA/;
6.4 AND 5;
7. 'randomized controlled trial'.pt;
8. 'controlled clinical trial'.pt;
9. 'randomized'.ab;
10. placebo.ab;
11. randomly.ab;
12. trial.ab;
13. groups.ab;
14. 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13;
15. expANIMALS/
16. 14 NOT 15
17. 3 OR 6
18. 16 AND 17;180 results

CINAHL (1981 to 21 May 2015):

1. cortisol.ti,ab;
2. saliva*".af;
3. HYDROCORTISONE/;
4. SALIVA/;
5. 1 AND 2;
6. 3 AND 4;
7. 5 OR 6;
8. 'randomized controlled trial'.pt;
9. 'controlled clinical trial'.pt;
10. 'clinical trial'.pt;
11. RANDOMIZED CONTROLLED TRIALS/OR CLINICAL TRIALS/OR INTERVENTION TRIALS/;
12. randomized.ab;
13. placebo.ab;
14. randomly.ab;
15. trial.ab;
16. groups.ab;
17. 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16;
18. 7 AND 17; 338 results

PsychINFO (1806 to 21 May 2015):

1 cortisol.ti,ab;
2. saliva*.af;
3. 1 AND 2;
4. HYDROCORTISONE/;
5. SALIVA/;
6. 4 AND 5;
7. 3 OR 6;
8. 'randomised controlled trial'.pt
9. 'controlled clinical trial'.pt
10. 'clinical trial'.pt
11. TREATMENT EFFECTIVENESS EVALUATION/OR CLINICAL TRIALS/;
12. randomized.ab;
13. placebo.ab;
14. randomly.ab;
15. trial.ab;
16. groups.ab;
17. 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16;
18. 7 AND 17; 879 results

AMED (1985 to 21 May 2015)

1 cortisol.ti,ab;
2. saliva*.af;
3. 1 AND 2;
4. HYDROCORTISONE/;
5. SALIVA/;
6. 4 AND 5;
8. 'randomized controlled trial'.pt
9. 'controlled clinical trial'.pt
10. 'clinical trial'.pt
11. CLINICAL TRIALS/OR RANDOMIZED CONTROLLED TRIALS;
12. randomized.ab;
13. placebo.ab;
14. randomly.ab;
15. trial.ab;
16. groups.ab;
17. 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16;
18. expANIMALS/
19. 17 NOT 18
20. 3 OR 6
21. 19 AND 20; 11 results

EMBASE (1974 to 21 May 2015)

1 cortisol.ti,ab;
2. saliva*.af;
3. 1 AND 2;
4. HYDROCORTISONE/;
5. SALIVA/OR SALIVA ANALYSIS/OR SALIVA COLLECTOR/
6. 4 AND 5
7. 3 OR 6
8. 'CLINICAL TRIAL (topic)'/OR CONTROLLED CLINICAL TRIAL/ OR 'CONTROLLED CLINICAL TRIAL (topic)'/OR 'PHASE 1 CLINICAL TRIAL (topic)'/OR 'PHASE 2 CLINICAL TRIAL (topic)'/OR 'PHASE 3 CLINICAL TRIAL (topic)'/OP 'PHASE 4 CLINICAL TRIAL (topic)'/OR 'RANDOMIZED CONTROLLED TRIAL (topic)'/
9. randomized.ab;
10. placebo.ab;
11. randomly.ab;
12. trial.ab;
13. groups.ab;

14. 8 OR 9 OR 10 OR 11 OR 12 OR 13

15. expANIMAL/

16. 14 NOT 15

17. 16 AND 7;109 results

Cochrane Central Register of Controlled Trials (up to 21 May 2015)

1. cortisol
2. saliva*
3. Mesh descriptor: hydrocortisone
4. Mesh descriptor: saliva
5. (1 AND 2) OR (3 AND 4); 857 results

Appendix B Quality and relevance assessment using Gough's framework.

