Stress
The International Journal on the Biology of Stress

ISSN: 1025-3890 (Print) 1607-8888 (Online) Journal homepage: informahealthcare.com/journals/ists20

A follow-up of cognitive performance and diurnal salivary cortisol changes in former burnout patients

K. Österberg, B. Karlson, B. Malmberg & Å.M. Hansen

To cite this article: K. Österberg, B. Karlson, B. Malmberg & Å.M. Hansen (2012) A follow-up of cognitive performance and diurnal salivary cortisol changes in former burnout patients, Stress, 15:6, 589-600, DOI: 10.3109/10253890.2011.648972

To link to this article: https://doi.org/10.3109/10253890.2011.648972

Published online: 24 Jan 2012.

Submit your article to this journal

Article views: 846

View related articles

Citing articles: 11 View citing articles

Full Terms & Conditions of access and use can be found at
https://informahealthcare.com/action/journalInformation?journalCode=ists20

Stress, November 2012; 15(6): 589–600
© Informa Healthcare USA, Inc.
ISSN 1025-3890 print/ISSN 1607-8888 online
DOI: 10.3109/10253890.2011.648972

informa
healthcare

A follow-up of cognitive performance and diurnal salivary cortisol changes in former burnout patients

K. ÖSTERBERG[1], B. KARLSON[1,2], B. MALMBERG[1], & Å.M. HANSEN[3]

[1]*Division of Occupational and Environmental Medicine, Department of Laboratory Medicine, Lund University, SE-221 85 Lund, Sweden,* [2]*Department of Psychology, Lund University, SE-221 00 Lund, Sweden, and* [3]*National Research Centre for the Working Environment, Lersø Parkallé 105, DK-2100 Copenhagen, Denmark*

(*Received 19 June 2011; revised 17 November 2011; accepted 8 December 2011*)

Abstract
The purpose of this study was to determine whether recovery from burnout is associated with improved cognitive functioning, and whether such improvement is associated with changes in hypothalamic–pituitary–adrenal axis activity and return to work. Forty-five former burnout patients were followed up after 1.5 years with a neuropsychological examination, diurnal salivary cortisol measurements, dexamethasone suppression test (DST), and self-ratings of cognitive problems. At follow-up, improved cognitive performance was observed on several tests of short-term memory and attention. Self-rated cognitive problems decreased considerably, but this decrease was unrelated to the improvement on neuropsychological tests. Diurnal salivary cortisol concentrations at awakening, 30 min after awakening, and in the evening, did not change from baseline to follow-up, nor did the cortisol awakening response. However, slightly, but significantly, stronger suppression of cortisol in response to the DST was observed at follow-up. Improvements in subjective or objective cognitive functioning and changes in diurnal cortisol concentration were unrelated to the extent of work resumption. However, a decreased DST response at follow-up was partially related to improved cognitive performance and work resumption. The clinical implications are that burnout seems to be associated with slight and significantly reversible cognitive impairment, and that self-rated cognitive change during recovery poorly reflects objective cognitive change.

Keywords: *Attention, burnout, cortisol, memory, mood, neuropsychological tests*

Introduction

Burnout is a reaction to long-term interpersonal work stressors and is characterized by overwhelming exhaustion, feelings of cynicism toward work, and reduced professional efficacy (Maslach et al. 2001). Subjective problems with memory and attention are some of the key clinical features of burnout (Socialstyrelsen 2003), verified through elevated scores on self-rating scales for cognitive problems (Öhman et al. 2007; Österberg et al. 2009). Moreover, objective signs of cognitive impairment revealed through neuropsychological tests have been demonstrated in clinical samples (Öhman et al. 2007, Sandström et al. 2005, 2011; Van der Linden et al. 2005; Rydmark et al. 2006; Österberg et al. 2009).

When trying to synthesize the above findings, one main problem is that outcomes on highly similar neuropsychological tests, measuring specific cognitive domains, vary considerably across studies. For example, concerning short-term memory, the studies by Sandström et al. (2005, 2011) reported lower performance on the Rey Complex Figure spatial memory test in burnout patients than among referents, while Öhman et al. (2007) did not find any group difference on the same test. In the latter study, lower performance among burnout patients was found on a verbal learning test (Buschke Concrete Nouns Selective reminding, Trials 2 and 3), while Sandström et al. (2005, 2011) reported normal results on a similar test (Claeson–Dahl verbal learning test). Rydmark et al. (2006) did not find any impairment among burnout patients on three short-term memory tests, and Österberg et al. (2009) found no impairment on either of two similar tests. Concerning

Correspondence: K. Österberg, Division of Occupational and Environmental Medicine, Department of Laboratory Medicine, Lund University, SE-221 85, Lund, Sweden. Tel: + 46 46 17 72 92. Fax: + 46 46 17 72 85. E-mail: kai.osterberg@med.lu.se

working memory, the referents in the study by Rydmark et al. (2006) outperformed burnout patients on a digit span test, while normal findings on similar tests were reported by Sandström et al. (2005, 2011) and Öhman et al. (2007).

Concerning tests for sustained attention and perceptual speed, the set of instruments used in the studies described above are so varied that comparisons of study outcomes are difficult. However, one well-known test, the Wechsler Adult Intelligence Scale Digit Symbol, was used in four of the studies, showing lower performance among burnout patients in the two studies by Öhman et al. (2007) and Österberg et al. (2009), but not in either of the two studies by Sandström et al. (2005, 2011). Moreover, three studies have reported lower performance in burnout patients on reaction-time tests (Sandström et al. 2005, 2011; Rydmark et al. 2006), which could not be reproduced by Österberg et al. (2009) using a similar test (the *k*-test; Levander and Elithorn 1987). Further comparisons of outcomes are problematic due to differences in the tests used, but the general impression of contradictory results persists.

Together these findings have given a rather confusing picture of the nature of cognitive impairment in burnout. Several factors may explain these incongruent findings. One is the small number of referents, ranging between 14 and 19 in four studies (Öhman et al. 2005; Sandström et al. 2005, 2011; Van der Linden et al. 2005), which increases the risk of chance findings. Small samples of burnout patients were also used in the studies by van der Linden et al. (2005; $n = 13$) and Öhman et al. (2007; $n = 19$). Notably, the single published study using large samples of burnout patients and referents failed to detect significant cognitive impairment (Österberg et al. 2009). Another potential factor is inadequate matching of patients and referents on predictors of cognitive performance. For example, in the Öhman et al.'s (2007) study, referents had an 8% higher mean score on a test of verbal ability (Table I in Öhman et al. 2007), reflecting a cognitive domain generally considered robust for minor brain impairment and primarily used as a measure of overall pre-morbid intelligence. Although this 8% difference on the pre-morbid intelligence test bordered on statistical significance, Öhman et al. (2007) made no attempt to adjust for this difference in the between-group comparisons of tests sensitive to impairment. Another example is the study by Sandström et al. (2005), in which none of the referents was older than 39 years, while the burnout group included participants up to 60 years of age. Methodological shortcomings of this kind may lead to underestimation of the cognitive performance of burnout patients.

