Cortisol as a predictor of psychological therapy response in depressive disorders: systematic review and meta-analysis

Susanne Fischer, Rebecca Strawbridge, Andres Herane Vives and Anthony J. Cleare

**Background**
Many patients with depressive disorders demonstrate resistance to psychological therapy. A frequent finding is hypothalamic–pituitary–adrenal (HPA) axis alterations. As cortisol is known to modulate cognitive processes, those patients may be less likely to profit from psychological therapy.

**Aims**
To conduct a systematic review and meta-analysis on cortisol as a predictor of psychological therapy response.

**Method**
The Cochrane Library, EMBASE, MEDLINE and PsycINFO databases were searched. Records were included if they looked at patients with any depressive disorder engaging in psychological therapy, with a pre-treatment cortisol and a post-treatment symptom measure.

Psychological therapies are among the most effective treatments for depressive disorders, with most evidence focusing on cognitive–behavioural therapy (CBT). However, around 50% of patients do not respond to psychotherapeutic interventions. Upon repeated non-response to a number of treatment modalities, those patients are sometimes referred to as having ‘treatment resistant’ illness. The illness course of this subgroup is most often relapsing or chronic; high levels of disability and mortality are common, and these patients make up the largest proportion of costs both direct (for example, treatment) and indirect (such as lost productivity). Among the most salient features of depressive disorders are alterations in the hypothalamic–pituitary–adrenal (HPA) axis, one of the major stress-responsive systems. Hypercortisolism (i.e. high circulating levels of cortisol) potentially because of impaired negative feedback sensitivity, is a common finding in depression. It is noteworthy that cortisol modulates cognitive processes. In patients with major depression, cortisol seems to be related to cognitive impairment, and may thus explain symptoms such as concentration and memory difficulties. As a consequence, these patients may be less able to engage in learning processes, such as during psychological therapy.

**Results**
Eight articles satisfied our selection criteria. The higher the cortisol levels before starting psychological therapy, the more symptoms patients with depression experienced at the end of treatment and/or the smaller their symptom change.

**Conclusions**
Our findings suggest that patients with depression with elevated HPA functioning are less responsive to psychological therapy.

**Declaration of interest**
None.

**Copyright and usage**
© The Royal College of Psychiatrists 2017.
relevant analyses. The reference sections of all articles were then searched for additional records.

Data extraction
For each identified study one of the study investigators (S.F.) collected information about the first author, its year of publication, number of treated patients, their gender and age, their primary diagnosis, eligibility criteria (for example comorbidity and medication), study design, type and intensity of psychological therapy, pre-treatment levels of cortisol in hair, urine, saliva or blood, post-treatment symptom scores of depression and adjustment for covariates. This information is reported in online Table DS1. When multiple cortisol measures were reported, we favoured hair, urinary and salivary cortisol over plasma cortisol. This was to obtain a measure of free (i.e. unbound cortisol, as found in these media) rather than total cortisol. Moreover, we were interested in obtaining stable measures, which is why long-term (for example 3 cm hair cortisol) or aggregate indices (such as areas under the curve) were preferred over single time-point assessments. Evening or afternoon measurements were to be preferred in preference to morning levels, as the diurnal rhythmicity of cortisol interindividual variability is lower later in the day. When psychological therapy response was assessed at multiple time-points, the assessment closest to the last therapy session was selected. Symptoms of depression were chosen as outcome variables. Whenever studies used multiple symptom measures as outcome variables (such as different questionnaires assessing depression), an average effect size (ES) was calculated and extracted. Risk of bias was assessed by means of an adapted version of a quality tool that was used in an earlier meta-analysis on the role of cortisol in functional somatic syndromes. We used eight of the original nine items (item two was excluded because of non-applicability) and scored them using the same three-point scale (0–2). An additional item regarding the duration (weeks) and frequency of psychological therapy (number of sessions) was added; if both were stated, the study was given a score of two, if duration or intensity was stated, it was given a score of one, and a score of zero was given to those not giving any information about either of these characteristics. The complete checklist can be found in online Table DS2. The maximum attainable quality score was 18.