Study ID	Weight of Evidence A (Generic quality of execution of study)	Weight of Evidence B (Appropriateness of the study design to the review aim)	Weight of Evidence C (Focus of the study content relative to the review aim)	Weight of Evidence D (Overall quality and relevance grade)
Banasik et al. 2011 [86]	Low	High	High	Moderate
Barbadoro et al. 2013[60]	Moderate	High	Moderate	Moderate
Bergen-Cico et al. 2014 [22]	Moderate	High	High	High
Billhult et al. 2008 [87]	Moderate	High	Low	Moderate
Boelens et al. 2009 [104]	Low	High	Moderate	Moderate
Bormann et al. 2009 [23]	Moderate	High	Moderate	Moderate
Bougea et al. 2013 [24]	Low	High	Moderate	Moderate
Bower et al. 2014 [88]	Moderate	High	High	High
Camfield et al. 2013 [61]	Low	High	High	Moderate
Campo et al. 2015 [90]	Low	High	High	Moderate
Carlson et al. 2013 [25]	High	High	High	High
Cash et al. 2014 [26]	Moderate	High	High	High
Chaborski et al. 2015 [62]	Moderate	High	Moderate	Moderate
Chan et al. 2006 [27]	Moderate	High	Moderate	Moderate
Chandwani et al. 2014 [92]	Moderate	High	High	High
Chen et al. 2013 [93]	Moderate	High	High	High
Corey et al. 2014 [94]	Moderate	High	High	High
Danucalov et al. 2013 [96]	Moderate	High	Moderate	Moderate
Delle Chiaie et al. 2012 [28]	Low	High	High	Moderate
Deuschle et al. 2003 [63]	Moderate	Low	Moderate	Moderate
Dudgeon et al. 2012 [105]	Low	High	Moderate	Moderate
Eijsbouts et al. 2008 [64]	Moderate	High	Moderate	Moderate
Emery et al. 2005 [106]	Low	High	Moderate	Moderate
Feicht al. 2013 [29]	Low	High	Moderate	Moderate
Gaab et al. 2006 [30]	Moderate	High	High	High
Garrison and Chambliss, 2006 [65] & Kalman et al. 2008 [66]	Low	High	Moderate	Moderate
Gex-Fabry et al. 2012 [31]	High	High	High	High
Hellweg et al. 2008 [67]	Moderate	High	Moderate	Moderate
Hinkelmann et al. 2012 [68]	Moderate	Moderate	High	Moderate
Hodgson and Lafferty, 2012 [97]	Moderate	High	Moderate	Moderate
Holt-Lunstad et al. 2008 [32]	Moderate	High	Moderate	Moderate
Hsiao et al. 2011 [33]	Moderate	High	Moderate	Moderate
Hsiao et al. 2012 [34]	Low	High	High	Moderate
Hsiao et al. 2014 [35]	Low	High	High	Moderate
Huang et al. 2012 [98]	Low	High	High	Moderate
Jaggers et al. 2014 [107]	Low	High	Moderate	Moderate
Jensen et al. 2012 [36]	Moderate	High	Low	Moderate
Klatt et al. 2009 [37]	Low	High	Moderate	Moderate

(continued)

Knorr et al. 2012 [70]	High	High	High	High
Krajewski et al. 2010 [38]	Low	Moderate	Moderate	Moderate
Lenze et al. 2011 [71] & Lenze et al. 2012 [72]	Low	High	Moderate	Moderate
Letourneau et al. 2011 [39]	Low	High	Moderate	Moderate
Lindh-Astrand et al. 2013 [41]	Moderate	High	Moderate	Moderate
Lipschitz et al. 2013 [42]	Moderate	High	High	High
Lieverse et al. 2011 [108]	High	High	Moderate	High
Limm et al. 2011 [40]	High	High	Moderate	High
Lok et al. 2012 [44]	High	High	Moderate	High
Lopresti et al. 2015 [73]	Moderate	High	Moderate	Moderate
Mocking et al. 2014 [75]	Moderate	High	Moderate	Moderate
Nickel, C et al. 2007 [46] & Nickel, M.K., 2007 [47]	Moderate	High	Moderate	Moderate
Nonino-Borges et al. 2007 [76]	Moderate	High	Moderate	Moderate
Nunes et al. 2007 [48]	Moderate	High	High	High
Oken et al. 2010 [49]	Moderate	High	Moderate	Moderate
Pacella et al. 2014 [50]	Moderate	Moderate	Moderate	Moderate
Plag et al. 2014 [51]	Moderate	High	Moderate	Moderate
Raghavendra et al. 2009 [99]	Low	High	High	Moderate
Richter et al. 2012 [53]	Low	High	High	Moderate
Ruhe et al. 2015 [77]	Moderate	Moderate	Moderate	Moderate
Saxton et al. 2014 [109]	Moderate	High	High	High
Scharnholz et al. 2010 [78]	Moderate	High	Moderate	Moderate
Scherder et al. 2003 [111]	Moderate	High	High	High
Scherder et al. 2006 [112]	Moderate	High	High	High
Schmidt et al. 2015 [79]	Moderate	High	High	High
Schneider et al. 2007 [100]	Low	High	High	Moderate
Schubert et al. 2011 [80]	Moderate	High	Moderate	Moderate
Sears et al. 2007 [54]	Low	High	Moderate	Moderate
Talbott et al. 2013a [81]	Low	High	Moderate	Moderate
Talbott et al. 2013b [82]	Low	High	High	Moderate
Tam et al. 2014 [113]	Moderate	High	Moderate	Moderate
Taylor et al. 2009 [55]	Moderate	High	High	High
Tornhage et al. 2013 [102]	Low	Moderate	Moderate	Moderate
Tucker et al. 2004 [83]	Moderate	High	Moderate	Moderate
Urizar and Munoz, 2011 [56]	Moderate	High	Moderate	Moderate
Walsh et al. 2006 [84]	Moderate	Moderate	Moderate	Moderate
Wilcox et al. 2014 [57]	Moderate	High	Moderate	Moderate
Witbracht et al. 2013 [85]	Moderate	High	Moderate	Moderate
Woods et al. 2009[103]	Low	Moderate	Moderate	Moderate
Yang et al. 2009 [59]	Low	High	High	Moderate

