Aus dem Max Planck Institut für Psychiatrie München

Are aberrant cortisol levels prognostic factors for the development of depression in the adult population? A systematic review

Dissertation

zum Erwerb des Doktorgrades der Humanbiologie

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

vorgelegt von

Daniela Stotz geb. Müller

aus

Ravensburg

2019

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatterin:	PD Dr. Caroline Jung-Sievers
Mitberichterstatterin:	PD Dr. Karoline Lukaschek
Mitbetreuung durch die promovierte Mitarbeiterin:	Dr. med. Mareike Stieg
Dekan:	Prof. Dr. med. dent. Reinhard Hickel
Tag der mündlichen Prüfung:	25.10.2019

Abbreviations

ADS-K	Allgemeines Depressionsskala (short form)
ACTH	Adrenocorticotropic hormones
ADS	Allgemeine Depressionsskala
ADH	Antidiuretic hormone
APA	American Association
AUC	Area under the curve
BDI	Beck-Depressions Inventar
CBG	Corticosteroid-binding globulin
CES-D	Center for Epidemiologic Studies Depression
CMDQ	Common Mental Disorder Questionnaire
CRH	Corticotropin-releasing hormone
DASS	Depression, anxiety and stress scale
DEX-CRH	Dexamethasone-Corticotropin-Releasing Hormone
DHEA	Dehydroepiandrosterone
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, 4. Edition
DSM-V	Diagnostic and Statistical Manual of Mental Disorders, 5. Edition
DSQ	Depression Screening Questionnaire
EPDS	Edinburgh Postnatal Depression Scale
GDS-15	Geriatric Depression Scale
GH	Growth hormone
GR	Glucocorticoid receptor
HADS	Hospital Anxiety and Depression Scale
HAM-D	Hamilton rating scale for depression
HPA	Hypothalamic-pituitary-adrenal
ICD-10	International Classification of Disease, Injuries and Causes of death, 10th Edition
IDS	Inventory of Depressive Symptomatology
IL-1	Interleukin-1
IL-6	Interleukin-6
LH	Luteinizing hormone
MADRS	Montgomery-Asberg Depression Rating Scale
MDA	Malondialdehyde
MDI	Major depression Inventory

MHI-5	Mental Health Inventory
MR	Mineralocorticoid receptor
POMS	Profile of mood states
PTSD	Post-traumatic stress disorder
RNA	Ribonucleic acid
SCAN	Schedules for Clinical Assessment in Neuropsychiatry
SCID	Structured Clinical Interview
SCL-90	Symptom-Checklist-90
SDS	Self-Rating Depression Scale
SOD	Superoxide dismutase
TNF	Tumor necrosis factor
TSH	Thyroid stimulation hormone
WHO	World Health Organization

Danksagung

„Leider lässt sich eine wahrhafte Dankbarkeit mit Worten nicht ausdrücken“ – so Johann Wolfgang von Goethe. An dieser Stelle bleiben mir trotzdem nur Worte, um mich bei all denen zu bedanken, die mich auf dem Weg meiner Promotion begleitet haben. Eine Promotion ist selten der Weg einer einzelnen Person, sondern immer das mehrerer. In diesem Sinne gebührt mein ganz besonderer Dank meiner Doktormutter Caroline Sievers, die es mir ermöglich hat, meine Arbeit in einem interessanten und von mir bevorzugten Fachgebiet zu schreiben. Ich schätzte besonders ihre ruhige Art, ihre konstruktiven Kommentare sowie ihre hilfreiche Unterstützung und ihre motivierenden Worte. Ich möchte auch dem endokrinologischen Team des Max-Planck-Instituts dafür danken, dass es mir geholfen hat, das Thema zu finden und die Arbeit zu gestalten. Meinem Mann Markus, danke ich aus tiefstem Herzen, dass er mich auf diese Weise unterstützt und mich in schwierigen Momenten motiviert hat. Seine unermüdliche Unterstützung half mir, meine Doktorarbeit erfolgreich abzuschließen.