Effect size calculation
To quantify the relationship between cortisol levels at baseline and psychological therapy response, correlation coefficients ($r$) were directly extracted by one of the study investigators (S.F.) or, if not available, calculated based on summary statistics, such as frequency tables, means, standard deviations, and sample sizes or $t$-values and sample sizes. This was done in accordance with the procedures outlined by Lipsey & Wilson. In brief, statistics are converted into target effect sizes (in this case $r$) by means of mathematical formulae. For instance, the standardised mean effect size $d$ is calculated based on means, standard deviations and sample sizes and $d$ in turn is converted to $r$ using $r = \frac{d}{\sqrt{d^2 + \frac{4}{n}}}$, with $a$ being a correction factor that is used in the case of unequal sample sizes, and $a = (n_1 + n_2)^2/n_1n_2$. Whenever correlation coefficients ($r$) were directly extracted, we used either correlations between cortisol values and symptoms of depression upon completion of treatment (controlled for pre-treatment levels of the same symptoms), or correlations between cortisol values and change in symptoms (delta values). We extracted data that were adjusted for relevant covariates rather than unadjusted data whenever possible. In cases where no or insufficient statistical parameters were reported regarding our research question, the authors were contacted. If we were unable to gather additional data from the authors and it was stated that no significant relationship between pre-treatment cortisol levels and response to psychological therapy was found, we included a conservative effect size of zero into our meta-analysis.

Statistical analysis
We calculated Fisher’s Zt and 95% confidence intervals for each study and weighed studies based on their sample size. Studies with an extreme effect size (below or above two standard deviations) were to be excluded. Separate analyses for basal cortisol and post-challenge measures were conducted. An aggregated effect size including a 95% confidence interval was calculated for each analysis, using SPSS 21 and the macros developed by David B. Wilson (http://mason.gmu.edu/~dwilsonb/ma.html). As significant statistical heterogeneity according to the Q and $I^2$ statistics was predicted in all analyses, random- rather than fixed-effects models were considered appropriate. Sensitivity analyses were applied for studies with unclear eligibility. Publication bias was planned to be examined by visual inspection of funnel plots. In addition, Egger’s test and a trim and fill procedure were to be used for quantification of publication bias.

Results

Search results
Our search yielded 25,991 records, of which 84 were considered potentially relevant based on their title or abstract. Of these, 76 were excluded because they were not original research (for example reviews), not conducted in patients with depression, did not assess cortisol levels, did not include any treatment at all, administered drugs as part of the treatment package, used specific interventions that did not meet our definition of psychological therapy (for example aerobic exercise), were retrospective (for example compared cortisol levels of patients whose condition was treatment resistant with healthy controls) or did not report predictor analyses. In one study, patients and healthy controls or patients receiving different kinds of treatments had been collapsed for statistical analyses. As the data-set of this study was not available, it was excluded. In one paper, the authors stated that no relationship between pre-treatment cortisol levels and psychological therapy response was found. As no standardised regression coefficients were available for this report, we imputed an effect size of 0 (95% CI −0.512 to 0.512). Two reports were likely to have patient overlap according to the authors, but as this could not be quantified, both were included into the initial analyses and assessed in sensitivity analyses. Another sensitivity analysis was conducted excluding the study that used clomipramine in a pharmacological challenge test of the HPA axis (all remaining studies used the dexamethasone challenge test). In total, eight studies were eligible for data extraction. Tables DS1 shows the characteristics of these studies.

Systematic review and meta-analysis
Eight studies were included with a total of 212 participants. Two studies focused on adolescents and one looked at older adults. The majority of patients had a major depressive disorder according to the DSM, and in most instances, patients had moderate to severe depression. Patients with major comorbid mental disorders, including substance use, psychotic and bipolar disorders, were mostly excluded. By contrast, patients with comorbid medical diseases were often still considered eligible, that is, if their illness was unlikely to affect their depression, HPA axis functioning or treatment itself. The handling of psychotropic medication was rather heterogeneous across studies: intake of