References

1. de Kloet ER, Joels M, Holsboer F, Joëls M. Stress and the brain: from adaptation to disease. *Nat Rev Neurosci*. 2005; 6: 463-475.
2. Kumari M, Shipley M, Stafford M, Kivimaki M. Association of diurnal patterns in salivary cortisol with all-cause and cardiovascular mortality: Findings from the Whitehall II study. *J Clin Endocrinol Metab*. 2011; 96: 1478-1485.
3. Sephton SE, Sapolsky RM, Kraemer HC, Spiegel D. Diurnal cortisol rhythm as a predictor of breast cancer survival. *J Natl Cancer Inst.* 2000; 92: 994-1000.
4. Sephton SE, Lush E, Dedert EA, et al. Diurnal cortisol rhythm as a predictor of lung cancer survival. *Brain Behav Immun*. 2013; 30(Suppl): S163-S170.
5. Kumari M, Badrick E, Sacker A, et al. Identifying patterns in cortisol secretion in an older population. Findings from the Whitehall II study. *Psychoneuroendocrine*. 2010; 35: 1091-1099.
6. Fries E, Hesse J, Hellhammer J, Hellhammer DH. A new view on hypocortisolism. *Psychoneuroendocrine*. 2005; 30: 1010-1016.
7. McEwen BS. The neurobiology of stress: From serendipity to clinical relevance. *Brain Res*. 2000; 886: 172-189.
8. Kirschbarum C, Hellhammer DH. Salivary cortisol in psychoneuroendocrine research: Recent developments and applications. *Psychoneuroendocrine*. 1994; 19: 313-333.
9. Clow A. Cortisol as a biomarker of stress. *J Holist Healthc*. 2004; 1: 10-14.