Table of contents

Abstract

According to the World Health Organization, (WHO, 2017), 322 million people suffered from depression in 2015, representing 4.4 percent of the world's population. Predictors of depression would help to identify patients at risk and shorten the time for an effective treatment. The aim of this systematic review is to determine whether cortisol levels predict the development of depression in the adult population. Our analysis considered various methods for measuring cortisol levels (salivary, urine, hair and serum), various timepoints and conditions (awakening cortisol, cortisol profiles, DEX-CRH test), the use of different screening instruments for depression as well as different study designs. Prospective observational studies and interventional cohort studies were investigated. The search was conducted via PubMed, Medline, PsychInfo, Embase, Science direct and Web of Science between March and April 2015 using appropriate indexed terms. All observational and prospective cohort studies between 1992 and 2015 were selected if they met all of the predefined inclusion and none of the exclusion criteria. Studies were included according to the PRISMA four-phase flow diagram. The quality of primary studies was assessed using the quality appraisal tool “Newcastle Ottawa Scale” for observational studies and “Jadad scale” for interventional studies. Our main outcome measure was the incidence of depression defined according to ICD-10, DSM-IV or DSM-V criteria. Standardized self-rating and foreign-rating scales were accepted for measuring the outcome variable. From a total of 4,619 articles in English language, we were able to identify 41 articles that met our criteria, of which 17 were observational studies and 24 interventional studies. Eleven observational and five interventional studies reported on the association between cortisol and follow-up depression. The present work shows that in women, elevated cortisol levels may act as a predictor of depression, whereas the current research situation in men is unclear. Salivary cortisol seems to be the most practicable measure. The best time for evaluation is “morning cortisol”. Measurements of the cortisol-awakening response are promising, although they are often not suitable because this type of measurement can be influenced by a variety of factors. Confirmatory studies based on our results are recommended before salivary morning cortisol measurements can be considered as clinical routine for predicting depression in women.

Abstract

Laut der Weltgesundheitsorganisation (WHO, 2017) litten im Jahr 2015 insgesamt 322 Millionen Menschen an Depressionen; 4,4 Prozent der Weltbevölkerung. Prädiktoren für das Auftreten einer Depression würden helfen Risikopatienten zu identifizieren und die Zeit für eine wirksame Behandlung zu verkürzen. Ziel dieses systematischen Reviews ist es festzustellen, ob der Cortisolspiegel die Entwicklung einer Deprression vorhersagt. In unserer Analyse wurden verschiedene Methoden zur Messung des Cortisolspiegels (Speichel, Urin, Haare und Serum), verschiedene Zeitpunkte und Bedingungen (Awakening Cortisol, Cortisolprofile, DEX-CRH-Test), verschiedene Screening-Instrumente für Depressionen sowie verschiedene Studiendesigns untersucht (prospektive Beobachtungsstudien und interventionelle Kohortenstudien). Die Suche wurde über PubMed, Medline, PsychInfo, Embase, Science direct und Web of Science zwischen März 2015 und April 2015 unter Verwendung geeigneter indizierter Begriffe durchgeführt. Alle beobachtenden und prospektiven Kohortenstudien zwischen 1992 und 2015 wurden ausgewählt, wenn sie alle vordefinierten Einschlusskriterien und keines der Ausschlusskriterien erfüllten. Die Studien wurden gemäß dem PRISMA Vierphasen Flussdiagramm aufgenommen. Die Qualität der Primärstudien wurde mit dem Qualitätsbewertungsinstrument "Newcastle Ottawa Scale" für Beobachtungsstudien und "Jadad Scale" für interventionelle Studien bewertet. Unsere Hauptergebnisvariable war die Inzidenz von Depressionen, definiert nach ICD-10-, DSM-IV oder DSM-V. Standardisierte Eigen- und Frembewertungsinstrumente wurden für die Messung der Ergebnisvariablen akzeptiert. Aus insgesamt 4619 Artikeln in englischer Sprache konnten wir 41 Artikel identifizieren, die unsere Kriterien erfüllten, darunter 17 Beobachtungsstudien und 24 Interventionsstudien. In elf Beobachtungs- und fünf Interventionsstudien wurde die Assoziation von Cortisol und Depression untersucht und berichtet. Die vorliegende Arbeit zeigt, dass bei Frauen erhöhte Cortisolspiegel als Prädiktor für eine Depression dienen könnten. Die aktuelle Forschungssituation bei Männern ist unklar. Speichelcortisol scheint für die Messung am praktikabelsten. Die beste Zeit für eine Probenentnahme ist das Morgencortisol. Messungen der Cortisol „Aufwachantwort“ scheinen ebenfalls vielversprechend, sie sind jedoch häufig nicht geeignet, da diese Art der Messung durch verschiedene Faktoren beeinflusst werden kann. Bestätigende Studien sind jedoch empfohlen, bevor Speichel-Cortisol-Morgenmessungen bei Frauen als klinische Routine für die Vorhersage von Depressionen genutzt werden.

1 Introduction

About 38% of the EU population suffers from a mental disorder. Among these 164 million people who are affected each year, major depression is claimed to be the most frequent disorder (Wittchen et al., 2011). In order improve the management of this disease or even prevent it, attempts have been made to identify possible predictors defining patients at risk. Researchers from a European consortium identified several risk factors for the development of depression, leading to the development of the so-called PredictD algorithm (King et al., 2008). Predictive factors that have been identified are age, gender, level of education, previous depressive episodes or a familial predisposition (King et al., 2008). However, given that these factors are mostly unmodifiable, they do not reflect the influence of environmental factors such as critical lifetime events. Nonetheless, it is well known that hormonal changes can be observed depending on the lifetime situation.

