„Are aberrant cortisol levels prognostic factors for the development of depression in the adult population? A systematic review study protocol and metaanalysis”

Daniela Stotz, Mareike Stieg, Anna Kopczak, Christina Dimopoulou, Sarah Leistner, Matthias Auer, Günter K Stalla and Caroline Sievers

Clinical Neuroendocrinology, Clinical Research Department, Max Planck Institute of Psychiatry, Kraepelinstrasse 2-10, 80804 Munich, Germany

Correspondence to: Caroline Sievers csievers@mpipsykl.mpg.de

Abstract

Introduction The role of cortisol as a non-invasive biomarker to predict depressive disorders is unclear until now. On the one hand, prospective observational studies suggest that individual differences in cortisol levels are a possible risk factor for the development and/or relapse of depression. On the other hand, various methods for measuring the cortisol levels (salivary, urine, hair, serum) at various timepoints and conditions (awakening cortisol, cortisol profiles, DEX-CRH test) and the use of different screening instruments for depression as well as different study designs allow no clear conclusion. Primary aim of this project is to systematically review the results of prospective cohort studies to determine whether cortisol levels predict the development of depression in the adult population.

Methods and analysis Through a systematic literature research (PubMed, Medline, PsychInfo, Embase), we will identify prospective cohort studies between March 2015 and April 2015 in which the relationship between cortisol levels and depression has been prospectively investigated. All studies will be taken into account that meet the predefined inclusion and none of the exclusion criteria. Studies will be included according to the PRISMA four-phase flow diagram. The results of two persons will be compared and discussed. Using the quality appraisal tool “Newcastle Ottawa Scale” for cohort studies the quality of primary studies will be assessed. A funnel plot will be created to check publication bias. If possible, a random-effects meta-analysis will be performed presenting results in a forest plot. Other forms of evidence synthesis (Text/Table/harvest plot) will be performed for those data where a meta-analysis cannot be performed. Incidence of depression defined according to ICD-10, DSM-IV or DSM-V criteria.

Ethics and dissemination Ethics approval is not required because this is a protocol for a systematic review. The results of this systematic review will be disseminated via peer-reviewed publications.
Background

About 38% of the EU population suffers from a mental disorder. Among these 164 million people that are affected each year, major depression is claimed to be the most frequent disorder (1). To prevent this disease it is essential to identify possible predictors to define patients at risk. Researchers identified risk factors for the development of depression by a European consortium leading to the development of the PredictD algorithm (2). Predictive factors that have been identified were age, gender, level of education, previous depressive episodes or familial predisposition (2). However, these factors are mostly unmodifiable. On the other hand, numerous researchers recognized hormonal risk factors, such as cortisol, testosterone, thyroid hormones, IGF-1 etc. as risk factors associated with depression (3-5).

One of the most studied predictive candidates is cortisol (6), as affected patients show hyperactivity of the hypothalamic-pituitary-adrenal (HPA) axis (7-8). At the same time depressive patients differ in terms of peripheral serum cortisol levels (9). On the basis of cross-sectional data it has been shown that early morning cortisol is increased in patients with depression when adjusted for the time of awakening (3). Other groups could demonstrate that urinary free cortisol values of depressive patients compared with those of healthy controls are significantly higher in a cross-sectional study design with 10 melancholic depressives and 15 healthy controls (10). A study in 42 students yielded that hair cortisol concentrations are associated with depressive symptoms, suggesting that hair cortisol measurement can serve as a biomarker for mental disorders (11). Furthermore other groups found that the plasma levels of cortisol are increased in patients with depression after the dexamethasone test (12-13). Based on results in our own institute, it has been suggested that altered cortisol levels measured with the DEX/CRH test at two different time points might be the best HPA axis related biomarker that can predict clinical outcome at follow-up (14). All these findings suggest that individual differences in cortisol levels and an altered HPA-axis are a possible risk factor for the development and/or relapse of depression.