10. Pruessner JC, Wolf OT, Hellhammer DH, et al. Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. *Life Sci*. 1997; 61: 2539-2549.
11. Adam EK, Kumari M. Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrine*. 2009; 34: 1423-1436.
12. Heim C, Ehlert U, Hellhammer DH. The potential role of hypocortisolism in the pathophysiology of stress-related bodily disorders. *Psychoneuroendocrine*. 2000; 25: 1-35.
13. De Kloet ER. Hormones and the stressed brain. *Ann N Y Acad Sci*. 2004; 1018: 1-15.
14. Clow A, Hucklebridge F, Stalder T, Evans P, Thorn L. The cortisol awakening response: More than a measure of HPA axis function. *Neurosci Biobehav Rev*. 2010; 35: 97-103.
15. Thorn L, Hucklebridge F, Esgate A, Evans P, Clow A. The effect of dawn simulation on the cortisol response to awakening in healthy participants. *Psychoneuroendocrine*. 2004; 29: 925-930.
16. Fries E, Dettenborn L, Kirschbaum C. The cortisol awakening response (CAR): Facts and future directions. *Int J Psychophysiol*. 2009; 72: 67-73.
17. Doane LD, Chen FR, Sladek MR, Van Lenten SA, Granger DA. Latent trait cortisol (LTC) levels: Reliability, validity, and stability. *Psychoneuroendocrine*. 2015; 55: 21-35.
18. Kraemer HC, Giese-Davis J, Yutsis M, et al. Design decisions to optimize reliability of daytime cortisol slopes in an older population. *Am J Geriatr Psychiatr*. 2006; 14: 325-333.
19. Smyth N, Clow A, Thorn L, Hucklebridge F, Evans P. Delays of 5–15 min between awakening and the start of saliva sampling matter in assessment of the cortisol awakening response. *Psychoneuroendocrine*. 2013; 38: 1476-1483.
20. Kovach CR, Woods DL, Devine EC, Logan BR, Raff H. Biobehavioral measures as outcomes: A cautionary tale. *Res Gerontol Nurs*. 2014; 7: 56-65.
21. Gough D. Weight of evidence: A framework for the appraisal of the quality and relevance of evidence. *Res Pap Educ*. 2007; 22: 213-228.
22. Bergen-Cico D, Possemato K, Pigeon W. Reductions in cortisol associated with primary care brief mindfulness program for veterans with PTSD. *Med Care*. 2014; 52: S25-S31.
23. Bormann E, Aschbacher K, Wetherell L, Roesch S, Redwine L. Effects of faith/assurance on cortisol levels are enhanced by a spiritual mantram intervention in adults with HIV: A randomized trial. *J Psychosom Res*. 2009; 66: 161-171.
24. Bougea AM, Spandideas N, Alexopoulos EC, et al. Effect of the emotional freedom technique on perceived stress, quality of life, and cortisol salivary levels in tension-type headache sufferers: A randomized controlled trial. *Explore (NY)*. 2013; 9: 91-99.
25. Carlson LE, Doll R, Stephen J, et al. Randomized controlled trial of mindfulness-based cancer recovery versus supportive expressive group therapy for distressed survivors of breast cancer. *J Clin Oncol*. 2013; 31: 3119-3126.
26. Cash E, Salmon P, Weissbecker I, et al. Mindfulness meditation alleviates fibromyalgia symptoms in women: Results of a randomized clinical trial. *Ann Behav Med*. 2015; 49: 319-330.
27. Chan CL, Ho RT, Lee PW, et al. A randomized controlled trial of psychosocial interventions using the psychophysiological framework for Chinese breast cancer patients. *J Psychosoc Oncol*. 2006; 24: 3-26.
28. Delle Chiaie R, Trabucchi G, Girardi N, et al. Group psychoeducation normalizes cortisol awakening response in stabilized bipolar patients under pharmacological maintenance treatment. *Psychother Psychosom*. 2013; 82: 264-266.
29. Feicht T, Wittmann M, Jose G, et al. Evaluation of a seven-week web-based happiness training to improve psychological well-being, reduce stress, and enhance mindfulness and flourishing: A randomized controlled occupational health study. *Evid Based Complement Alternat Med*. 2013; 2013: 676953. doi:10.1155/2013/676953.
30. Gaab J, Sonderegger L, Scherrer S, Ehlert U. Psychoneuroendocrine effects of cognitive-behavioral stress management in a naturalistic setting-a randomized controlled trial. *Psychoneuroendocrine*. 2006; 31: 428-438.
31. Gex-Fabry M, Jermann F, Kosel M, et al. Salivary cortisol profiles in patients remitted from recurrent depression: One-year follow-up of a mindfulness-based cognitive therapy trial. *J Psychiatr Res*. 2012; 46: 80-86.
32. Holt-Lunstad J, Birmingham A, Light C. Influence of a "warm touch" support enhancement intervention among married couples on ambulatory blood pressure, oxytocin, alpha amylase, and cortisol. *Psychosom Med*. 2008; 70: 976-985.
33. Hsiao FH, Jow GM, Lai YM, et al. The long-term effects of psychotherapy added to pharmacotherapy on morning to evening diurnal cortisol patterns in outpatients with major depression. *Psychother Psychosom*. 2011; 80: 166-172.
34. Hsiao FH, Jow GM, Kuo WH, et al. The effects of psychotherapy on psychological well-being and diurnal cortisol patterns in breast cancer survivors. *Psychother Psychosom*. 2012; 81: 173-182.
35. Hsiao FH, Lai YM, Chen YT, et al. Efficacy of psychotherapy on diurnal cortisol patterns and suicidal ideation in adjustment disorder with depressed mood. *Gen Hosp Psychiatry*. 2014; 36: 214-219.
36. Jensen CG, Vangkilde S, Frokjaer V, Hasselbalch SG. Mindfulness training affects attention–or is it attentional effort? *J Exp Psychol Gen*. 2012; 14: 106-123.
37. Klatt D, Buckworth J, Malarkey B. Effects of low-dose mindfulness-based stress reduction (MBSR-ld) on working adults. *Health Educ Behav*. 2009; 36: 601-614.
38. Krajewski J, Sauerland M, Wieland R. Relaxation-induced cortisol changes within lunch breaks—an experimental longitudinal worksite field study. *J Occup Organ Psychol*. 2011; 84: 382-394.
39. Letourneau N, Stewart M, Dennis CL, et al. Effect of home-based peer support on maternal-infant interactions among women with postpartum depression: A randomized, controlled trial. *Int J Ment Health Nurs*. 2011; 20: 345-357.
40. Limm H, Gündel H, Heinmüller M, et al. Stress management interventions in the workplace improve stress reactivity: A randomised controlled trial. *Occup Environ Med*. 2011; 68: 126-134.
41. Lindh-Åstrand L, Nedstrand E. Effects of applied relaxation on vasomotor symptoms in postmenopausal women: A randomized controlled trial. *Menopause*. 2013; 20: 401-409.
42. Lipschitz L, Kuhn R, Kinney Y, Donaldson W, Nakamura Y. Reduction in salivary α-amylase levels following a mind–body intervention in cancer survivors—an exploratory study. *Psychoneuroendocrine*. 2013; 38: 1521-1531.
43. Nakamura Y, Lipschitz DL, Kuhn R, Kinney AY, Donaldson GW. Investigating efficacy of two brief mind-body intervention programs for managing sleep disturbance in cancer survivors: A pilot randomized controlled trial. *J Cancer Surviv*. 2013; 7: 165-182.
44. Lok A, Mocking RJ, Ruhé HG, et al. Longitudinal hypothalamic-pituitary-adrenal axis trait and state effects in recurrent depression. *Psychoneuroendocrine*. 2012; 37: 892-902.
45. Bockting CLH, Schene AH, Spinhoven P, et al. Preventing relapse/recurrence in recurrent depression with cognitive therapy: A randomized controlled trial. *J Consult Clin Psychol*. 2005; 73: 647-657.
46. Nickel C, Tanca S, Kolowos S, et al. Men with chronic occupational stress benefit from behavioural/psycho-educational group training: A randomized, prospective, controlled trial. *Psychol Med*. 2007; 37: 1141-1149.
47. Nickel K. Behavioral/psychoeducational group training in therapy for overtaxed men. *J Psychosom Res*. 2007; 62: 597.

