

Cortisol awakening response and subsequent depression: prospective longitudinal study

Rebecca Carnegie, Ricardo Araya, Yoav Ben-Shlomo, Vivette Glover, Thomas G. O'Connor, Kieran J. O'Donnell, Rebecca Pearson and Glyn Lewis

Background

Some studies have found an association between elevated cortisol and subsequent depression, but findings are inconsistent. The cortisol awakening response may be a more stable measure of hypothalamic–pituitary–adrenal function and potentially of stress reactivity.

Aims

To investigate whether salivary cortisol, particularly the cortisol awakening response, is associated with subsequent depression in a large population cohort.

Method

Young people (aged 15 years, $n=841$) from the Avon Longitudinal Study of Parents and Children (ALSPAC) collected salivary cortisol at four time points for 3 school days. Logistic regression was used to calculate odds ratios for developing depression meeting ICD-10 criteria at 18 years.

Results

We found no evidence for an association between salivary

cortisol and subsequent depression. Odds ratios for the cortisol awakening response were 1.24 per standard deviation (95% CI 0.93–1.66, $P=0.14$) before and 1.12 (95% CI 0.73–1.72, $P=0.61$) after adjustment for confounding factors. There was no evidence that the other cortisol measures, including cortisol at each time point, diurnal drop and area under the curve, were associated with subsequent depression.

Conclusions

Our findings do not support the hypothesis that elevated salivary cortisol increases the short-term risk of subsequent depressive illness. The results suggest that if an association does exist, it is small and unlikely to be of clinical significance.

Declaration of interest

None.

Research into the neurobiology of depressive illness has frequently implicated the hypothalamic–pituitary–adrenal (HPA) axis, which has been found to be abnormal in some people with both current^{1,2} and recently remitted^{2,3} depression. Elevated cortisol levels have also been found in cross-sectional studies of individuals at high risk for depression,^{4–9} suggesting that HPA abnormalities may play a role in the aetiology of depressive illness. Given the high prevalence of depressive symptoms in hypercortisolaemic states, both pathological as in Cushing's syndrome,¹⁰ as well as iatrogenic such as in long-term exogenous steroid use,¹¹ it seems plausible that abnormal HPA function could increase the risk of subsequent depression.

The increasing availability of salivary cortisol assays has facilitated studies of HPA function on a larger scale. This has enabled a handful of studies to analyse baseline cortisol levels in individuals followed longitudinally to assess the onset of depression.^{12–18} Abnormalities in baseline cortisol levels among individuals developing subsequent depression have strengthened support for the hypothesis that the HPA axis has an aetiological role in depression, and that elevated cortisol levels could increase the risk of developing depressive illness.^{13,17} There are a number of hypotheses for the mechanisms underlying this association,¹⁹ and it is difficult to specify precisely how HPA axis abnormalities might increase risk of depression, or exactly which cortisol measures might confer this risk. The lack of a clear hypothesis has led to a mixture of cortisol measures and methods for analysis being used. The consistent replication of results is a particular problem.^{20,21} Methodological differences, including variation in salivary collection protocols, sample populations, outcome measures and varying adjustment for confounding factors²² may account for some of the inconsistency. It may also be a consequence of the inherent difficulties in reliably measuring an

endocrinological system with marked diurnal as well as pulsatile variations.¹⁹ Some have also suggested that the relationship between cortisol and depression may be U-shaped rather than linear, further complicating analysis.¹⁸ Finally, it is possible that modest sample sizes combined with a wide range of potential ways to analyse cortisol may have increased type 1 errors. Using a more comprehensive approach to cortisol assessment and a larger sample size may enable a more thorough and inclusive analysis of HPA function, including the consideration of non-linear relationships. A cortisol awakening response (CAR) is increasingly regarded as a useful measure in cortisol research. The CAR is a normal physiological response to waking, and potentially captures information about stress reactivity in a minimally invasive way.^{19,23} There is also evidence that it is a relatively stable measure, with evidence of heritability.²⁴ The CAR may therefore be a particularly useful measure for prospective research in depression.

A handful of studies have looked at CAR in cross-sectional studies of individuals considered to be at high risk of depression. These studies have shown an elevated CAR in individuals with a family history of depression,^{6,8} and in those with high hopelessness reactivity.⁹ Conversely, a depressed CAR has been associated with other psychological vulnerabilities to depression such as rumination.²⁵ To our knowledge, only one study has used a prospective longitudinal design to assess whether CAR at baseline is associated with subsequent onset of depressive illness.¹⁷ In that study of 230 healthy participants, there was some evidence that participants who went on to develop depression at 1 year had a higher CAR at baseline than those not developing depression (odds ratio (OR) = 2.96 per standard deviation, 95% CI 1.06–8.26, $P=0.04$). Unfortunately, despite oversampling of individuals with symptoms of neuroticism to increase the incidence of depression, only 18 participants developed depression in the

cohort. Continuing follow-up of these participants found decreasing evidence for an association between CAR and subsequent depression over time.²⁶ It is possible that the abnormalities may be early biological changes of an emerging or subclinical depressive process, rather than pre-existing markers observable in healthy individuals. Further studies using longer follow-up periods would help to clarify the association between CAR and subsequent depression. In our study we investigated whether CAR was associated with subsequent depression in a longitudinal study of a population cohort of British adolescents. We also examined other salivary cortisol measures used in previous HPA research. To our knowledge, this is the largest longitudinal study using a comprehensive cortisol protocol to date. Using an established population birth cohort has the advantage of being able to account for confounding factors without recall bias, as well as maximising on the proven collaboration of participants, which have both been suggested as limitations in previous cortisol research.^{19,27}