106

https://doi.org/10.1192/bjp.bp.115.180653 Published online by Cambridge University Press
medication was excluded a priori, discontinued shortly before treatment or kept stable over the course of treatment. The distribution of out-patients to in-patients was 50:50. All patients received cognitive therapy or CBT, except for the two studies of adolescents where patients were treated with interpersonal therapy. Out-patients were treated on a weekly basis for a total duration of 8–16 weeks. In-patients had three to five sessions of psychological therapy per week, lasting between 3 and 6 weeks. Basal levels of cortisol were determined in urine, blood or saliva. The earlier studies were primarily interested in post-challenge cortisol levels and used the dexamethasone suppression test to assess this. This test allows the assessment of negative feedback sensitivity by administering 1 mg of a synthetic analogue of cortisol; higher concentrations of cortisol on the day after dexamethasone administration indicate alterations (hyperactivity) in the HPA axis. Patients were often divided into non-responders and responders. One study challenged the HPA axis by administering the tricyclic antidepressant clomipramine, and here, lower cortisol concentrations were considered maladaptive. The Beck Depression Inventory and Hamilton Rating Scale for Depression were the most frequently employed therapy response measures. Again, some studies divided patients into groups of non-responders and responders.

Quality scores ranged from 7 to 13 out of a maximum of 18 points. The quality of the diagnostic assessment, stating of eligibility criteria, the description of psychological therapy and the statistical reporting were very satisfactory in general. On the other hand, virtually no information was provided on the duration of patients’ illness and on the masking of therapists and study personnel rating treatment responses to patients’ cortisol levels. The description of HPA axis assessment and the handling of confounders varied greatly.

Results were fairly consistent: half of the studies found that patients with higher basal or post-challenge cortisol levels retained higher levels of depression upon completion of psychological therapy. Two more studies were in line with these findings, although only with regard to some of the employed cortisol and depression measures. More specifically, in the study by Thase et al, only basal levels of cortisol (but not post-challenge levels) were associated with treatment response, whereas basal cortisol levels in the study by Holland et al only predicted change in depression levels when the three used outcome scales were combined to create a global depression index. Two studies did not find any association between cortisol levels and psychological therapy response. Subsequent meta-analysis confirmed the overall finding: a significant relationship between basal cortisol levels (five independent effect sizes, mean ES = 0.264, 95% CI 0.047–0.481, Z = 2.382, P = 0.017) and post-challenge cortisol levels (six independent effect sizes, mean ES = 0.421, 95% CI 0.095–0.748, Z = 2.528, P = 0.012) and response to psychological therapy emerged, indicating that higher cortisol levels pre-treatment were associated with more severe symptoms post-treatment or smaller changes in symptoms (see Figs. 1 and 2 for forest plots).

Significant heterogeneity was present regarding post-challenge cortisol levels (Q = 15.22, P = 0.010, I² = 67.15%), but not regarding basal levels (Q = 5.55, P = 0.235, I² = 27.93%). However, because of the low number of included studies, the latter result cannot be regarded as definitive evidence for the absence of heterogeneity. For the same reason, Egger’s test and the trim and fill procedure were not undertaken. Sensitivity analyses were conducted by repeating analyses without the study where an effect size had been imputed, without the study that was likely to have some degree of case overlap with a later report and without the one using clomipramine to challenge the HPA axis. Excluding the study that possibly overlapped with another included study rendered our previously significant association between basal cortisol levels and outcomes significant by trend only (four independent effect sizes, mean ES = 0.215, 95% CI −0.029 to 0.460, Z = 1.726, P = 0.084). However, the positive finding of post-challenge cortisol levels predicting treatment response remained significant and no other changes to our results occurred as a result of excluding the other two studies.

### Discussion

#### Main findings

The main finding of the present meta-analysis is that the higher the basal and post-challenge cortisol levels before starting psychological therapy, the more symptoms individuals with depression
experienced at the end of treatment and/or the smaller their symptom change. This finding is in line with our initial hypothesis that HPA axis alterations, in this case indicated by high cortisol levels, are associated with worse psychological treatment outcomes in people with depression. Importantly, basal levels of cortisol were associated with response to psychological therapy by trend only when one study was excluded that had potential patient overlap with another study. As it was not possible to determine whether this overlap actually occurred or not, this finding needs to be interpreted with caution. Cortisol in response to the dexamethasone suppression test, on the other hand, did predict treatment response in the present meta-analysis. This is interesting from a clinical point of view, as the dexamethasone test offers the possibility of assessing HPA axis integrity in a highly standardised manner. In addition, the dexamethasone suppression test can now be undertaken using salivary cortisol measures making it a more practical test that can be undertaken at home without the need for venepuncture or hospital attendance. Importantly, however, our findings may not be applicable to patients with atypical depression, who have been found to present with low rather than high cortisol values.