48. Nunes DF, Rodriguez AL, Silva Hoffmann F, et al. Relaxation and guided imagery program in patients with breast cancer undergoing radiotherapy is not associated with neuroimmunomodulatory effects. *J Psychosom Res*. 2007; 63: 647-655.
49. Oken BS, Fonareva I, Haas M, et al. Pilot controlled trial of mindfulness meditation and education for dementia caregivers. *J Altern Complement Med*. 2010; 16: 1031-1039.
50. Pacella L, Feeny N, Zoellner L, Delahanty L. The impact of PTSD treatment on the cortisol awakening response. *Depress Anxiety*. 2014; 31: 862-869.
51. Plag J, Gaudlitz K, Schumacher S, et al. Effect of combined cognitive-behavioural therapy and endurance training on cortisol and salivary alpha-amylase in panic disorder. *J Psychiatr Res*. 2014; 58: 12-19.
52. Gaudlitz K, Plag J, Dimeo F, Ströhle A. Aerobic exercise training facilitates the effectiveness of cognitive behavioral therapy in panic disorder. *Depress Anxiety*. 2015; 32: 221-228.
53. Richter J, Bittner A, Petrowski K, et al. Effects of an early intervention on perceived stress and diurnal cortisol in pregnant women with elevated stress, anxiety, and depressive symptomatology. *J Psychosom Obstet Gynaecol*. 2012; 33: 162-170.
54. Sears SF, Sowell LD, Kuhl EA, et al. The ICD shock and stress management program: A randomized trial of psychosocial treatment to optimize quality of life in ICD patients. *Pacing Clin Electrophysiol*. 2007; 30: 858-864.
55. Taylor CB, Conrad A, Wilhelm FH, et al. Does improving mood in depressed patients alter factors that may affect cardiovascular disease risk? *J Psychiatr Res*. 2009; 43: 1246-1252.
56. Urizar G, Muñoz F. Impact of a prenatal cognitive-behavioral stress management intervention on salivary cortisol levels in low-income mothers and their infants. *Psychoneuroendocrine*. 2011; 36: 1480-1494.
57. Wilcox RR, Granger DA, Szanton S, Clark F. Cortisol diurnal patterns, associations with depressive symptoms, and the impact of intervention in older adults: Results using modern robust methods aimed at dealing with low power due to violations of standard assumptions. *Horm Behav*. 2014; 65: 219-225.
58. Clark F, Jackson J, Carlson M, et al. Effectiveness of a lifestyle intervention in promoting the well-being of independently living older people: Results of the well elderly 2 randomised controlled trial. *J Epidemiol Community Health*. 2012; 66: 782-790.
59. Yang TT, Hsiao FH, Wang KC, et al. The effect of psychotherapy added to pharmacotherapy on cortisol responses in outpatients with major depressive disorder. *J Nerv Ment Dis*. 2009; 197: 401-407.
60. Barbadoro P, Annino I, Ponzio E, et al. Fish oil supplementation reduces cortisol basal levels and perceived stress: A randomized, placebo-controlled trial in abstinent alcoholics. *Mol Nutr Food Res*. 2013; 57: 1110-1114.
61. Camfield DA, Wetherell MA, Scholey AB, et al. The effects of multivitamin supplementation on diurnal cortisol secretion and perceived stress. *Nutrients*. 2013; 5: 4429-4450.
62. Chaborski K, Bitterlich N, Alteheld B, Parsi E, Metzner C. Placebo-controlled dietary intervention of stress-induced neurovegetative disorders with a specific amino acid composition: A pilot-study. *Nutr J*. 2015; 14: 43. doi:10.1186/s12937-015-0030-3.
63. Deuschle M, Hamann B, Meichel C, et al. Antidepressive treatment with amitriptyline and paroxetine: Effects on saliva cortisol concentrations. *J Clin Psychopharmacol*. 2003; 23: 201-205.
64. Eijsbouts AM, Kempers MJ, Kramer RS, et al. Effect of naproxen on the hypothalamic-pituitary-adrenal axis in healthy volunteers. *Br J Clin Pharmacol*. 2009; 67: 22-28.
65. Garrison R, Chambliss WG. Effect of a proprietary Magnolia and Phellodendron extract on weight management: A pilot, double-blind, placebo-controlled clinical trial. *Altern Ther Health Med*. 2006; 12: 50-55.