Method

Participants

The Avon Longitudinal Study of Parents and Children (ALSPAC) is an ongoing prospective longitudinal study of children born in the former county of Avon, UK.^{28,29} Pregnant women with expected delivery dates between 1 April 1991 and 31 December 1992 were recruited, with catch-up recruitment at age 7. The complete ALSPAC sample consists of 14 701 children alive at 1 year. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the local research ethics committees.

Participants were recruited from attendees of a research clinic at 15 years, which assessed 5501 of the ALSPAC participants (online Fig. DS1). Around 3020 participants were invited at random to participate in the cortisol data collection, 1175 declined. Of the 1845 participants agreeing to enter the study, 1035 returned their saliva sampling packs. However, 75 participants provided empty cortisol samples and 119 provided insufficient data for analysis of CAR. Of the remaining 841 participants, 173 participants had no outcome data for depression at 18, either because of non-attendance at the follow-up research clinic ($n=122$) or non-completion of the Revised Clinical Interview Schedule (CIS-R)³⁰ diagnostic tool at the follow-up clinic ($n=51$).

This left 668 participants (22.1% of invited) with both exposure and outcome measure at 18, on which the primary analysis was undertaken. For 208 participants there were one or more covariates missing from the ALSPAC data-set, leaving 460 complete cases. Sociodemographic characteristics of active participants were compared with the invited participants, using chi-squared tests to evaluate any selective attrition. Non-participants were defined as those randomly invited to participate at the baseline clinic but either declining to participate or failing to provide sufficient data.

Cortisol sampling

As described in detail elsewhere,³¹ each participant agreeing to take part in the study was given a cortisol sampling pack during their clinic visit. Each pack contained detailed sampling instructions, 12 Salivette collection devices (Sarstedt, Germany), a sample collection diary sheet and a pre-paid envelope in which to return the samples to the laboratory. Participants were instructed to provide four samples per day over 3 consecutive school days, and to record the time of the sample. The samples were collected immediately on waking, at 30 min after waking, after school and before bed. To avoid contamination, participants

were instructed not to eat or brush their teeth for at least 30 min prior to collection. On completion, samples were returned to the laboratory and stored at -20°C . Salivary cortisol was determined using a commercially available enzyme immune assay. Inter-assay and intra-assay variation were 7.9 and 8.9 respectively.

Cortisol samples were excluded if they were greater than 82 nmols/l (the upper limit of the cortisol assay) or greater than 4 standard deviations above the mean. Individual samples were also excluded if the timing deviated from the specified protocol. For the CAR, samples were excluded if taking the initial sample after 10.00 h or taking the second sample less than 20 or more than 60 min after the first. For the mean levels at each time point, the second sample was excluded if taken after 11.00 h, the 'after school' samples were excluded if before 15.00 h or after 18.00 h, and the final 'before bed sample' was excluded if before 19.00 h. Although previous literature suggested an absent CAR in a proportion of the population, a secondary analysis excluded individuals with an absent or negative CAR, in case this was the result of measurement error or non-adherence.

The primary cortisol measure of interest was the CAR, derived using the difference between the two morning cortisol samples. Additionally, the mean cortisol level at each time point, the estimated daily secretion as given by the area under the curve and the diurnal drop in cortisol, and estimated total morning cortisol were analysed. The diurnal drop in cortisol was taken as the difference between mean morning and bedtime cortisol levels. Total morning cortisol was calculated using the sum of the two morning samples. Each cortisol variable was calculated as a mean of the available samples over the study period.

Confounding factors

Potential confounders were decided *a priori* based on the available literature. Confounders available from assessment at 15 years included current depressive symptoms (using bands of the Development and Well-Being Assessment, DAWBA)³² body fat percentage, oral contraceptive use, smoking and alcohol use. Historical assessments also provided potentially confounding background characteristics such as gender, social class, birth weight, gestation, prior depressive symptoms (using continuous scores of the Short Mood and Feelings Questionnaire (SMFQ)³³ assessed at 10, 13 and 14 years), maternal education, maternal history of depression and exposure to postnatal depression (assessed using the Edinburgh Postnatal Depression Scale, EPDS).³⁴

Following the development of a full set of covariates, certain variables were dropped from further consideration as they were neither associated with the cortisol measures nor had an impact on the effect size. This also helped to minimise numbers lost through incomplete data. This subset included anxiety symptoms at 15 years, neuroticism, cannabis use, friendships and social support, early significant life events, hours of sleep and income. Additionally, to avoid overadjustment and preserve sample size, groups of similar potential confounders were reduced to a single representative variable, with the others dropped from the final model. This subset included baseline fitness level and exercise, which was replaced by adiposity alone and antenatal depressive and anxiety symptoms, which was replaced by postnatal depressive symptoms alone.