One way to circumvent problems with matching patients and referents is a prospective study design, ideally with patients examined before and after onset of burnout, allowing them to serve as their own controls. For obvious reasons, this is not feasible, so an alternative prospective approach is to follow burnout patients from shortly after illness onset until substantial recovery has been achieved. If recovered burnout patients were to show marked cognitive improvement, this would give strong support for the existence of initial deficits and also eliminate the uncertainty as to whether cognitive impairment was caused by burnout or preceded burnout (constituting a vulnerability factor). To date, only one such study has been published: a follow-up of the Rydmark et al. (2006) study sample, showing what the authors interpreted as a normalization of cognitive performance (Wahlberg et al. 2009). However, their follow-up study did not take advantage of a repeated measures within-group analysis of the neuropsychological data, but relied on a second cross-sectional comparison of patients and referents. A comparison of the baseline and retest data (Table III in Rydmark et al. 2006; Table II in Wahlberg et al. 2009) indicates that the absence of a group difference in neuropsychological test results at follow-up was due to reduced performance among referents rather than to improvement in the burnout group. To date, no longitudinal study has been published using a repeated measures within-group analysis of neuropsychological test results in patients recovered from burnout.

In several studies, physiological correlates of burnout symptoms have been explored, most often the dynamics of the hypothalamic–pituitary–adrenal (HPA) axis through measurement of salivary cortisol. Compromised functioning of the HPA axis may be highly relevant to cognitive problems, in view of findings from numerous animal model studies, showing that prolonged stress with increased levels of glucocorticoids may lead to structural remodeling of the brain, particularly in the hippocampus, prefrontal cortex, and amygdala (McEwen 2007). However, studies of diurnal cortisol variation in burnout have shown either decreased, increased, or normal cortisol concentrations compared to referents (Melamed et al. 1999; Pruessner et al. 1999; De Vente et al. 2003; Grossi et al. 2003, 2005; Moch et al. 2003; Mommersteeg et al. 2006a,b; Sonnenschein et al. 2007). The feedback sensitivity of the HPA axis has also been studied, producing varying outcomes. Using the dexamethasone suppression test (DST), some studies have shown hypersuppression in burnout (Pruessner et al. 1999; Sonnenschein et al. 2007), while other studies have shown normal findings (Mommersteeg et al. 2006a). One factor that might explain this variation in findings is time since onset of the stressor. A meta-analytic study concluded that more recent stressors were associated with higher morning levels of cortisol, while lower cortisol levels were seen when the onset of stress was more distant (Miller et al. 2007). Again, a prospective study in

which burnout patients are followed well into the recovery phase may provide retrospective evidence for initial disturbances in the HPA axis, possibly a normalization of increased cortisol morning values in early stages of burnout, and a reduced suppression of cortisol in response to a DST, despite the fact that our cross-sectional study at baseline did not reveal major differences in cortisol concentrations between burnout patients and referents (Österberg et al. 2009).

In a recent paper, we described cognitive performance and diurnal salivary cortisol concentrations shortly after the onset of sick leave due to work-related burnout (Österberg et al. 2009). In this study, we reassessed the former patients after substantial recovery from burnout, testing the following hypotheses:

(1) Improved cognitive functioning in vital cognitive domains such as short-term memory, cognitive speed, and sustained attention will be observed at follow-up.
(2) The improved cognitive functioning will be related to a subjective reduction in everyday cognitive problems.
(3) Decreased morning cortisol concentrations and decreased cortisol suppression in response to a DST will be observed at follow-up, and these changes will be related to objective and subjective improvement in cognitive functioning.
(4) A successful return to work will be associated with objective and subjective improvement in cognitive functioning, decreased morning cortisol concentrations, and decreased cortisol suppression in response to a DST.

Methods

Participants

The former burnout patients were previous participants in a broad study focused on a workplace intervention to facilitate return to work (Karlson et al. 2010). In that study, eligibility for participation was based on data from social insurance records, requiring recent sick leave with work strain as a probable cause and having any of the International Classification of Diseases (ICD-10) diagnoses within the category F43 (reaction to severe stress, and adjustment disorders), with the exception of post-traumatic stress disorder (F43.1). Prior to the workplace intervention, a baseline medical and social investigation had been carried out, including a neuropsychological examination (Österberg et al. 2009). One to two years after the intervention, in conjunction with a general follow-up of health and well-being, patients were invited to a neuropsychological reexamination (this study). Of the 65 patients examined at baseline, 45 agreed to participate in the neuropsychological follow-up. Most participants were women (71%), and the mean age was 50.8 years (Table I). Mean time-lag from baseline to follow-up was 18 months (range 14–25), and mean time-lag from onset of sick leave to follow-up was 22 months (range 17–28). All participants were reassessed by a physician and a social worker at the Department of Occupational and Environmental Medicine, Lund University Hospital, Lund, Sweden. The participants' health status was generally good, although 31% had minor somatic disorders (Table II). At follow-up, 11% met the criteria for exhaustion disorder suggested by the Swedish National Board of Health and Welfare (Socialstyrelsen 2003), a decrease from 87% at baseline. Screening with the Primary Care Evaluation of Mental Disorders (Prime-MD; Spitzer et al. 1994) showed that the prevalence of a Prime-MD diagnosis had decreased from 76% at baseline to 47% at follow-up, reflecting a marked fall in the prevalence of depression and anxiety disorder. Scores on self-rating scales for depressive symptoms, Symptom Checklist-90 (SCL-90; Derogatis 1992) and Beck Depression Inventory (BDI; Beck et al. 1988) also decreased markedly from baseline (SCL-90 mean change: depression: -0.72; anxiety: -0.61, BDI mean change: 7.0, all $p < 0.002$). Yet BDI and SCL-90 scores were still elevated compared to means commonly found in the general population, and a minority still reported considerable depressive symptoms (Table II). At follow-up, antidepressants, anxiolytic drugs, or sleeping pills were being used by 49% (including in combination), and 27% were undergoing psychotherapy. Also at follow-up, 38% had returned to the same degree of employment as before the onset of burnout, 33% were on part-time sick leave, partially retired or partially unemployed, or had settled for part-time work (without sickness

Table I. Demographic characteristics of the participating former burnout patients ($n = 45$).

Characteristic	
Age (years)	
Mean (SD)	50.8 (9.5)
Range	32–64
	n (%)
Gender	
Men	13 (29)
Women	32 (71)
Education	
Nine-year compulsory schooling	7 (16)
Upper secondary school	13 (29)
University studies	25 (56)
Supervisory/managerial position	
Preceding sick leave due to burnout	14 (31)
At follow-up	6 (19)*
Employment	
Private sector	12 (27)
Public sector	33 (73)

* Figures based on those having returned to work to any extent at follow-up ($n = 31$).

Table II. Medical characteristics of the participating former burnout patients ($n = 45$).