In detail, it was shown that there is an association of different hormones such as cortisol, testosterone, thyroid hormones, IGF-1, etc. with the risk of depression (Bhagwagar et al., 2005; Giltay et al., 2012; Sievers et al., 2014). One of the most studied promising candidates in this context is cortisol (Velders et al., 2011), as affected patients present with abnormal activity of the hypothalamic-pituitary-adrenal (HPA) axis (Holsboer et al., 2000; Pariante, 2003; Piwowarska et al., 2012). Known as a stress hormone, cortisol appears to be an important factor in the pathophysiology of depression, among other factors such as structural and functional changes in certain areas of the brain, serotonin deficiency, depression as an inflammatory process, the release of various hormones, etc. It is well known that an increased release of cortisol occurs in case of stress. Repeated stress or critical life events connected with a stress response and, thus, an increased release of cortisol, can lead to depression (Möller et al., 2009).

Cortisol is a very susceptible hormone that can be affected by various factors between and within individuals (Nicholson, 2008). In cross-sectional studies it has been clearly shown that early morning cortisol was increased in serum in patients with depression when adjusted for the time of awakening (Bhagwagar et al., 2005). Scott and Dinan (1998) demonstrated in a small sample that urinary-free cortisol levels of depressive patients are significantly higher compared with healthy controls. In a study with 42 students, hair cortisol concentrations were associated with the presence of depressive symptoms, suggesting that hair cortisol measurement can also serve as a potential biomarker for this kind of mental disorder (Gerber et al., 2013).

Disturbances of hormonal axes cannot only be identified by reduced baseline levels. In order to identify impaired hormonal axes, suppression tests are necessary to investigate whether elevated baseline levels will appropriately be reduced upon suppression. In fact, patients with depression presented with a blunted suppression of plasma cortisol levels following the intake of dexamethasone indicating disturbances in HPA axis negative feedback regulation (Fang et al., 1981; Halbreich et al., 1985). Based on results from our institute, it has been suggested that altered HPA axis regulation measured by the DEX/CRH test at two different time points might be the best HPA axis related biomarker for predicting the clinical outcome at follow-up. All of these findings indicate that individual differences in cortisol levels as well as an altered HPA axis reactivity are possible risk factors respectively markers for the development of depression.

However, there are some inconsistencies and contradictory results regarding the time, method and applicability of cortisol measurements. For instance, Grynderup et al. (2013) examined 4,231 healthy people and found a correlation between the delta in morning and evening salivary cortisol concentration and the risk of developing depression while this was not the case for morning or evening salivary cortisol levels alone. This result contradicts findings by Harris et al. (2000) stating that individual differences in morning salivary cortisol levels (but not evening salivary cortisol levels) in females can be considered a risk factor for major depressive disorder (Harris et al., 2000). Herbert et al. (2012) came to the conclusion that an association between the adjusted morning cortisol levels predicts the risk for depression onset during a 12-month follow-up period best. According to a recent study, high cortisol values determined in the late morning in women at the age 45 years are predictive for subsequent depressive symptoms at the age of 50 years. In men, lower instead of higher cortisol values were prospectively associated with the occurrence of depressive symptoms (Geoffroy et al., 2013).

In conclusion, various biomaterials (salivary, urine, hair, serum) at various time points 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 thus far. The objective of this systematic review is therefore to clarify whether there is an evidence-based, pragmatic way of using cortisol measurements as additional indicators/predictors for physicians to evaluate whether their patients might develop a depression in the near future or not.

.

2 Background

2.1 Depression

The mental state of depressed persons is well described by the term “press down”, while the Latin word for this is deprimere (O’Toole, 2016). In everyday life, the word depression is often used if we do not feel well, do not feel like doing something or are sad, although according to Hegerl et al. (2005) it can lead to misunderstandings. because we all experience depressive symptoms in our lives, such as sadness or drivelessness. Often such experiences emerge during stressful phases of life, illness or stressful events, such as the death of a relative. However, these symptoms alone do not warrant the diagnosis of depression in the clinical sense. Being depressed does not mean being completely sad or perhaps unmotivated; rather, it is a disorder that can lead to impairments on the *"emotional, cognitive, physiological, motor, social-interactive and behavioral"* levels of the human body (Wittchen & Hoyer, 2006).