66. Kalman DS, Feldman S, Feldman R, et al. Effect of a proprietary Magnolia and Phellodendron extract on stress levels in healthy women: A pilot, double-blind, placebo-controlled clinical trial. *Nutr J*. 2008; 7: 11. doi:10.1186/1475-2891-7-11.
67. Hellweg R, Ziegenhorn A, Heuser I, Deuschle M. Serum concentrations of nerve growth factor and brain-derived neurotrophic factor in depressed patients before and after antidepressant treatment. *Pharmacopsychiatry*. 2008; 41: 66-71.
68. Hinkelmann K, Moritz S, Botzenhardt J, et al. Changes in cortisol secretion during antidepressive treatment and cognitive improvement in patients with major depression: A longitudinal study. *Psychoneuroendocrine*. 2012; 37: 685-692.
69. Otte C, Hinkelmann K, Moritz S, et al. Modulation of the mineralocorticoid receptor as add-on treatment in depression: A randomized, double-blind, placebo-controlled proof-of-concept study. *J Psychiatr Res*. 2010; 44: 339-346.
70. Knorr U, Vinberg M, Gether U, et al. The effect of escitalopram versus placebo on perceived stress and salivary cortisol in healthy first-degree relatives of patients with depression-A randomised trial. *Psychiatry Res*. 2012; 200: 354-360.
71. Lenze EJ, Mantella RC, Shi P, et al. Elevated cortisol in older adults with generalized anxiety disorder is reduced by treatment: A placebo-controlled evaluation of escitalopram. *Am J Geriatr Psychiatr*. 2011; 19: 482-490.
72. Lenze EJ, Dixon D, Mantella RC, et al. Treatment-related alteration of cortisol predicts change in neuropsychological function during acute treatment of late-life anxiety disorder. *Int J Geriatr Psychiatry*. 2012; 27: 454-463.
73. Lopresti AL, Maes M, Meddens MJM, et al. Curcumin and major depression: A randomised, double-blind, placebo-controlled trial investigating the potential of peripheral biomarkers to predict treatment response and antidepressant mechanisms of change. *Eur Neuropsychopharmacol*. 2015; 25: 38-50.
74. Lopresti AL, Maes M, Maker GL, Hood SD, Drummond PD. Curcumin for the treatment of major depression: A randomised, double-blind, placebo controlled study. *J Affect Disord*. 2014; 167: 368-375.
75. Mocking JT, Wever A, Pflanz CP, et al. Effects of short-term varenicline administration on cortisol in healthy, non-smoking adults: A randomized, double-blind, study. *Psychopharmacology*. 2014; 231: 143-148.
76. Nonino-Borges CB, Borges RM, Bavaresco M, et al. Influence of meal time on salivary circadian cortisol rhythms and weight loss in obese women. *Nutrition*. 2007; 23: 385-392.
77. Ruhé G, Khoenkhoen J, Ottenhof W, et al. Longitudinal effects of the SSRI paroxetine on salivary cortisol in major depressive disorder. *Psychoneuroendocrine*. 2015; 52: 261-271.
78. Scharnholz B, Weber-Hamann B, Lederbogen F, et al. Antidepressant treatment with mirtazapine, but not venlafaxine, lowers cortisol concentrations in saliva: A randomised open trial. *Psychiatry Res*. 2010; 177: 109-113.
79. Schmidt K, Cowen J, Harmer J, et al. Prebiotic intake reduces the waking cortisol response and alters emotional bias in healthy volunteers. *Psychopharmacology*. 2015; 232: 1793-1801.
80. Schubert M, Contreras C, Franz N, Hellhammer J. Milk-based phospholipids increase morning cortisol availability and improve memory in chronically stressed men. *Nutr Res*. 2011; 31: 413-420.
81. Talbott M, Talbott A, George A, Pugh M. Effect of Tongkat Ali on stress hormones and psychological mood state in moderately stressed subjects. *J Int Soc Sports Nutr*. 2013; 10: 28. doi:10.1186/1550-2783-10-28.
82. Talbott M, Talbott A, Pugh M. Effect of Magnolia officinalis and Phellodendron amurense (Relora®) on cortisol and psychological