Characteristic	*n* (%)
Somatic disorders (total)	14 (31)
Hypertension (treated)	5 (11)
Type I diabetes	1 (2)
Type II diabetes	1 (2)
Other*	8 (18)
Medication	
Antidepressants: SSRI, SNRI, or NaSSA†	20 (44)
Anxiolytics or sleeping pills	3 (7)
Medication for somatic disorders‡	21 (47)
Hormonal replacements/contraceptives	8 (18)
No medication	9 (20)
Psychotherapy	12 (27)
Alcohol consumption	
Low to moderate	45 (100)
High	0
Prime-MD	
Major depression	4 (9)¶
Mild depression or dysthymia	2 (4)§
Anxiety disorder	3 (7)
Anxiety disorder + major depression	5 (11)‖
Somatization disorder only	4 (9)
Somatoform syndrome NUD only	3 (7)
BDI	
Severe depression	5 (11)
Moderate depression	7 (16)
Slight depression	12 (27)
Minimal depression	21 (47)
	Mean (SD)
BDI total score	13.2 (10.4)
SCL-90 subscales	
Anxiety	1.1 (0.8)
Somatization	1.0 (0.8)
Depression	1.3 (1.0)

Notes: Prime-MD, Primary Care Evaluation of Mental Disorders (Spitzer et al. 1994); MBI-GS, Maslach Burnout Inventory-General Survey (Maslach et al. 1996); SCL-90, Symptom Checklist 90 (Derogatis 1992); BDI, Beck Depression Inventory (BDI; Beck et al. 1988); *Two patients had thyroid disorder (treated), one in combination with slight B12 deficiency (treated), and the other patients each had one of: airways allergy with sleep apnea (treated), atrial fibrillation, endometriosis, gluten intolerance (treated), hypocalcemia plus iron deficiency (treated), and recurrent diverticulitis; †Selective serotonin reuptake inhibitor (SSRI), serotonin–norepinephrine reuptake inhibitor (SNRI), or noradrenergic and specific serotonergic antidepressant (NaSSA); ‡Mostly antihypertensive and antihyperlipidemic drugs and vitamin or mineral supplements; ¶One subject also fulfilled the criteria for somatization disorder; §One subject also fulfilled the criteria for somatoform syndrome (non-ulcer dyspepsia; NUD); ‖Subjects also fulfilled the criteria for: somatoform syndrome NUD ($n = 2$), somatization disorder ($n = 2$; one in combination with obsessive–compulsive disorder), and obsessive–compulsive disorder combined with social phobia ($n = 1$).

benefit) in order to cope, 18% remained on full-time sick leave, and 13% were on parental leave or had been given severance pay awaiting retirement.

Neuropsychological tests

An experienced psychologist tested each participant individually in the afternoon. All tests were administered according to the standard procedures given in the test manuals. A colleague cross-checked manual scorings. All tests from the baseline study (Österberg et al. 2009) were repeated at follow-up.

SRB:1 vocabulary from the DS-battery (Dureman et al. 1971), a verbal knowledge or intelligence test containing 30 items of multiple-choice type; *Cronholm–Molander verbal memory* (Cronholm and Molander 1957), an associative learning task comprising immediate and delayed recall sections; *Austin Maze* with the Milner pathway (Milner 1965; Walsh 1985), a test of spatial learning and strategy, in a computerized version (Österberg et al. 2000). Ten trials were used, and error and performance time refer to the nine memory trials; *Digit Symbol* from the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler 1981), a test of perceptual and fine motor speed; *WAIS-R NI incidental learning* (Kaplan et al. 1991), a test of memory for the digit–symbol combinations of the Digit Symbol test. The modified version used showed only the stimulus key row with the symbols blanked out (scoring range 0–9); *APT k-test* from the Automated Psychological Test System (APT; Levander and Elithorn 1987), measuring sustained attention. The *k*-test required continuous visual scanning of computer screen images, containing a scatter of 10 letters, for the presence/absence of a critical stimulus (*k*). Around 100 stimulus sets were shown over a period of 5 min. The task was to respond correctly and quickly to each stimulus set by pressing the corresponding keys ("*k*" vs. "no *k*"). Results were expressed (a) according to signal detection theory, as d' (accuracy) and β (balance between missed hits to targets and false hits to non-targets; a negative β indicating that missed hits were more common than false hits), (b) as reaction time (RT) level in relation to stimulus laterality (the mean of single RT responses for correct hits, left–center–right), (c) as error ratio (proportion of missed and false hits), and (d) as search strategy [the RT ratio for correct yes/no responses with "*k*" present/absent on a scale from sequential (obsessive) to global (flexible), a higher value indicating a more global strategy]. (In addition, a number of other tests were carried out only at follow-up, the results of which will be published separately.)

Self-ratings of cognitive problems

Euroquest-9, comprising the Memory and Concentration Scale of the larger Euroquest (Chouanière et al. 1997), was used to assess subjective cognitive problems at baseline and follow-up. Six items address memory problems, and three items address problems with attention and concentration. The verbal anchors for the four-point (1–4) response range are "seldom/never," "sometimes," "often," and "very often." The scale has been validated as a sensitive measure of slight cognitive impairment (Karlson et al. 2000; Carter et al. 2002).

Salivary cortisol

The participants collected saliva in sampling tubes with cotton swabs (Salivette®; Sarstedt Ltd, Leicester, UK) on a regular weekday at awakening, 30 min later, and around 21:00 h in the evening. On the following regular weekday, a DST was also carried out, by ingesting 0.5 mg dexamethasone (Oradexon®; MSD, Whitehouse Station, NJ, USA) at bedtime and taking two additional saliva samples the following morning. Oral and written instructions emphasized that swabs should be kept in the mouth until thoroughly hydrated (at least 2 min but not exceeding 5 min) and to avoid smoking and eating heavy meals 1 h prior to sampling. To avoid contamination due to microbleeding from the gums, tooth brushing and eating were not allowed until after the morning samples. The participants were shown how to label each sampling tube with the exact time of measurement, and to record the exact time of awakening. Samples were stored in the refrigerator until returned by mail, and then frozen at −20°C until analysis.

Determination of cortisol in saliva was carried out using a competitive radioimmunoassay (RIA) designed for quantitative *in vitro* measurement of cortisol in serum, plasma, urine, and saliva (Spectria Cortisol Coated Tube RIA, purchased from Orion Diagnostica, Espoo, Finland), according to the manufacturer's specifications. The sample volume was 150 μL, the range of the standard solutions prepared was 1.0–100.0 nmol/L, and the incubation time was 30 min at 37°C. The specifications given by the manufacturer were a sensitivity of twice the SD of the zero binding value in saliva (0.8 nmol/L), a bias of 110% (103–115%), an intra-assay variation of 5.4%, and an inter-assay variation of 7.3%. The cross-reactivity to cortisone was $<0.2\%$. A 1470 Wizard gamma counter (Wallac, Turku, Finland) was used for the measurement of radioactivity. A method evaluation of certified reference material in water, performed by our laboratory, showed no bias in the method, with recovery being 97% (95% CI: 94.0–100.9). Limit of detection was 1.59 nmol/L. Between-run coefficients of variation were 19% at 11.5 nmol/L and 16% at 49.2 nmol/L (Hansen et al. 2003). To show equivalence between different runs, natural saliva samples at two concentrations (5.9 and 18.5 nmol/L) were used as control materials and analyzed together with the samples. Westgard control charts were used to document that the analytical method remained under analytical and statistical control, so that the accuracy and precision of the analytical methods remained stable (Westgard et al. 1981). The performance of the methods has been validated using inter-laboratory comparison schemes (Garde et al. 2003; Hansen et al. 2003).