On the other hand, some groups found contradictory results. Grynderup and colleagues, for instance, examined 4231 healthy people and found a relationship between the difference in morning and evening salivary cortisol concentration and the risk of depression while neither morning nor evening salivary cortisol levels alone predicted the risk for depression (15). This result contradicts findings by Harris and colleagues stating that individual differences in morning salivary cortisol levels (but not evening salivary cortisol levels) in females were considered a risk factor for major depressive disorder (16). Other researchers came to the conclusion that an U-shaped association between the adjusted morning cortisol levels predicts the risk for depression onset in 12-month follow-up best (17). According to a recent study, high cortisol values measured in late morning in women at age 45 years predicts subsequent depressive symptoms at age 50 years. In men, lower cortisol values were prospectively associated with the occurrence of depressive symptoms (18). In regard to depression relapse, Ising and colleagues found that an exaggerated cortisol response to the DEX/CRH test might be an important risk factor that predicts the recurrence of depressive disorder within the first 6 months. Especially the change in the cortisol response between admission and hospital discharge and the effect on cortisol of the CRH infusion compared to baseline levels before discharge had predictive value (19). In a further study, Hardevelt and colleagues examined whether different HPA-axis parameters like cortisol awakening response (CAR), dexamethasone suppression test (DST) and evening cortisol predict the recurrence of a depressive episode in remitted subjects. A total of 549 people with a lifetime diagnosis of Major Depression participated in this investigation. All of them were in remission for at least six months. The researchers found an association between the cortisol awakening response and the recurrence of depression. Only higher cortisol awakening response but not cortisol levels measured by the DST nor the evening cortisol levels were risk factors for recurrence of depression in their study (20).
In conclusion, various methods for measuring cortisol levels (salivary, urine, hair, serum) at various timepoints and under different conditions (awakening cortisol, cortisol profiles, DEX-CRH tests) and the use of different screening instruments for depression as well as different study designs allow no clear conclusion to the study question. The objective of this systematic review is therefore to clarify whether there is predictive value of cortisol levels (measured at which condition) in respect of the development of a depressive disorder.

Methods

Types of studies The design of primary studies will be prospective in order to examine the predictive value of cortisol status in the development of depression in the cohort of healthy adults. All studies will be identified in which the relationship between cortisol levels and depression has been examined and which meet the inclusion criteria and none of the exclusion criteria.

Inclusion criteria Observational studies, prospective cohort studies with a minimum length of 6 weeks, studies in which cortisol levels are measured in a local laboratory, studies of 1992 to 2015, studies in which depression is measured at two time points.

Exclusion criteria Case reports, studies with reports in another language than English, systematic reviews or meta-analysis, studies with subjects with a lifetime diagnosis of psychotic disorder, studies with subjects with antidepressant treatment, subjects with medications that interfere with cortisol metabolism, subjects with hormone replacement therapy including oral estrogens.

Study population Adults ≥ 18 years of age with measured cortisol levels and measures of depression, analyzed in prospective cohort studies between 1992 – 2015 or adults with a remitted depression. Patients in remission must be stable for at least 6 months without taking antidepressants.

Types of predictors and outcomes The intervention is the cortisol assessment as a prognostic test.

Table 1 Predictors and Outcomes

<table>
<thead>
<tr>
<th>Hormone of interest</th>
<th>Neuroendocrine axis</th>
<th>Context of interest</th>
<th>Disease of interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salivary cortisol</td>
<td>Hypothalamic-pituitary-adrenal axis</td>
<td>Mental health</td>
<td>Major Depressive disorder</td>
</tr>
<tr>
<td>Urinary cortisol</td>
<td></td>
<td></td>
<td>Recurrent Depression</td>
</tr>
<tr>
<td>Hair cortisol</td>
<td></td>
<td></td>
<td>Dysthymic Disorder</td>
</tr>
<tr>
<td>Serum cortisol</td>
<td></td>
<td></td>
<td>Depressive symptoms</td>
</tr>
<tr>
<td>CRH-Test</td>
<td></td>
<td></td>
<td>Adjustment disorder</td>
</tr>
<tr>
<td>DEX-CRH Test</td>
<td></td>
<td></td>
<td>Suicide in Depression</td>
</tr>
</tbody>
</table>

Outcome measures The outcome depression will be defined according to ICD-10 criteria (first appeared 1992, released by the World Health Organization), or DSM-IV criteria (first appeared 1994; released by the American Psychiatric Association in 2013) or DSM-V (first appeared 2015) that are internationally established diagnostic frames.
The outcome variable will be measured by standardized depression rating scales, for example the

1. Hamilton Depression Scale (HAMD) scale completed by researcher
2. Beck-Depressions-Inventory (BDI) scale completed by patients (self-report)
3. Center for Epidemiologic Studies Depression Scale (CES-D) - Self-report depression scale
4. Structural clinical interview (SCID-I)
5. Diagnostic Interview Schedule (structured interview designed to diagnose in a reliable and valid fashion the major psychiatric disorders according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV)
6. Symptom-Checklist-90®-Standard
7. Edinburgh Depression Scale (EDS).
8. Montgomery-Asberg Depression Scale (MADRS)

Search strategy

Search methods for identification of studies Systematic review of the peer-reviewed literature using different databases with the predefined search terms listed in Appendix 1, grey literature, hand search, expert consultation. Appropriate articles in English language will be extracted by title and abstract.