mood state in moderately stressed subjects. *J Int Soc Sports Nutr*. 2013; 10: 37. doi:10.1186/1550-2783-10-37.
83. Tucker P, Ruwe D, Masters B, et al. Neuroimmune and cortisol changes in selective serotonin reuptake inhibitor and placebo treatment of chronic posttraumatic stress disorder. *Biol Psychiatry*. 2004; 56: 121-128.
84. Walsh JK, Randazzo AC, Stone K, et al. Tiagabine is associated with sustained attention during sleep restriction: Evidence for the value of slow-wave sleep enhancement? *Sleep*. 2006; 29: 433-443.
85. Witbracht MG, Loan M, Adams SH, Keim NL, Laugero KD. Dairy food consumption and meal-induced cortisol response interacted to influence weight loss in overweight women undergoing a 12-week, meal-controlled, weight loss intervention. *J Nutr*. 2013; 143: 46-52.
86. Banasik J, Williams H, Haberman M, Blank E, Bendel R. Effect of Iyengar yoga practice on fatigue and diurnal salivary cortisol concentration in breast cancer survivors. *J Am Acad Nurse Pract*. 2011; 23: 135-143.
87. Billhult A, Lindholm C, Gunnarsson R, Stener-Victorin E. The effect of massage on cellular immunity, endocrine and psychological factors in women with breast cancer-a randomized controlled clinical trial. *Auton Neurosci*. 2008; 140: 88-95.
88. Bower E, Greendale G, Crosswell D, et al. Yoga reduces inflammatory signaling in fatigued breast cancer survivors: A randomized controlled trial. *Psychoneuroendocrine*. 2014; 43: 20-29.
89. Bower JE, Garet D, Sternlieb B, et al. Yoga for persistent fatigue in breast cancer survivors: A randomized controlled trial. *Cancer*. 2012; 118: 3766-3775.
90. Campo A, Light C, O'Connor K, et al. Blood pressure, salivary cortisol, and inflammatory cytokine outcomes in senior female cancer survivors enrolled in a tai chi chih randomized controlled trial. *J Cancer Surviv*. 2015; 9: 115-125.
91. Campo RA, O'Connor K, Light KC, et al. Feasibility and acceptability of a Tai Chi Chih randomized controlled trial in senior female cancer survivors. *Integr Cancer Ther*. 2013; 12: 464-474.
92. Chandwani D, Perkins G, Nagendra HR, et al. Randomized, controlled trial of yoga in women with breast cancer undergoing radiotherapy. *J Clin Oncol*. 2014; 32: 1058-1066.
93. Chen Z, Meng Z, Milbury K, et al. Qigong improves quality of life in women undergoing radiotherapy for breast cancer: Results of a randomized controlled trial. *Cancer*. 2013; 119: 1690-1698.
94. Corey M, Epel E, Schembri M, et al. Effect of restorative yoga vs. stretching on diurnal cortisol dynamics and psychosocial outcomes in individuals with the metabolic syndrome: The PRYSMS randomized controlled trial. *Psychoneuroendocrine*. 2014; 49: 260-271.
95. Kanaya AM, Araneta MRG, Pawlowsky SB, et al. Restorative yoga and metabolic risk factors: The practicing restorative yoga vs. stretching for the metabolic syndrome (PRYSMS) randomized trial. *J Diabetes Complicat*. 2014; 28: 406-412.
96. Danucalov MAD, Kozasa EH, Ribas KT, et al. A yoga and compassion meditation program reduces stress in familial caregivers of Alzheimer's disease patients. *Evid Based Complement Alternat Med*. 2013; 2013: 513149. doi:10.1155/2013/513149.
97. Hodgson NA, Lafferty D. Reflexology versus swedish massage to reduce physiologic stress and pain and improve mood in nursing home residents with cancer: A pilot trial. *Evid Based Complement Alternat Med*. 2012; 2012: 456897. doi:10.1155/2012/456897.
98. Huang W, Taylor A, Howie J, Robinson N. Is the diurnal profile of salivary cortisol concentration a useful marker for measuring reported stress in acupuncture research? A randomized controlled pilot study. *J Altern Complement Med*. 2012; 18: 242-251.
99. Raghavendra RM, Vadiraja HS, Nagarathna R, et al. Effects of a yoga program on cortisol rhythm and mood states in early breast cancer patients undergoing adjuvant radiotherapy: A randomized controlled trial. *Integr Cancer Ther*. 2009; 8: 37-46.
100. Schneider A, Weiland C, Enck P, et al. Neuroendocrinological effects of acupuncture treatment in patients with irritable bowel syndrome. *Complement Ther Med*. 2007; 15: 255-263.
101. Schneider A, Enck P, Streitberger K, et al. Acupuncture treatment in irritable bowel syndrome. *Gut*. 2006; 55: 649-654.