In addition to the raw cortisol concentrations, the following indices were used: (i) the cortisol awakening response (CAR), defined as the increase in cortisol concentration from the awakening sample to the +30 min sample, expressed in absolute (nmol/L) and proportional (%) changes; (ii) the cortisol morning peak (CMP), defined as the highest concentration of cortisol in saliva found in either of the two morning samples; (iii) the cortisol range (CR), defined as the decrease in cortisol concentration from the CMP to the 21:00 h evening sample; and (iv) the DST response, defined as the proportional (%) decrease in cortisol concentration after ingestion of dexamethasone, in comparison with the normal weekday morning values.

Data analysis

The statistical analyses were carried out using SPSS, version 18.0 (SPSS Inc., Chicago, IL, USA). Visual inspection showed that neuropsychological test scores and Euroquest-9 scores were approximately normally distributed, except for skewed distributions for the variables WAIS-R NI incidental learning and APT *k*-test error ratio, necessitating ranking prior to statistical analysis. All cortisol raw values had positively skewed distributions, hence they were log-transformed (ln) prior to analysis and computation of the CAR, CMP, and CR indexes. The baseline versus follow-up comparisons of DST proportional decrease were carried out on ranked data.

Within-group comparisons of cortisol, test, and questionnaire data were carried out using Linear Mixed Models of SPSS, with Timepoint (baseline vs. follow-up) as the main factor, together with the covariates age, gender, medication with psychopharmacological drugs, and SCL-90 depression score. In addition, for neuropsychological test scores, education level was explored as a covariate, and for cortisol measures, time of awakening. Initially, all covariates were entered, and then removed one at a time (based on highest *p*-value), until only covariates with $p < 0.20$ remained (Mickey and Greenland 1989). In case a significant effect of Timepoint was observed and gender and/or age remained a covariate in the final model, the interactions Timepoint by Gender and Timepoint by Age were added (separately) to explore possible differential effects of gender and/or age on the change from baseline to follow-up. To explore whether the change in test scores from baseline to follow-up was related to having resumed work, the difference between test scores at follow-up and baseline (T2 minus T1) was compared with the degree of work resumption at follow-up compared to before burnout onset (from 0% to 100%, due to sick leave, voluntary reduction in work hours, or other causes). For this analysis, Pearson's partial correlation was used, controlling for age and gender as a baseline and, when justified, for remaining covariates. The same method was used for analyzing relations between

changes in other variables, such as cortisol concentrations and self-ratings of cognitive problems.

To ensure that the follow-up participants ($n = 45$) did not differ to any appreciable extent from those refraining from participation ($n = 20$), baseline test scores and medical characteristics in these groups were compared using independent-samples t tests.

In case of statistical significance for the factor Timepoint, the effect size is given as partial eta square (η_p^2). Two-tailed p-values <0.05 were considered statistically significant.

Ethics

All participants gave written informed consent to participate in the study. The study protocols were approved by the Lund University Ethics Committee (dnr. 784-03) and the Regional Ethical Review Board in Lund (dnr. 583/2006).

Results

Neuropsychological tests

At follow-up (Table III), improved test scores were observed in two of the three memory tests: the Cronholm–Molander verbal memory test [immediate recall: $F(1, 43.3) = 24.0$, $p < 0.001$, $\eta_p^2 = 0.36$; delayed recall: $F(1, 44.8) = 11.5$, $p = 0.001$, $\eta_p^2 = 0.20$] and the Austin Maze spatial memory test [errors: $F(1, 45.5) = 24.0$, $p < 0.001$, $\eta_p^2 = 0.31$]. Moreover, improved test scores were found on both tests of sustained attention and processing speed: the WAIS-R Digit Symbol [$F(1, 48.8) = 9.4$, $p = 0.004$, $\eta_p^2 = 0.16$], and the APT k-test [RT, center part of screen: $F(1, 43.9) = 4.56$, $p = 0.038$, APT = 0.09; RT, right part of screen: $F(1, 43.6) = 11.0$, $p = 0.002$, $\eta_p^2 = 0.20$]. The interaction Timepoint by Gender [explored for the Cronholm–Molander verbal memory test (immediate and delayed), WAIS-R Digit Symbol, and APT k-test RT (center and right parts of screen)] did not reveal any differential effect of gender (interaction p-values ranging from 0.45 to 0.88). However, the interaction Timepoint by age (explored for the Cronholm–Molander verbal memory test delayed, Austin Maze errors, and WAIS-R Digit Symbol) was significant for the scores in the latter two tests (interaction p-values 0.023 and 0.015, respectively). Austin Maze error scores tended to improve (decrease) more among *older* participants (−13.0 vs. −5.8) when comparing similar sized age subgroups [older: ≥52 years ($n = 22$); younger: <52 years ($n = 23$)], while WAIS-R Digit Symbol scores tended to improve (increase) mainly among *younger* participants (means: −0.64 vs. +5.7, respectively).

Table III. Neuropsychological test raw scores for the former burnout patients, at baseline and at follow-up ($n = 45$).

	Baseline		Follow-up		ANOVA		
Cognitive domain and test	Mean	SD	Mean	SD	$F^{\star}$	$p^{\dagger}$	η_p^2
Verbal intellectual ability							
SRB:1 vocabulary	24.8	3.2	25.1	2.8	0.58	$0.45^{A,E}$	–
Memory							
Cronholm–Molander verbal memory, immediate recall	21.0	4.4	24.4	4.2	23.96	$<0.001^{E,G}$	0.356
Cronholm–Molander verbal memory, delayed recall	17.3	4.9	19.8	6.3	11.47	$0.001^{A,E,G}$	0.204
Austin Maze, total errors‡	38.0	22.5	28.8	16.3	20.19	$<0.001^{A,P}$	0.307
Austin Maze, total performance time (s)‡	297	115	289	100	2.25	$0.14^{A,E}$	–
Incidental learning WAIS-R NI, number correct	6.8	2.1	6.7	2.2	<0.01	0.97^{A}	–
Sustained attention and processing speed							
WAIS-R Digit Symbol	52.4	11.4	55.0	11.5	9.38	$0.004^{A,G}$	0.161
APT k-test¶, d'	4.08	0.44	4.09	0.42	0.06	$0.81^{A,E,G}$	–
APT k-test, β§	−0.60	0.69	−0.77	0.66	1.94	0.17^{G}	–
APT k-test, RT level for correct responses, left part of screen (ms)‡	1141	279	1081	207	0.44	$0.51^{A,D,G}$	–
APT k-test, RT level for correct responses, center part of screen (ms)‡	1027	265	943	213	4.56	0.038^{G}	0.094
APT k-test, RT level for correct responses, right part of screen (ms)‡	1261	306	1147	216	11.01	0.002^{G}	0.201
APT k-test, search strategy‖	0.56	0.10	0.58	0.07	2.35	$0.13^{A,G}$	–
APT k-test, error ratio (%)‡	2.2	2.4	2.3	2.2	0.04	$0.83^{D,G}$	–

Abbreviations: APT, Automated Psychological Test System; RT, reaction time; SRB:1, Synonyms; WAIS-R, Wechsler Adult Intelligence Scale-Revised; ⋆ Df ranged from [1, 43.3] to [1, 48.8], except for APT k-test RT level for correct responses, left part of screen (ms) df = [1, 57.7]; † Linear Mixed Models fixed effects p-values for within-group comparisons, corrected for age (A), Symptom Checklist-90 depression score (D), level of education (E), gender (G), and medication with psychopharmacological drugs (P) in case $p < 0.20$ for each covariate; ‡ A lower score signifies better performance; ¶ $N = 44$ at baseline; for one case, k-test data were not available due to patient misunderstanding the instructions; § Qualitative variable; a negative β indicates that false rejection (missed "k"s) was more common than false alarms (alarms when k is not present); ‖ Qualitative variable; a lower value indicates a more sequential search strategy, while a higher value indicates a more global search strategy.