Data sources and Search Strategy The following databases will be searched to identify relevant studies: PubMed, Medline, Embase, Psychnfo. Two subjects will independently and systematically search for literature in different databases.

Data collection and analysis

Selection of studies Titles and abstracts of the studies will be screened for relevance and suitability by two independent reviewers. If necessary, disagreements will be resolved by a third reviewer. Included articles will then be retrieved in full text and assessed by using our inclusion and exclusion criteria.

Data extraction and management Data extraction will simultaneously be performed by two persons. Studies will be included according to the PRISMA four-phase flow diagram. Included studies will be examined and selected by study design, participants, intervention, data collection, and analysis, results and relevant contextual information. Results of the two persons will be compared and discussed afterwards to build up standard validation of data. The inter-rater reliability will be reported.

Measures of outcome The outcome of this review is the occurrence of depression or symptoms of depression. We will compare the outcome scores of adults with cortisol levels within range and adults with levels out of range at two different time points (T1, T2). If possible we will analyze the data as continuous. The continuous data will be reported as mean differences (MD) if the scales for the measurement are the same. If there are different ones we will report the standardized mean differences (SMD) (weighted mean difference, WMD – in meta-analysis). When cut-off values are reported, we will analyze the data as dichotomous and will calculate the risk ratio (RR) (or Odds Ratio). Ordinal outcome data are classified as follows: mild, moderate or severe depression will be reported. If there is an established, defensible cut-off point for the ordinal outcome data we will analyze them as dichotomous data (risk ratio or odds ratio). When ordinal scales are summarized using methods for continuous data, we will report the mean or standardized mean difference. The dichotomous and continuous data will be combined by statistical approaches (ORs as SMD and vice versa) (Deeks,
Higgins, & Altman, 2008). For RD (risk difference), OR and RR the 95% confidence intervals will be reported. Observer-rated scales and self-report scales will be considered separately.

**Dealing with missing data** When we encounter missing data, we will first contact the author and try to get the missing data. If clarification is not possible we will not include the study. The dropout rate will be reported in a table in which the characteristics of included studies are listed.

**Assessment of reporting bias and study quality** Quality of primary studies will be assessed by using the quality appraisal tool “Newcastle Ottawa Scale” for cohort studies. To check for publication bias, a funnel plot will be created.

**Assessment of heterogeneity** Clinical heterogeneity will be documented and assessed through sensitivity analyses and statistical heterogeneity will be assessed through Cochrane’s Q statistic ($X^2$ hypothesis test $p < 0.1$) and I² statistic (0-100%). For interpretation of I² we will use the following thresholds from Deeks, Higgins & Green 2011:

- 0% to 40%: might not be important
- 30% to 60%: may represent moderate heterogeneity
- 50% to 90%: may represent substantial heterogeneity
- 75% to 100%: considerable heterogeneity

**Data Synthesis** If possible, a random-effects meta-analysis will be performed presenting results in a forest plot. Other forms of evidence synthesis (Text/Table/harvest plot) will be performed for those data where a meta-analysis cannot be performed.
Contributors All authors were involved in the development of the protocol. They contributed to the selection of the predictors and outcomes and the overall study design. Daniela Stotz wrote the first draft of study protocol which was reviewed and revised by all co-authors. Caroline Sievers designed and managed the project.

Funding None.

Competing Interests There are no competing interests.

Data sharing statement We do not plan to distribute additional unpublished data to further parties.
Reference List


Appendix 1: Detailed search syntax. The database PubMed, PsycInfo, Medline, Embase will be studied.

**Pubmed**
(Cortisol OR DEX/CRH OR "HPA axis") AND (Depression OR Depressive Disorder OR Adjustment Disorder OR Suicide)

**PsycInfo**
(Cortisol OR DEX/CRH OR "HPA axis") AND (Major Depression OR depression OR depressive disorder OR Adjustment Disorder OR Suicide)

**Medline**
(Cortisol OR DEX/CRH OR "HPA axis") AND (Depression OR Depressive Disorder OR Adjustment Disorder OR Suicide)

**Embase**
(Cortisol OR DEX/CRH OR "HPA axis") AND (Depression OR Depressive Disorder OR Adjustment Disorder OR Suicide)