102. Tornhage CJ, Skogar O, Borg A, et al. Short- and long-term effects of tactile massage on salivary cortisol concentrations in Parkinson's disease: A randomised controlled pilot study. *BMC Complement Altern Med*. 2013; 13: 357. doi:10.1186/1472-6882-13-357.
103. Woods DL, Beck C, Sinha K. The effect of therapeutic touch on behavioral symptoms and cortisol in persons with dementia. *Forsch Komplementmed*. 2009; 16: 181-189.
104. Boelens A, Reeves R, Replogle H, Koenig G. A randomized trial of the effect of prayer on depression and anxiety. *Int J Psychiatry Med*. 2009; 39: 377-392.
105. Dudgeon WD, Jaggers JR, Phillips KD, et al. Moderate-intensity exercise improves body composition and improves physiological markers of stress in HIV-infected men. *ISRN AIDS*. 2012; 2012: 145127. doi:10.5402/2012/145127.
106. Emery CF, Kiecolt-Glaser JK, Glaser R, Malarkey WB, Frid DJ. Exercise accelerates wound healing among healthy older adults: A preliminary investigation. *J Gerontol A Biol Sci Med Sci*. 2005; 60: 1432-1437.
107. Jaggers JR, Hand GA, Dudgeon WD, et al. Aerobic and resistance training improves mood state among adults living with HIV. *Int J Sports Med*. 2015; 36: 175-182.
108. Lieverse R, Someren EJ, Nielen MM, et al. Bright light treatment in elderly patients with nonseasonal major depressive disorder: A randomized placebo-controlled trial. *Arch Gen Psychiatry*. 2011; 68: 61-70.
109. Saxton JM, Scott EJ, Daley AJ, et al. Effects of an exercise and hypocaloric healthy eating intervention on indices of psychological health status, hypothalamic-pituitary-adrenal axis regulation and immune function after early-stage breast cancer: A randomised controlled trial. *Breast Cancer Res*. 2014; 16: R39. doi:10.1186/bcr3643.
110. Scott E, Daley AJ, Doll H, et al. Effects of an exercise and hypocaloric healthy eating program on biomarkers associated with long-term prognosis after early-stage breast cancer: A randomized controlled trial. *Cancer Causes Control*. 2013; 24: 181-191.
111. Scherder E, Knol D, van Someren E, et al. Effects of low-frequency cranial electrostimulation on the rest-activity rhythm and salivary cortisol in Alzheimer's disease. *Neurorehabil Neural Repair*. 2003; 17: 101-108.
112. Scherder E, Knol D, van Tol MJ, et al. Effects of high-frequency cranial electrostimulation on the rest-activity rhythm and salivary cortisol in Alzheimer's disease. A pilot study. *Dement Geriatr Cogn Disord*. 2006; 22: 267-273.
113. Tam CS, Frost EA, Xie W, et al. No effect of caloric restriction on salivary cortisol levels in overweight men and women. *Metabolism*. 2014; 63: 194-198.
114. Heilbronn L, de Jonge L, Frisard M, et al. Effect of 6-month calorie restriction on biomarkers of longevity, metabolic adaptation, and oxidative stress in overweight individuals. *JAMA*. 2015; 295: 1539-1548.
115. Coste J, Strauch G, Letrait M, Bertagna X. Reliability of hormonal levels for assessing the hypothalamic-pituitary-adrenocortical system in clinical pharmacology. *Br J Clin Pharmacol*. 1994; 38: 474-479.
116. Hellhammer J, Fries E, Schweisthal OW, et al. Several daily measurements are necessary to reliably assess the cortisol rise after awakening: State- and trait components. *Psychoneuroendocrine*. 2007; 32: 80-86.

117. Smyth JM, Ockenfels MC, Gorin AA, et al. Individual differences in the diurnal cycle of cortisol. *Psychoneuroendocrine*. 1997; 22: 89-105.
118. Ice GH, Katz-Stein A, Himes J, Kane RL. Diurnal cycles of salivary cortisol in older adults. *Psychoneuroendocrine*. 2004; 29: 355-370.
119. Wilhelm I, Born J, Kudielka BM, Schlotz W, Wüst S. Is the cortisol awakening rise a response to awakening? *Psychoneuroendocrine*. 2007; 32: 358-366.
120. Golden SH, Sánchez BN, Wu M, et al. Relationship between the cortisol awakening response and other features of the diurnal cortisol rhythm: The multi-ethnic study of atherosclerosis. *Psychoneuroendocrine*. 2013; 38: 2720-2728.
121. Ross KM, Murphy MLM, Adam EK, Chen E, Miller GE. How stable are diurnal cortisol activity indices in healthy individuals? Evidence from three multi-wave studies. *Psychoneuroendocrine*. 2013; 39: 184-193.