No significant changes in test scores at follow-up were observed for Verbal intellectual ability, Austin Maze performance time, Incidental learning WAIS-R NI, or for the five APT *k*-test variables d', β, RT left part of screen, search strategy, and error ratio.

Concerning return to work, no relationship was found between change in any neuropsychological test variable from baseline to follow-up and the extent of work resumption ($r \leq 0.23$ throughout; $p \geq 0.14$).

Self-ratings of cognitive problems

The Euroquest-9 scores fell from a mean total score of 3.00 (SD 0.62) at baseline to 2.43 (SD 0.70) at follow-up [$F(1, 57.5) = 17.1$, $p < 0.0001$, $\eta_p^2 = 0.23$]. The concentration and attention problems subscale score tended to decrease more [baseline mean 3.07, SD 0.67 vs. follow-up mean 2.30, SD 0.73; $F(1, 58.3) = 18.7$, $p < 0.0001$, $\eta_p^2 = 0.24$] than the memory problems subscale score [baseline mean 2.97, SD 0.67 vs. follow-up mean 2.49, SD 0.74; $F(1, 56.6) = 11.4$, $p = 0.001$, $\eta_p^2 = 0.17$]. The interactions Timepoint by Gender and Timepoint by Age were not significant for the Euroquest-9 total score or either of the subscale scores, thus failing to support any differential age or gender effects on the changes in Euroquest-9 scores from baseline to follow-up.

Concerning relationships with a change in any of the neuropsychological test scores from baseline to follow-up, only the Austin Maze test variables were related to Euroquest-9 scores, however in the opposite direction of what might have been expected: the better the Austin Maze performance time at follow-up compared to baseline, the less reduction in self-ratings of cognitive problems was observed, regarding total score ($r = 0.35$; $p = 0.025$) as well as the concentration and attention subscale ($r = 0.35$; $p = 0.021$). The change in Austin Maze errors showed a similar inverse relationship with the change in Euroquest-9 total score ($r = 0.38$; $p = 0.012$) and memory subscale score ($r = 0.42$; $p = 0.006$). It was also observed that the changes in all Euroquest-9 scales correlated with the change in the depression subscale of the SCL-90 (r ranged from 0.31 to 0.33, p ranged from 0.027 to 0.041), and that the change in SCL-90 depression scores did not correlate with the change in any neuropsychological test.

Concerning return to work, no relationship was found between change in any of the Euroquest-9 scores from baseline to follow-up and the extent of work resumption ($r \leq 0.20$ throughout; $p \geq 0.18$).

Salivary cortisol concentrations

Analyzable cortisol samples from ordinary weekdays at both baseline and follow-up were obtained from 34 to 37 participants, depending on sampling point (Table IV). Six participants failed to return any weekday samples at baseline and/or follow-up, and 10 participants had missing data for 1–3 sampling points due to insufficient amounts of saliva (dry swabs). Complete DST response data at both baseline and follow-up were obtained from 28 participants. Missing data were due to failure to return any DST samples at baseline and/or follow-up ($n = 7$), dry swabs ($n = 8$), erroneous sampling time ($n = 1$), or values considered as outliers (> 100 nmol/L, $n = 1$). Based on self-reports, the median time lapse from awakening to the first morning sample was 3.5 min at follow-up ($n = 34$, Table IV). A few participants ($n = 4$) reported a time lapse exceeding 15 min but still showed a substantial CAR (medians: +46.2%, +8.5 nmol/l), which is why their data were not excluded from any analysis. The self-reports indicated a median time lapse from the first morning sample to the +30 min sample of 30 min (range 28–40 min), which was considered satisfactory.

The raw cortisol concentrations and the CAR, CMP, and CR index values were similar at baseline and follow-up (Table IV). However, the DST response, i.e. the DST-induced proportional (%) decrease in cortisol concentration from the morning values on an ordinary weekday, was slightly stronger at follow-up, concerning both the awakening measurement [$F(1, 29.9) = 5.7$, $p < 0.024$, $\eta_p^2 = 0.16$] and the +30 min measurement [$F(1, 25) = 4.4$, $p < 0.046$, $\eta_p^2 = 0.15$]. In addition, the raw cortisol concentrations after the DST were also significantly lower at follow-up at the awakening measurement [$F(1, 32) = 4.5$, $p < 0.042$, $\eta_p^2 = 0.12$], though not at the +30 min measurement [$F(1, 32) = 2.1$, $p = 0.16$]. No interaction with age or gender was observed for any of the cortisol measures.

Change in any of the neuropsychological measures from baseline to follow-up was not found to be related to changes in CAR and CMP ($r < 0.33$; $p > 0.09$). Concerning CR, only a relationship with the Incidental learning WAIS-R NI score was found; an improvement in a number of correctly recalled items was related to an increase in CR ($r = 0.34$; $p = 0.048$). A decrease from baseline to follow-up in the cortisol suppression in response to DST (i.e. compared to a normal weekday) at awakening was related to greater improvement in the Cronholm–Molander verbal memory test (immediate recall; $r = 0.43$; $p = 0.019$), and to a shift toward a more global search strategy in the *k*-test ($r = 0.40$; $p = 0.037$). The change in DST response at +30 min, however, was not related to a change in any of the neuropsychological measures.

Concerning relations with the Euroquest-9 scores of subjective cognitive problems, only the CMP was found to be related to the concentration and attention problems subscale score; a decrease in CMP from baseline to follow-up was related to a more marked

Table IV. Salivary cortisol concentrations and computed index values (medians and quartiles) on a normal weekday, and after dexamethasone treatment, among the former burnout patients, at baseline and follow-up.