Finally, because of the composition of the sample, other variables with negligible prevalence were dropped as covariates. First, as none of the individuals in the sample population were pre-pubertal (taken as a Tanner stage <2), pubertal stage was dropped as a confounder. Similarly, as only three participants reported steroid use prior to the study period, none of whom developed depression on follow-up, this was also dropped from the final model.

Outcome measures

The main outcome measure was ICD-10 diagnosis of depression,³⁵ which was assessed using the CIS-R³⁰ at a research clinic attended by participants at a mean age of 17.7 years. Responses to the questions about 14 symptom groups, as well as their onset and duration, were given via computer self-assessment at a clinic visit. The five symptom groups used to assess depression were sleep problems, poor concentration, fatigue and depression (which were scored on a scale of 0–4) and depressive thoughts (scored from 0 to 5). Responses to questions were aggregated by the specified algorithm to make a binary variable of ICD-10 diagnosis of depression.

A secondary analysis was undertaken using a binary cut off of 'depressive symptoms', to account for subclinical symptoms of depression. A 'depressive symptoms' score was derived from the sum of the five depressive symptom groups in the CIS-R. This score ranged from 0 to 21. A score of ≥ 8 depressive symptoms was chosen, aiming for a prevalence of $\sim 15\%$. Secondary analysis also looked at depressive symptoms after 1 year using a cut-off of ≥ 11 on the SMFQ,³³ assessed by postal questionnaire at a mean age of 16.7 years.

Statistical analysis

Statistical analysis was performed using Stata 12 on Windows. Logistic regression was used to calculate odds ratios between each cortisol variable and the subsequent development of depression. As CAR measured on a single day has been suggested to be related to current circumstances rather than a stable interpersonal state

marker,³⁶ the number of days used to calculate a mean CAR was included as a covariate. The waking cortisol level was also included as a covariate as this has been shown to affect the size of the CAR.³⁷

The possibility of a non-linear relationship between cortisol and depression was investigated using a quadratic term. We postulated *a priori* that there may be an interaction between gender and CAR and tested this using an appropriate interaction term in the regression model. An interaction term in the regression model was also used to investigate whether the association between cortisol and subsequent depression differed for participants with previous depression, baseline depressive symptoms or those with a positive family history of depression.

Missing data

Multiple imputation using chained equations (MICE) was used to address missing data.³⁸ The comprehensiveness of data collection in this cohort means that the missing values were assumed to be dependent on other observed data. The MICE procedure was used to address data missing on two levels: first, participants with one or more missing confounding factors had this data imputed; and second, a depression outcome at 18 years was imputed using previous measures of depression in earlier adolescence. The data were imputed using the 'ice' command of Stata, using 45 imputation models set at 20 cycles. The imputation model included all variables identified as potential confounders, including those discarded from the final analysis model. Additionally, previously

Table 1 Differences in characteristics of participants with exposure and outcome data compared with those invited to participate from baseline assessment clinic

	Non-participants (<i>n</i> = 2363)	Participants with exposure and outcome (<i>n</i> = 668)	χ^2	<i>t</i> -test	<i>P</i>
Gender, <i>n</i> (%)					
Male	1159 (49.1)	301 (45.1)	3.3		0.07
Female	1204 (51.0)	367 (54.9)			
Maternal education, <i>n</i> (%)					
< O-level	426 (19.4)	99 (15.5)	7.2		0.03
O-level	747 (34.0)	247 (38.7)			
A-level or higher	1022 (46.6)	293 (45.8)			
Unknown	168	29			
Social class, ^a <i>n</i> (%)					
I	75 (3.6)	30 (4.8)	4.7		0.32
II	563 (26.7)	177 (28.4)			
III	1133 (53.8)	327 (52.5)			
IV	281 (13.3)	70 (11.2)			
V	54 (2.6)	19 (3.1)			
Unknown	257	45			
Ethnicity of child, <i>n</i> (%)					
White	2075 (96.1)	605 (96.2)	0.0		0.93
Black and minority ethnic	84 (3.9)	24 (3.8)			
Unknown	204	39			
Depression at 15 years, <i>n</i> (%)					
Yes	222 (9.5)	42 (6.3)	6.6		0.01
No	2118 (90.5)	625 (93.7)			
Unknown	23	1			
Previous depression, <i>n</i> (%)					
Yes	345 (14.8)	82 (12.4)	2.5		0.12
No	1980 (85.2)	579 (87.6)			
Unknown	38	7			
Depression at 18 years, <i>n</i> (%)					
Yes	108 (7.3)	46 (6.9)	0.1		0.74
No	1376 (92.7)	622 (93.1)			
Unknown	879				
Body fat, %: mean (s.d.)	22.2 (10.3)	23.2 (10.3)		−2.2	0.03

a. I, professional occupations; II, managerial and technical; III, skilled occupations (includes manual and non-manual); IV, partly-skilled occupations; V, unskilled occupations.