**Interpretation and integration of findings**

One explanation for our findings is evidence that HPA axis alterations may be linked with cognitive functioning. According to this line of reasoning, the more pronounced a person with depression’s HPA axis alterations, the more severe his/her cognitive impairment, which in turn renders him/her less capable of engaging in psychological therapy. Based on our findings, it may be sensible to combine psychological therapy with psychotropic medication, which specifically targets hypercortisolism, such as antidepressants25 or antiglucocorticoid agents (such as mifepristone).26 However, considering the low overall number of published studies, the possibility of publication bias, the age of certain studies and the heterogeneity detected, further research on cortisol as a predictor of psychological therapy responses is clearly warranted before treatment suggestions can be made based on the findings reported here.

Strong evidence suggests that HPA axis alterations in depression originate from experiences of childhood trauma and chronic stress.27,28 Apart from direct effects on treatment efficacy via cognitive impairment, hypercortisolism may thus indicate a specific subgroup of patients with depression, and one that has its aetiology in earlier life trauma or chronic stress.29,30 In fact, a recent meta-analysis showed that those patients with depression who reported childhood maltreatment had poorer responses across treatment modalities.31 Apart from the above-mentioned psychotropic agents, clinicians may therefore consider interventions that have a strong focus on trauma and/or interpersonal functioning, with the latter being an important source of chronic stress. Moreover, findings of a recent trial suggest that mindfulness-based cognitive therapy may be specifically useful for patients who report childhood trauma.32 Taken together, this line of evidence suggests it would be highly commendable for future studies on cortisol as a predictor of treatment response to stratify patients according to their levels of childhood trauma, and to test whether the findings we report here are driven by a specific subgroup of patients with depression.

**Limitations**

Several limitations need to be considered when interpreting our results. First, our search yielded very few suitable studies. As a result of this, publication bias could not be assessed. Second, in some studies, important information was not reported, such as whether the assessors of treatment response were unaware of the participants’ cortisol levels. Another point is that covariates (such as psychotropic medication and baseline levels of depression) influencing cortisol levels and/or its association with treatment response have not been consistently taken into account. Third, there was substantial heterogeneity in eligibility criteria, cortisol and treatment response assessment within the included studies. Unfortunately, the small number of studies prevented us from exploring which factors were most strongly related to positive findings.

**Implications and directions for future research**

Taken together, our findings suggest that pre-treatment cortisol concentrations may predict responses to psychological therapy in patients with depression. Accumulating evidence suggests that there may be a subtype of depression that has its origins in early life stress, which could take its toll on stress-responsive bodily systems, such as the HPA axis, and in turn mediate some aspects
of the treatment resistance seen in these patients. HPA axis markers are thus a promising avenue in research on treatment resistance. More studies that address the research question outlined in this meta-analysis are warranted. Future research may consider using long-term measures of HPA axis functioning, such as fingerprint or hair cortisol concentrations, that most accurately represent the patients’ pre-treatment state, controlling for confounding factors, in particular depression severity and childhood trauma, and using continuous response scores when evaluating treatment success.

Susanne Fischer, PhD, Rebecca Strawbridge, MSc, Department of Psychological Medicine, Centre for Affective Disorders, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, UK; Andres Herane Vives, MD, DFM, MSc, Universidad Católica del Norte, Coquimbo, Chile and Department of Psychological Medicine, Centre for Affective Disorders, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, UK.

Correspondence: Susanne Fischer, King’s College London, Institute of Psychiatry, Psychology & Neuroscience, Department of Psychological Medicine, Box P074, 103 Denmark Hill, London SE5 8AF, UK. Email: susanne.fischer@kcl.ac.uk

First received 23 Dec 2015, final revision 26 Jul 2016, accepted 23 Aug 2016

Funding

S.F. acknowledges funding by the Swiss National Science Foundation. R.S. is supported by the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre. A.H.V. is funded by the Comisión Nacional de Investigación Científica y Tecnológica and the Psychiatric Research Trust. A.I.C. is supported by the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre, South London and Maudsley NHS Foundation Trust and King’s College London. The views expressed are those of the authors and not necessarily those of the NIHR, the NHS or the Department of Health.

Acknowledgements

We thank the authors of the included studies for generously providing us with additional information on their studies, and the anonymous reviewers for their helpful comments and suggestions on earlier versions of this manuscript.

References