Cortisol measure (values in nmol/L unless otherwise stated; *n* refers to number of participants in the burnout group with valid data for both timepoints)	Baseline		Follow-up		ANOVA		
	MD	Q1/Q3	MD	Q1/Q3	$F^{\star}$	$p^{\dagger}$	η_p^2
Raw values							
At awakening ($n = 34$)	10.1	6.1/17.0	9.8	6.9/13.9	0.18	0.67^{P}	–
30 min after awakening ($n = 34$)	14.4	9.1/21.8	15.5	10.7/18.9	0.07	0.80	–
At 21.00 h ($n = 37$)[‡]	1.4	1.0/1.7	1.1	0.7/1.6	0.42	$0.52^{D,G}$	–
After dexamethasone treatment							
At awakening ($n = 33$)[‡]	0.8	0.4/1.2	0.5	0.3/0.8	4.51	0.042	0.124
30 min after awakening ($n = 29$)[‡]	0.7	0.5/1.1	0.6	0.4/1.0	2.09	0.16	–
Indexes							
CAR ($n = 31$)[¶,§]							
Absolute increase	3.8	−1.5/9.1	5.3	2.0/9.7	0.62	$0.44^{A,G,T}$	–
Proportional increase (%)	42.6	−12.9/100.9	47.5	23.0/117.1	1.76	0.19	–
CMP value ($n = 37$)	14.4	8.9/22.1	15.6	10.9/19.6	0.62	0.44^{P}	–
Cortisol decrease from morning (peak value) to evening (CR; $n = 36$)	−13.2	−20.9/−6.7	−14.9	−19.0/−10.6	1.12	$0.30^{D,G}$	–
Proportional cortisol decrease after dexamethasone treatment compared with the corresponding measurement on a normal weekday							
At awakening (%; $n = 31$)	−90.7	−97.0/−86.0	−95.1	−96.7/−92.0	5.65	0.024^{P}	0.159
30 min after awakening (%; $n = 26$)	−93.1	−96.7/−88.0	−96.2	−98.1/−92.8	4.41	0.046	0.150

Abbreviations: CAR, cortisol awakening response; CMP, cortisol morning peak; CR, cortisol range; * *Df* ranged from [1, 25] to [1, 45.5]; [†] Mixed models ANOVA *p*-values for within-group comparisons, corrected for age (A), SCL-90 depression score (D), gender (G), medication with psychopharmacological drugs (P), and time of awakening (T) in case $p < 0.20$ for each covariate; [‡] Most cortisol concentrations were below the detection limit (1.59 nmol/L) and are thus approximate; [¶] Computed as the increase from the awakening value to the value 30 min after awakening; [§] A negative CAR was found among 32% ($n = 10/31$) at baseline and among 13% ($n = 4/31$) at follow-up.

reduction in concentration and attention problems ($r = 0.42$; $p = 0.012$).

None of the changes in the cortisol measures were significantly related to the extent of work resumption, although a decrease from baseline to follow-up in cortisol suppression in response to the DST at awakening showed a tendency to be related to a higher degree of work resumption, bordering on significance ($r = 0.35$; $p = 0.065$). Those who had returned to the same degree of employment as prior to burnout onset ($n = 11$) had almost unchanged DST suppression at awakening (median: −0.3%), while those with less than full return to work ($n = 20$) had increased DST suppression at follow-up (median: +7.2%; $p = 0.039$; Mann–Whitney *U* test).

Participants as representative of the baseline group

The follow-up participants ($n = 45$) had characteristics quite similar to those refraining from participation ($n = 20$), concerning baseline demographic and medical characteristics as well as scores on self-rating scales and neuropsychological tests. The single significant difference observed was a slightly lower SRB:1 vocabulary score among non-participants (22.8 vs. 24.8; $p = 0.047$).

Discussion

This study showed that persons with work-related burnout improved their cognitive performance on several tests of short-term memory and attention at a follow-up approximately 2 years after sick leave onset. Self-rated cognitive problems also decreased considerably from baseline to follow-up, but self-rated improvement was unrelated to actual objective improvement in neuropsychological tests. Diurnal salivary cortisol concentrations showed no significant change from baseline to follow-up, though the reduction in cortisol in response to a DST was slightly more pronounced at follow-up. Neither improvements in subjective or objective cognitive functioning nor changes in diurnal cortisol concentrations were found to predict the extent of work resumption after burnout. However, unchanged or decreased DST suppression was partially related to improved test performance and work resumption, and a stronger reduction in self-rated concentration and attention problems was related to a decreased CMP.

The observed improvements in neuropsychological tests shed new light on the results of our baseline study of the burnout patients (Österberg et al. 2009). In the baseline comparison with a matched reference group, only a single test parameter (WAIS-R Digit Symbol)

showed lower performance, which hardly supported the presence of any major cognitive impairment. We briefly discussed whether this result might reflect a type II error due to extraordinary pre-morbid capacities among persons with burnout who had managed to shoulder an exceptional workload. The results of this study seem to support this interpretation, implying that matching of burnout patients with referents using vocabulary test scores (or demographic data on education, age, and gender) might not be sufficient. To successfully match burnout patients with referents, the complexity of work tasks and total workload managed over a number of years probably need to be taken into consideration. The question of how we should define suitable criteria for these aspects is, however, hardly within the scope of this study, but one solution may be to avoid a pure cross-sectional study design.

The improved scores on tests for short-term memory and attention at follow-up support the presence of cognitive impairment in early stages of burnout. In a broad sense, the findings confirm the observations of (a) verbal learning deficits reported by Öhman et al. (2007), (b) spatial memory deficits reported by Sandström et al. (2005, 2011), and (c) deficits in sustained attention and perceptual speed, as measured by the WAIS-R Digit Symbol test (Öhman et al. 2007; Österberg et al. 2009) and reaction-time tests (Sandström et al. 2005, 2011; Rydmark et al. 2006). Moreover, the Austin Maze spatial memory test score tended to improve more among older persons, while WAIS-R Digit Symbol scores improved mainly among younger persons. Speculatively, this might indicate that older persons' spatial memory functions (hippocampal dependant) are more vulnerable to burnout, while among younger persons, attention and cognitive speed (executive functions, associated with prefrontal areas) are more vulnerable. In addition, this study indicates that initial cognitive impairment in burnout is largely reversible during a 1.5-year recovery period. Most burnout patients had received a combination of treatments (e.g. workplace intervention, psychotherapy, antidepressants, and sleeping pills), which probably contributed to the recovery process and possibly to improvement in cognitive functions. The reduction in problems with self-rated attention and memory was, however, generally unrelated to the actual improvement observed in the neuropsychological tests. The clinical implication of this may be that the individual patient's report of improved/unimproved cognitive functioning during the recovery phase is a poor indicator of objective improvement. Moreover, the extent of reduction in self-rated cognitive problems paralleled reduced self-rated depression, indicating that subjective cognitive problems are largely related to depressed mood, as previously shown (Österberg et al. 2009).

The present findings add to the results of several recently published studies, showing normal diurnal cortisol concentrations in burnout patients, and no change in diurnal cortisol even after substantial recovery (Mommersteeg et al. 2006b). This indicates that basal HPA axis functioning is unaffected in burnout. Thus, it is not surprising that we generally failed to find associations between the diurnal cortisol parameters and improved cognitive performance. The only exception was total diurnal cortisol variation (CR), for which an increase at follow-up was weakly related to increased performance on the WAIS-NI Incidental learning test. Considering the number of correlation analyses carried out (11 cortisol parameters × 14 cognitive test parameters), this single relationship might be a chance finding. Notably, no relation was found between changes in CR/CMP and the improved performance seen on most of the cognitive speed parameters (the RT parameters of the *k*-test, the WAIS-R Digit Symbol, and the Austin Maze total performance time), in contrast to the clear relationships found between cognitive speed, CR, and CMP in our baseline study (Österberg et al. 2009).