collected values of the missing variables and background characteristics of the sample previously shown to affect missing data were included in the imputation model (Fig. DS1). Non-normally distributed variables (such as adiposity) were log-transformed to allow imputation. Highly skewed variables such as maternal measures of depression and previous depressive symptom scores were imputed using prediction matching. Categories with very small numbers (such as DAWBA bands 4 and 5, more than three siblings, household income bands, and frequency of exercise, cannabis and alcohol use) were collapsed. Monte Carlo errors were <10% of standard errors and the maximum fraction of missing information value was 0.43.³⁸

Results

Descriptive statistics

The characteristics of participants actively taking part in the cortisol study are shown in Table 1. Active study participants were more likely to be female, have higher maternal education and a lower prevalence of depression at baseline than non-participants. Distribution of social class, ethnicity, previous depression scores and prevalence of depression at 18 (where this was known) differed little between the groups.

Descriptive statistics for cortisol are given in Table 2. Around 75% of participants showed a rise in their morning cortisol, with

	<i>n</i>	Mean	s.d.	Range	Interquartile range	Mean sample time
Cortisol awakening response						
All participants	841	3.84	4.71	−14.60 to 26.11	1.04 to 6.40	0 h 35 min
With depression	46	5.07	4.61	−12.34 to 26.11	1.28 to 7.93	
Without depression	622	3.87	5.52	−4.77 to 24.53	1.13 to 6.40	
Waking cortisol						
All participants	887	7.62	3.72	0.03 to 26.74	5.10 to 9.22	07.03 h
With depression	49	7.11	3.03	2.38 to 17.53	4.75 to 9.22	
Without depression	659	7.59	3.80	0.29 to 26.74	5.08 to 9.05	
Cortisol sample 2						
All participants	891	11.30	5.20	0.44 to 33.75	7.81 to 14.25	07.41 h
With depression	49	11.85	5.70	2.79 to 33.06	7.82 to 15.10	
Without depression	665	11.29	5.24	0.44 to 33.75	7.78 to 14.15	
After school cortisol						
All participants	868	2.40	1.60	0.00 to 11.20	1.40 to 2.95	16.30 h
With depression	45	2.16	1.33	0.39 to 7.61	1.48 to 2.39	
Without depression	646	2.39	1.53	0.00 to 11.20	1.37 to 2.96	
Bedtime cortisol						
All participants	930	1.19	1.38	0.00 to 12.22	0.46 to 1.33	21.49 h
With depression	49	0.92	0.81	0.00 to 4.35	0.39 to 1.16	
Without depression	687	1.15	1.27	0.00 to 11.09	0.45 to 1.30	
Diurnal drop						
All participants	897	8.41	3.79	−2.30 to 28.46	5.95 to 10.50	16 h 5 min
With depression	49	8.57	3.37	2.12 to 17.55	6.65 to 10.85	
Without depression	687	8.42	3.89	−2.30 to 28.46	5.94 to 10.49	
Total daily cortisol (area under the curve, g)						
All participants	721	12275	4581	116 to 37811	9060 to 14885	
With depression	39	11388	4068	3712 to 26646	8813 to 14005	
Without depression	540	12221	4504	116 to 32485	8976 to 14811	

Attribute	<i>n</i>	CAR quartile, <i>n</i> (%)				<i>P</i>
		1 (<i>n</i> = 210)	2 (<i>n</i> = 210)	3 (<i>n</i> = 210)	4 (<i>n</i> = 211)	
Baseline and prior depression						
With depression at baseline	839	8.1	3.8	4.8	11.0	0.22
History of depression	832	11.1	12.9	12.4	13.0	0.60
Background characteristics						
Female	841	43.3	42.9	56.7	75.8	<0.01
Social class 4 or 5	774	15.3	12.6	15.6	15.3	0.80
Post O-level maternal education	798	42.0	49.5	41.7	48.7	0.43
Maternal history of depression	791	7.7	7.0	3.6	4.0	0.05
Exposed to postnatal depression	834	16.7	15.9	17.3	12.4	0.31
Low birth weight	809	3.0	5.9	5.0	3.0	0.88
Preterm birth (≤32 weeks)	841	0.5	1.0	2.4	0.0	1.00
Lifestyle factors						
Highest body fat quartile	837	22.0	22.6	24.9	40.3	<0.01
Taking oral contraceptive pill	811	4.4	3.9	4.5	5.5	0.58
Current daily smoker	828	3.5	4.8	1.4	6.5	0.41
Alcohol >2 days per week	704	2.9	2.2	4.1	7.3	0.03

a. Table includes all participants with a CAR, including those with missing outcome data.