The DST challenge showed that the suppression of cortisol at awakening had increased at follow-up compared to baseline. This indicates an increase in the negative feedback activity of the HPA axis at follow-up. The observation is contrary to what might have been expected, as in some previous studies, high DST-induced suppression of cortisol has been related to more severe burnout symptoms (Pruessner et al. 1999; Sonnenschein et al. 2007), while our patients had markedly lower levels of burnout symptoms at follow-up. Another cortisol follow-up study of burnout patients, however, did not find a change in the DST response from baseline to follow-up, nor any difference at baseline compared to referents (Mommersteeg et al. 2006a,b). Our finding of increased DST-induced suppression of cortisol secretion at follow-up should probably be interpreted with great caution, because most DST values were below the detection limit (1.6 nmol/L) and are thus approximate. Yet the follow-up morning cortisol concentrations after DST in the burnout group (medians 0.5–0.6 nmol/L) do indeed seem low in comparison with our previous findings among healthy referents (medians 2.0–2.4 nmol/L) using similar methods (Carlsson et al. 2006). At least, the DST results fail to support reduced (normalized) negative feedback activity of the HPA axis in parallel with reduced symptoms and improved cognitive performance, but rather suggest that high negative feedback activity may persist even after substantial recovery from burnout.

However, having unchanged or decreased DST suppression at follow-up, at awakening, was associated with improved immediate recall in the verbal memory test and a more global or flexible search

strategy in the attention test. Moreover, full work resumption tended to be associated with unchanged or decreased DST suppression. These observations give partial support to the hypothesis that more successful recovery from burnout is associated with reduced feedback activity in the HPA axis, although it should be kept in mind that only 2 out of 14 cognitive test parameters were significantly associated with the DST variable.

This study has several strengths. First, by following a large group of burnout patients, using a repeated measures design, statistical power was gained, increasing the chances of detecting even subtle improvements. Second, by using former patients as their own controls, any peculiarities in pre-morbid intellectual abilities were controlled for, eliminating the problems associated with composing a well-matched reference group. However, there are also some limitations inherent in the study design. One possible shortcoming concerns repeated testing using the same cognitive tests, which could give rise to learning effects, possibly leading to spurious findings of improvement. To counteract this risk, parallel versions of the memory tests, and the WAIS-R Digit Symbol test were used, i.e. identical tests that consisted of different sets of equally difficult stimuli. Moreover, in a previous follow-up study (Link et al. 2006), we found no significant improvement for any of the test variables used in this study at a 1-year follow-up in a healthy control group. Although we cannot rule out that learning effects contributed, to some degree, to the improvements observed among burnout cases in this study, such effects can hardly explain the general picture of improvement.

Another issue is that the measurements of diurnal salivary cortisol were based on salivary samples from one single day, while recent recommendations suggest taking samples from at least two consecutive days in order to provide more stable values (Hellhammer et al. 2007). The single-day sampling at follow-up was chosen because the same procedure had been used at baseline, implying that little could be gained by increasing the precision only at follow-up. A third day of measurements (in addition to one weekday and the DST day) was also considered to seriously jeopardize participation. The lower precision associated with single-day sampling may have reduced the possibility to detect subtle changes in the cortisol measures, and obscured possible relationships with the other measures.

We also made rather simple hypotheses regarding the baseline to follow-up change in cortisol concentrations and the DST response, and how such changes might be related to improvements in neuropsychological tests and self-ratings. Exposure to work stress at baseline was regarded as rather short term, and we therefore hypothesized that morning cortisol concentrations, CAR, CMP, and CR, would decrease at follow-up, after substantial recovery. However, it is possible that individual patients' diurnal cortisol variation had changed in different directions, due to the state of the individual at baseline. For example, a subset of patients might have experienced very long-term work stress before beginning sick leave, leading to a chronic stress response with lower than normal baseline morning cortisol, CAR, CMP, and CR (Miller et al. 2007). For such a subset of patients, the direction of change in the HPA axis parameters from baseline to follow-up might have been the opposite (from low to normal) of the hypothesized direction, which would be obscured in analyses of the total group. Due to our limited sample size, we were unable to reliably discern and analyze such potential subgroups.

Moreover, the fraction of the burnout group that still fulfilled the criteria for exhaustion disorder at follow-up (11%) was included in the analyses; i.e. a subset of participants might not be regarded as adequately recovered from burnout. However, these participants ($n = 5$) all had decreased self-ratings of cognitive problems (Euroquest-9) at follow-up, accompanied by the same pattern of reduced mental distress (SCL-90) and depression (BDI) as found among the other participants, which justified including them in an attempt to avoid any selection effect. There were also no obvious selection effects identified among the 69% of participants in the baseline study who agreed to participate in the follow-up; participants and non-participants were quite similar at baseline on a large number of demographic, medical, and psychological variables. We also found no evidence that non-participants refrained from participation due to less successful medical or psychological recovery in the years following the baseline examination. Sixteen of the 20 non-participants agreed to fill out follow-up questionnaires, which showed that non-participants' self-ratings of cognitive problems, depression (BDI) and mental distress (SCL-90) were similar to, or slightly better than, the participants' scores, including the extent of work resumption (data not shown).

To conclude, the findings indicate that acute stages of burnout are associated with slight and diffuse cognitive impairment, which is mainly reversible over a 1.5-year period. Whether the improved cognitive performance at follow-up represents full or partial recovery from burnout is, however, a question that cannot be answered based on the present data. For example, the rather brief nature of the tests, like most other neuropsychological tests, might provide an opportunity to compensate for fatigue through a temporary extra effort, which cannot be sustained in tests of longer duration. This and other issues will be addressed in a forthcoming paper presenting results from a more comprehensive neuropsychological assessment of former burnout patients.

Acknowledgments

This study was supported by the Swedish Council for Working Life and Social Research (2006-0213). Thanks are due to Birgitta Pålsson, B.Sc., for interviews, to Jonatan Axelsson, M.D., for medical assessments, and to Lisbeth Prahl, for administrative work.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Beck AT, Steer RA, Garbin M. 1988. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clin Psychol Rev 8:77–100.

Carlsson F, Persson R, Karlson B, Österberg K, Hansen AM, Garde AH, Ørbæk P. 2006. Salivary cortisol and self-reported stress among persons with environmental annoyance. Scand J Work Environ Health 32:109–120.

Carter N, Iregren A, Söderman E, Anshelm Olson B, Karlson B, Lindelöf B, Lundberg I, Österberg K. 2002. EUROQUEST—a questionnaire for solvent related symptoms. Factor structure, item analysis, and predictive validity. Neurotoxicology 23: 711–717.

Chouanière D, Cassitto MG, Spurgeon A, Verdier A, Gilioli R. 1997. An international questionnaire to explore neurotoxic symptoms. Environ Res 73:70–72.

Cronholm B, Molander L. 1957. Memory disturbances after electroconvulsive therapy. 1. Conditions six hours after electroshock treatment. Acta Psychiatr Neurol Scand 32: 280–306.

De Vente W, Olff M, Van Amsterdam JG, Kamphuis JH, Emmelkamp PM. 2003. Physiological differences between burnout patients and healthy controls: Blood pressure, heart rate, and cortisol responses. Occup Environ Med 60(Suppl. 1): 54–61.

Derogatis LR. 1992. SCL-90-R. Administration, scoring & procedures. Manual-II. Towson, MD: Clinical Psychometric Research Inc.