Table 4 Odds ratios (ORs) for depression at 18 years for each standard deviation increase in cortisol awakening response (CAR), total morning cortisol, mean waking and bedtime cortisol, diurnal drop and area under the curve^a

	Unadjusted						Complete cases						Cumulative adjustments									
	OR (95% CI)			P			OR (95% CI)			P			Baseline and prior depression ^b			Background characteristics ^c			Lifestyle factors ^d			
	OR	n	P	OR	n	P	OR	n	P	OR	n	P	OR	n	P	OR	n	P	OR	n	P	
CAR	1.24	668	0.14	1.18	460	0.38	1.19	460	0.28	1.19	460	0.34	1.19	460	0.34	1.12	460	0.61	460	0.61	460	
CAR (excluding flat or negative results)	1.25	562	0.12	1.21	387	0.29	1.17	387	0.32	1.17	387	0.27	1.21	387	0.27	1.31	387	0.21	387	0.21	387	
Total early morning cortisol	1.49	668	0.10	1.40	460	0.27	1.38	460	0.21	1.43	460	0.22	1.43	460	0.22	1.30	460	0.47	460	0.47	460	
Mean waking cortisol	0.87	708	0.39	0.85	488	0.41	0.86	488	0.38	0.86	488	0.28	0.81	488	0.28	0.86	488	0.50	488	0.50	488	
Mean bedtime cortisol	0.77	708	0.25	0.74	488	0.29	0.76	488	0.24	0.76	488	0.12	0.64	488	0.12	0.57	488	0.09	488	0.09	488	
Diurnal drop	1.05	710	0.76	0.99	487	0.94	1.01	487	0.97	1.01	487	0.84	0.97	487	0.84	0.95	487	0.82	487	0.82	487	
Area under the curve, g	0.81	579	0.26	0.84	401	0.41	0.80	401	0.24	0.80	401	0.20	0.77	401	0.20	0.76	401	0.22	401	0.22	401	
Imputed data-sets																						
CAR ^e	1.24	668	0.14	1.22	668	0.19	1.22	668	0.19	1.22	668	0.52	1.11	668	0.52	1.12	668	0.51	668	0.51	668	
CAR ^f	1.20	841	0.24	1.20	841	0.24	1.17	841	0.30	1.17	841	0.68	1.07	841	0.68	1.09	841	0.61	841	0.61	841	

a. All results are adjusted for mean cortisol at Time 1 and number of CAR samples, except for results that are unadjusted.

b. Includes additional adjustment for depressive symptoms at 15 years and prior depression.

c. Includes additional adjustment for gender, social class, birth weight, gestation, maternal education, maternal history of depression and exposure to postnatal depression.

d. Includes further adjustments for oral contraceptive use, body fat, alcohol and smoking.

e. Data-set with imputed confounders.

f. Data-set with imputed confounders and outcome.

the remainder showing no response or a negative response after waking. The four daily samples displayed the expected diurnal variation. The CAR was normally distributed with very slight right deviation (skewness 0.58).

Confounding factors

Table 3 shows the characteristics of participants in each quartile group of CAR. The biggest differences were seen for gender, body fat, alcohol and maternal history of depression. There was a higher proportion of baseline depression in the top quartile, however this quartile also had a higher proportion of women.

Association with depression

Of the participants, 46 (6.89%) were diagnosed with depression at 18 years, 33 of whom were female (71.74%), and 34 of whom (73.91%) were considered 'new onset' as they had no previous recorded episodes.

Results of logistic regression for the main cortisol parameters of interest are provided in Table 4. There was no statistical evidence for an association between any of the cortisol measures and depression at follow-up, either before or after adjustment for confounding factors. The unadjusted odds ratio for depression at 18 was 1.24 per standard deviation of CAR (95% CI 0.93–1.66, $P=0.14$.) Adjusting for confounding factors and lifestyle factors decreased the effect size to 1.12 (95% CI 0.73–1.72, $P=0.61$). There was still no evidence for an association using the first day sample only (which gave the highest value of CAR) (OR = 1.09, 95% CI 0.80–1.48, $P=0.58$). A similar result was found when CIS-R 'depressive symptoms' was used as a binary outcome (OR = 1.07, 95% CI 0.78–1.48, $P=0.67$). Although there was weak evidence for an association between CAR and depressive symptoms at 16 years (OR = 1.22, 95% CI 0.99–1.51, $P=0.06$) this was markedly attenuated by confounding factors (online Table DS1). Adjusting for gender alone reduced the odds ratio to 1.11 (95% CI 0.89–1.37, $P=0.36$).

There was no evidence that any of the other cortisol parameters including the area under the curve, diurnal drop or total morning cortisol secretion were associated with subsequent depression. There was no evidence that the association between CAR and depression varied by gender ($P=0.69$), previous depression ($P=0.44$) or maternal history of depression ($P=0.67$). There was some evidence for a possible interaction between baseline depression score and CAR ($P=0.06$). However, the apparent interaction disappeared after excluding the 42 people with depression at baseline ($P=0.87$), suggesting a possible association between CAR and depression resolution or recurrence, but not with depression incidence.

Table 5 shows the characteristics of the cortisol awakening response divided into quartiles. The proportion with depression appears higher in the top and bottom quartiles, suggesting a U-shaped distribution. However, there was no statistical evidence for this possibility when investigated using a quadratic term ($P=0.38$). The relationships were attenuated following adjustment for gender, and for adiposity and alcohol use, which were also most common in the top quartile. These factors also provided the biggest attenuation of the relationship between cortisol and depression in the regression analyses.

Missing data

Imputation of missing confounders increased the final adjusted model numbers to 668. Imputation of missing outcome data increased the sample size to 841, with depression prevalence of 7.61%. Results of logistic regression using both imputed data-sets are provided in Table 4. Despite gains in statistical power, imputation of the missing data did not alter our findings.

Table 5 Cortisol (nmols/l) and depressive symptoms score given by quartile of cortisol awakening response (CAR)

CAR quartile	n	Cortisol mean (s.d.)	Depressive symptoms score, mean (s.d.)	% cases ^a	CIS-R depression (n = 668)		CIS-R depression after adjustments (n = 462)	
					OR (95% CI)	P	OR (95% CI)	P
1	210	-1.48 (2.59)	2.97 (3.69)	6.88	1.00		1.00	
2	210	2.17 (0.64)	3.10 (3.62)	5.95	0.77 (0.31–1.90)	0.578	0.47 (0.14–1.58)	0.224
3	210	4.75 (0.84)	3.22 (3.81)	5.33	0.68 (0.27–1.72)	0.578	0.57 (0.18–1.75)	0.324
4	211	9.88 (3.53)	3.31 (3.71)	9.36	1.27 (0.31–1.91)	0.564	0.85 (0.30–2.47)	0.772

CIS-R, Revised Clinical Interview Schedule.
a. Percentage of individuals with an ICD-10 diagnosis of depression based on the specified algorithm for the CIS-R.

Discussion

The findings from this large prospective cohort study do not support the hypothesis that the CAR, or any other basal salivary cortisol measures including cortisol at each time point, diurnal drop and area under the curve, increase the risk of developing a depressive illness. The small unadjusted association with depression at 16 years in this study may initially appear to corroborate previous suggestions of a time-limited association between the CAR and depression.²⁶ However, as adjusting for confounding factors nullifies this association, we find no evidence for a unique association between cortisol and depression at either time point.

Strengths and limitations

This study has several strengths including a large sample size, which is to our knowledge the largest prospective study using CAR to date. However, this study may still have lacked sufficient statistical power. Although the confidence intervals in this study do overlap those in previous research, the point estimate of the previous study (OR = 2.96) lies well outside our estimates here.¹⁷ The upper limit of our confidence interval (OR = 1.7) suggests that if a true difference in CAR does exist, it is small.

Recruitment from the ALSPAC cohort has enabled this study to consider a wealth of potential confounders. The inability of previous studies to account for these without inducing recall bias has been suggested as a potential factor underlying some of the inconsistency among previous studies.¹⁹ In our study there was some evidence that adjusting for certain confounding factors, such as alcohol use, would attenuate an apparent relationship with subsequent depression.

It is important to consider both selection and loss to follow-up bias as only a subset of participants provided cortisol samples and there was some attrition over time. This raises the concern that participants who are susceptible to depression, may be less likely to participate in or to complete the study protocol. However, there was a similar prevalence of depression at follow-up in this sample to the rest of the ALSPAC cohort. Additionally, controlling for potentially confounding factors such as social class and baseline depression (which also related to whether data were missing) would have minimised this bias. Finally, we investigated any influence missing data might have made using multiple imputation, and this had no overall impact on our results.

A further strength of this study is the comprehensive protocol for cortisol assessment, which exceeded a 'minimum protocol' for salivary cortisol collection in several ways.²³ It is possible that a two-sample protocol for CAR calculation may lose some sensitivity compared with a four-sample protocol.^{23,36} However, a two-sample protocol is well accepted in the literature, and any calculation of CAR in a prospective study has advantages over much of the literature to date. Nevertheless, measurement error is likely to have

occurred because of inaccurate reporting of sample collection times, potentially reducing the strength of an association.²⁷ This problem has led to some researchers starting to use automatic timing devices for salivary cortisol collection. However, there is no reason that poor adherence would occur more frequently than in previous studies, especially given the proven levels of cooperation of ALSPAC participants over the years. Measurement bias may also have been introduced by choosing point prevalences of depression at 16 and 18 years, although we do not think this would have introduced a major bias especially given the surge in depressive prevalence in this age group.³⁹

Implications

The findings from this large prospective study do not support the hypothesis that salivary cortisol is associated with the subsequent onset of depression at a population level. Our results suggest that if a true association exists, it is smaller than estimated by previous studies, is markedly attenuated by confounding factors and is unlikely to have any clinical significance. There are a number of reasons why this result appears to contradict previous study findings, which have implications for future research.

It is possible that abnormal HPA function may have a minor influence on the complex aetiology of depression, making the overall population difference small, or that HPA abnormalities are only found in a small subset of depression, which is no doubt a heterogeneous disorder.⁴⁰ Future prospective studies will need adequate power, requiring large sample sizes to detect small differences, or to investigate distinct subgroups.

The biological complexities of the HPA axis may also account for research inconsistencies. Salivary cortisol alone may be too crude a measure for prospective prediction and a more comprehensive approach may be required.⁴¹ Using laboratory-based tests of stress responsiveness, although impractical for studies of this size, may provide different parameters with which the association with depression becomes clearer.