Dureman I, Kebbon L, Österberg E. 1971. *Manual till DS-batteriet*. (Manual for the DS battery.) [in Swedish]. Stockholm, Sweden: Psykologiförlaget.

Garde AH, Hansen ÅM, Nikolajsen TB. 2003. An inter-laboratory comparison for determination of cortisol in saliva. Accred Qual Assur 8:16–20.

Grossi G, Perski A, Ekstedt M, Johansson T, Lindström M, Holm K. 2005. The morning salivary cortisol response in burnout. J Psychosom Res 59:103–111.

Grossi G, Perski A, Evengard B, Blomkvist V, Orth-Gomer K. 2003. Physiological correlates of burnout among women. J Psychosom Res 55:309–316.

Hansen ÅM, Garde AH, Christensen JM, Eller N, Netterstrøm B. 2003. Evaluation of a radioimmunoassay and establishment of a reference interval for salivary cortisol in healthy subjects in Denmark. Scand J Clin Lab Invest 63:303–310.

Hellhammer J, Fries E, Schweisthal OW, Schlotz W, Stone AA, Hagemann D. 2007. Several daily measurements are necessary to reliably assess the cortisol rise after awakening: State- and trait components. Psychoneuroendocrinology 32:80–86.

Kaplan E, Fein D, Morris R. 1991. WAIS-R NI manual. San Antonio, TX: The Psychological Corporation.

Karlson B, Jönsson P, Pålsson B, Åbjörnsson G, Malmberg B, Larsson B, Österberg K. 2010. Return to work after a workplace-oriented intervention for patients on sick-leave for burnout—a prospective controlled study. BMC Public Health 10:301.

Karlson B, Österberg K, Ørbæk P. 2000. Euroquest – the validity of a new symptom questionnaire. Neurotoxicology 21: 783–790.

Levander S, Elithorn A. 1987. The automated psychological test system. Manual. Trondheim, Norway: Department of Psychiatry and Behavioral Medicine, University of Trondheim.

Link K, Moëll C, Österberg K, Persson R, Ørbæk P, Garwicz S, Cavallin-Ståhl E, Erfurth EM. 2006. Adult survivors of childhood acute lymphoblastic leukaemia with GH deficiency have normal self-rated quality of life but impaired neuropsychological performance 20 years after cranial irradiation. Clin Endocrinol 65:617–625.

Maslach C, Jackson SE, Leiter MP. 1996. Maslach burnout inventory manual. Palo Alto, CA: Consulting Psychologists Press, Inc.

Maslach C, Schaufeli WB, Leiter MP. 2001. Job burnout. Annu Rev Psychol 52:397–422.

McEwen BS. 2007. Physiology and neurobiology of stress and adaptation: Central role of the brain. Physiol Rev 87: 873–904.

Melamed S, Ugarten U, Shirom A, Kahana L, Lerman Y, Froom P. 1999. Chronic burnout, somatic arousal and elevated salivary cortisol levels. J Psychosom Res 46:591–598.

Mickey RM, Greenland S. 1989. The impact of confounder selection criteria on effect estimation. Am J Epidemiol 129: 125–137.

Miller GE, Chen E, Zhou ES. 2007. If it goes up, must it come down? Chronic stress and the hypothalamic-pituitary-adrenocortical axis in humans. Psychol Bull 133:25–45.

Milner B. 1965. Visually guided maze learning in man: Effects of bilateral hippocampal, bilateral frontal, and unilateral cerebral lesions. Neuropsychologia 3:317–338.

Moch SL, Panz VR, Joffe BI, Havlik I, Moch JD. 2003. Longitudinal changes in pituitary-adrenal hormones in South African women with burnout. Endocrine 21:267–272.

Mommersteeg PM, Heijnen CJ, Verbraak MJ, van Doornen LJ. 2006a. Clinical burnout is not reflected in the cortisol awakening response, the day-curve or the response to a low-dose dexamethasone suppression test. Psychoneuroendocrinology 31:216–225.

Mommersteeg PM, Heijnen CJ, Verbraak MJ, van Doornen LJ. 2006b. A longitudinal study on cortisol and complaint reduction in burnout. Psychoneuroendocrinology 31:793–804.

Öhman L, Nordin S, Bergdahl J, Slunga Birgander L, Stigsdotter Neely A. 2007. Cognitive function in outpatients with perceived chronic stress. Scand J Work Environ Health 33: 223–232.

Österberg K, Karlson B, Hansen ÅM. 2009. Cognitive performance in patients with burnout, in relation to diurnal salivary cortisol. Stress 12:70–81.

Österberg K, Ørbæk P, Karlson B, Bergendorf U, Seger L. 2000. A comparison of neuropsychological tests for the assessment of chronic toxic encephalopathy. Am J Ind Med 38: 666–680.

Pruessner JC, Hellhammer DH, Kirschbaum C. 1999. Burnout, perceived stress and cortisol responses to awakening. Psychosom Med 61:197–204.

Rydmark I, Wahlberg K, Ghatan PH, Modell S, Nygren A, Ingvar M, Åsberg M, Heilig M. 2006. Neuroendocrine, cognitive and structural imaging characteristics of women on longterm sick leave with job stress-induced depression. Biol Psychiatry 60: 867–873.

Sandström A, Nyström Rhodin I, Lundberg M, Olsson T, Nyberg L. 2005. Impaired cognitive performance in patients with chronic burnout syndrome. Biol Psychol 69:271–279.

Sandström A, Peterson J, Sandström E, Lundberg M, Nystrom IL, Nyberg L, Olsson T. 2011. Cognitive deficits in relation to personality type and hypothalamic-pituitary-adrenal (HPA) axis

dysfunction in women with stress-related exhaustion. Scand J Psychol 52:71–82.

Socialstyrelsen. 2003. *Utmattningssyndrom*. (Exhaustion disorder.) [in Swedish]. Stockholm, Sweden: Bjurner & Bruno. http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/10723/2003-123-18_200312319.pdf.

Sonnenschein M, Mommersteeg PM, Houtveen JH, Sorbi MJ, Schaufeli WB, Van Doornen LJ. 2007. Exhaustion and endocrine functioning in clinical burnout: An in-depth study using the experience sampling method. Biol Psychol 75: 176–184.

Spitzer RL, Williams JB, Kroenke K, Linzer M, DeGruy FV, Hahn SR, Brody D, Johnson JG. 1994. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. J Am Med Assoc 272:1749–1756.

Van der Linden D, Keijsers GPJ, Eling P, Van Schaijk R. 2005. Work stress and attentional difficulties: An initial study on burnout and cognitive failures. Work Stress 19:23–36.

Wahlberg K, Ghatan PH, Modell S, Nygren A, Ingvar M, Åsberg M, Heilig M. 2009. Suppressed neuroendocrine stress response in depressed women on job-stress-related long-term sick leave: A stable marker potentially suggestive of preexisting vulnerability. Biol Psychiatry 65:742–747.

Walsh K. 1985. Understanding brain damage. Edinburgh: Churchill Livingstone.

Wechsler D. 1981. Wechsler adult intelligence scale—revised manual. San Antonio, TX: The Psychological Corporation.

Westgard JO, Barry PL, Hunt MR, Groth T. 1981. A multi-rule Shewhart chart for quality control in clinical chemistry. Clin Chem 27:493–501.