At present a lack of clear hypothesis about which cortisol parameters are associated with future depressive illness may be causing spurious differences in cortisol being identified through multiple testing and selective reporting. Future research should publish both negative as well as positive results of all cortisol measures collected, to ensure a balanced view in literature and avoid publication bias. Furthermore, research to clarify any possible biological link between the HPA axis and depression, including mechanisms and potential parameters for measurement, will facilitate the improved design of any future epidemiological studies.

Funding

The UK Medical Research Council (Grant reference: 74882), the Wellcome Trust (Grant reference: 076467) and the University of Bristol provide core support for ALSPAC. This research was supported by NIH grant R01 MH073842 to T.G.O'C., which funded the cortisol collection; Wellcome Trust grant 084268/2/07/Z to G.L., which funded the CIS-R; and the Wellcome Trust Institutional Strategic Support grant 097825.

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

Rebecca Carnegie, BSc, MBChB, **Ricardo Araya**, MRCPsych, PhD, Academic Unit of Psychiatry, School of Social and Community Medicine, University of Bristol, UK; **Yoav Ben-Shlomo**, BSc, MBBS, MSc, PhD, FFPH, School of Social and Community Medicine, University of Bristol, UK; **Vivette Glover**, MA, PhD, DSc, Institute of Reproductive and Developmental Biology, Imperial College London, UK; **Thomas G. O'Connor**, PhD, Wynne Center for Family Research, Department of Psychiatry, University of Rochester Medical Center, Rochester, New York, USA; **Kieran J. O'Donnell**, PhD, Douglas Mental Health University Institute, Department of Psychiatry, McGill University, Canada; **Rebecca Pearson**, PhD, Academic Unit of Psychiatry, School of Social and Community Medicine, University of Bristol, UK; **Glyn Lewis**, PhD, Mental Health Sciences Unit, University College London, UK

Correspondence: Rebecca Carnegie, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK. Email: rebecca.carnegie@bristol.ac.uk

First received 14 Jan 2013, final revision 1 Aug 2013, accepted 5 Sep 2013

References

- Bhagwagar Z, Hafizi S, Cowen PJ. Increased salivary cortisol after waking in depression. *Psychopharmacology* 2005; **182**: 54–7.
- Vreeburg SA, Hoogendijk WJ, van Pelt J, Derijk RH, Verhagen JC, van Dyck R, et al. Major depressive disorder and hypothalamic-pituitary-adrenal axis activity: results from a large cohort study. *Arch Gen Psychiatry* 2009; **66**: 617–26.
- Bhagwagar Z, Hafizi S, Cowen PJ. Increase in concentration of waking salivary cortisol in recovered patients with depression. *Am J Psychiatry* 2003; **160**: 1890–1.
- Dougherty LR, Klein DN, Olino TM, Dyson M, Rose S. Increased waking salivary cortisol and depression risk in preschoolers: the role of maternal history of melancholic depression and early child temperament. *J Child Psychol Psychiatry* 2009; **50**: 1495–503.
- Young EA, Vazquez D, Jiang H, Pfeffer CR. Saliva cortisol and response to dexamethasone in the children of depressed parents. *Biol Psychiatry* 2006; **60**: 831–6.
- Vreeburg SA, Hartman CA, Hoogendijk WJG, van Dyck R, Zitman FG, Ormel J, et al. Parental history of depression or anxiety and the cortisol awakening response. *Br J Psychiatry* 2010; **197**: 180–5.
- Vinberg M, Bennike B, Kyvik KO, Andersen PK, Kessing LV. Salivary cortisol in unaffected twins discordant for affective disorder. *Psychiatry Res* 2008; **161**: 292–301.
- Mannie ZN, Harmer CJ, Cowen PJ. Increased waking salivary cortisol levels in young people at familial risk of depression. *Am J Psychiatry* 2007; **164**: 617–21.
- van Santen A, Vreeburg SA, Van der Does AJ, Spinhoven P, Zitman FG, Penninx BW. Psychological traits and the cortisol awakening response: results from the Netherlands Study of Depression and Anxiety. *Psychoneuroendocrinology* 2011; **36**: 240–8.
- Pereira AM, Tiemensma J, Romijn JA. Neuropsychiatric disorders in Cushing's syndrome. *Neuroendocrinology* 2010; **92** (suppl 1): 65–70.
- Patten SB. Exogenous corticosteroids and major depression in the general population. *J Psychosom Res* 2000; **49**: 447–9.
- Halligan SL, Herbert J, Goodyer I, Murray L. Disturbances in morning cortisol secretion in association with maternal postnatal depression predict subsequent depressive symptomatology in adolescents. *Biol Psychiatry* 2007; **62**: 40–6.
- Ellenbogen MA, Hodgins S, Linnen AM, Ostiguy CS. Elevated daytime cortisol levels: a biomarker of subsequent major affective disorder? *J Affect Disord* 2011; **132**: 265–9.
- Goodyer IM, Tamplin A, Herbert J, Altham PME. Recent life events, cortisol, dehydroepiandrosterone and the onset of major depression in high-risk adolescents. *Br J Psychiatry* 2000; **177**: 499–504.
- Harris TO, Borsanyi S, Messari S, Stanford K, Brown GW, Cleary SE, et al. Morning cortisol as a risk factor for subsequent major depressive disorder in adult women. *Br J Psychiatry* 2000; **177**: 505–10.
- Goodyer IM, Bacon A, Ban M, Croudace T, Herbert J. Serotonin transporter genotype, morning cortisol and subsequent depression in adolescents. *Br J Psychiatry* 2009; **195**: 39–45.
- Adam EK, Doane LD, Zinbarg RE, Mineka S, Craske MG, Griffith JW. Prospective prediction of major depressive disorder from cortisol awakening responses in adolescence. *Psychoneuroendocrinology* 2010; **35**: 921–31.
- Herbert J, Ban M, Brown GW, Harris TO, Ogilvie A, Uher R, et al. Interaction between the *BDNF* gene Val/66/Met polymorphism and morning cortisol levels as a predictor of depression in adult women. *Br J Psychiatry* 2012; **201**: 313–9.
- Pariante CM, Lightman SL. The HPA axis in major depression: classical theories and new developments. *Trends Neurosci* 2008; **31**: 464–8.
- Knorr U, Vinberg M, Kessing LV, Wetterslev J. Salivary cortisol in depressed patients versus control persons: a systematic review and meta-analysis. *Psychoneuroendocrinology* 2010; **35**: 1275–86.
- Burke HM, Davis MC, Otte C, Mohr DC. Depression and cortisol responses to psychological stress: a meta-analysis. *Psychoneuroendocrinology* 2005; **30**: 846–56.
- Lopez-Duran NL, Kovacs M, George CJ. Hypothalamic-pituitary-adrenal axis dysregulation in depressed children and adolescents: a meta-analysis. *Psychoneuroendocrinology* 2009; **34**: 1272–83.
- Adam EK, Kumari M. Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology* 2009; **34**: 1423–36.
- Wust S, Federenko I, Hellhammer DH, Kirschbaum C. Genetic factors, perceived chronic stress, and the free cortisol response to awakening. *Psychoneuroendocrinology* 2000; **25**: 707–20.
- Kuehner C, Holzhauer S, Huffziger S. Decreased cortisol response to awakening is associated with cognitive vulnerability to depression in a nonclinical sample of young adults. *Psychoneuroendocrinology* 2007; **32**: 199–209.
- Vrshek-Schallhorn S, Doane LD, Mineka S, Zinbarg RE, Craske MG, Adam EK. The cortisol awakening response predicts major depression: predictive stability over a 4-year follow-up and effect of depression history. *Psychol Med* 2013; **43**: 483–93.
- Halpern CT, Whitels EA, Wagner B, Harris KM. Challenges of measuring diurnal cortisol concentrations in a large population-based field study. *Psychoneuroendocrinology* 2012; **37**: 499–508.
- Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort profile: the 'Children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2013; **42**: 111–27.
- Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, et al. Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* 2013; **42**: 97–110.
- Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychol Med* 1992; **22**: 465–86.
- O'Donnell KJ, Glover V, Jenkins J, Browne D, Ben-Shlomo Y, Golding J, et al. Prenatal maternal mood is associated with altered diurnal cortisol in adolescence. *Psychoneuroendocrinology* 2013; **38**: 1630–8.
- Goodman A, Heiervang E, Collishaw S, Goodman R. The 'DAWBA bands' as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. *Soc Psychiatry Psychiatr Epidemiol* 2011; **46**: 521–32.
- Angold A, Costello EJ, Messer SC, Pickles A, Winder F, Silver D. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *Int J Method Psych* 1995; **5**: 237–49.
- Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-Item Edinburgh Postnatal Depression Scale. *Br J Psychiatry* 1987; **150**: 782–6.
- World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. WHO, 1992.
- Hellhammer J, Fries E, Schweisthal OW, Schlotz W, Stone AA, Hagemann D. Several daily measurements are necessary to reliably assess the cortisol rise after awakening: state- and trait components. *Psychoneuroendocrinology* 2007; **32**: 80–6.
- O'Connor TG, Ben-Shlomo Y, Heron J, Golding J, Adams D, Glover V. Prenatal anxiety predicts individual differences in cortisol in pre-adolescent children. *Biol Psychiatry* 2005; **58**: 211–7.
- White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011; **30**: 377–99.
- Moffitt TE, Harrington H, Caspi A, Kim-Cohen J, Goldberg D, Gregory AM, et al. Depression and generalized anxiety disorder - cumulative and sequential comorbidity in a birth cohort followed prospectively to age 32 years. *Arch Gen Psychiatry* 2007; **64**: 651–60.
- Antonijevic IA. Depressive disorders – is it time to endorse different pathophysiologies? *Psychoneuroendocrinology* 2006; **31**: 1–15.
- Hellhammer DH, Wust S, Kudielka BM. Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology* 2009; **34**: 163–71.