2.1.1 Diagnose of Depression

The notion that the diagnosis of the depression is not mistakenly awarded a clear definition is indispensable in view of a clear clinical diagnosis, which should be based on operationalized criteria. In order to guarantee this, different diagnostic manuals have been introduced in the past in which internationally-recognized diagnostic definitions are operationalized. The currently-used classification schemes are the ICD-10/ICD-11 (International Classification of Diseases, 10th Edition), which is predominantly used in German-speaking countries, and the DSM-IV/DSM V used in English-speaking countries (Diagnostic and Statistical Manual of Mental Disorders, 4th Edition). The former was published by the WHO (1993), the latter by the American Psychiatric Association (APA, 1994). In the meantime, the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders, the so-called DSM-5 (APA, 2013) has been published. The American Psychiatric Association released this 5th edition in May 2013 (Ehret & Berking, 2013). In the present work, all three diagnostic systems will be considered.

2.1.2 The clinical picture of depression in ICD-10

In ICD-10, the disorder of depression is listed in the diagnostic category "Affective Disorders". Affective disorders are characterized by a depressed mood of unusual proportions, with or without fear. Its episodic occurrence is often associated with a distressing situation or event. The disease of depression is divided into three different degrees of severity. Depending on the severity, the disease is characterized by three major symptoms: 1.) depressed mood most of the day and almost every day 2.) loss of interest and enjoyment of activities and listlessness 3.) increased fatigue. In a mild or moderate depressive episode at least two of these symptoms must be present, in severe depression at least three. In addition, there are other symptoms, such as 4.) loss of trust or self-esteem 5.) unfounded self-blame or guilt 6.) suicidal thoughts or actions 7.) concentrations problems or indecision 8.) psychomotor agitation or inhibition 9.) sleep disorders 10.) appetite loss or increase. In mild depression, there must be at least two of these symptoms, three to four in a moderate, and at least four in a severe depression. Despite the patient's experience of mild depression, he can fulfill his professional obligations and continue to participate in social life. In the case of moderate depression, performing everyday activities is much more difficult for the patient, and in the case of severe depression, the patient is no longer able to carry out his day-to-day activities. In a mild or moderate depressive episode, symptoms must at least two weeks before the diagnosis can be given. In severe depression, this diagnosis can be given earlier, but this depends on the severity of the disease (Dilling, Mombour, Schmidt, 2014).

2.1.3 The clinical picture of depression in DSM-IV

The criteria of DSM-IV are different from those of ICD-10 in terms of clinical significance, duration of symptoms and severity of disease (Gruenberg, Goldstein, Pincus, 2005). In DSM-IV there is no subdivision by severity. Due to the lack of severity classification in the DSM-IV the diagnosis of depression can then be given, if there are five or more symptoms (Kölch, Fegert & Freyberger, 2011; Gruenberg, Goldstein & Pincus (2005) and the clinical significance is then given when the symptoms lead to *"impairments in social, occupational, or other important areas of functioning"* or if the symptoms *"cause clinically significant distress*" (Gruenberg, Goldstein & Pincus 2005; Kölch, Fegert & Freyberger, 2011). Furthermore, DSM-IV *"feelings of worthlessness or inappropriate guilt"* are listed together, and according to Gruenberg, Goldstein and Pincus (2005) this can be considered more serious in quality compared with the criterion "loss of self-confidence or self-esteem" as listed in the ICD-10. In both classification systems, the diagnosis of depression requires that the

symptoms persist for at least two weeks. However, the DSM-IV notes that these symptoms must persist "most of day, almost every day". In the ICD-10, only the symptom of depressed mood is required to be present almost every day and for most of the day (Kölch, Fegert & Freyberger, 2011, Gruenberg, Goldstein, Pincus, 2005). For clarity, the criteria of DSM-IV-TR (Faumann, 2001) and ICD-10 (Dilling, Mombour, Schmidt, 2014) are contrasted in Table 1.

Table 1.

Diagnostic criteria of depression ICD-10. Adapted to Gruenberg, Goldstein, Pincus (2005); Dilling, Mombour, Schmidt (2014); Faumann (2001).

	ICD-10	DSM-IV-TR
Symptoms	**Three major symptoms** 1. depressed mood most of the day and almost every day 2. loss of interest and enjoyment of activities and listlessness 3. increased fatigue In a mild or moderate depressive episode at least two of these symptoms must be present, in severe depression at least three **Other symptoms** 4. loss of trust or self-esteem 5. unfounded self-blame or guilt 6. suicidal thoughts or actions 7. concentrations problems or indecision 8. psychomotor agitation or inhibition 9. sleep disorders 10. appetite loss or increase In mild depression, there must be at least two of these symptoms, three to four in a moderate, and at least four in a severe depression.	**Symptoms** 1. depressed mood most of the day and almost every day 2. loss of interest or pleasure in all, or almost all, activities most of the day, nearly every day **3. Feelings of worthlessness or guilt feelings that are inappropriate and excessive** 4. suicidal thoughts or actions 5. concentrations problems or indecision nearly every day 6. increased fatigue nearly every day 7. psychomotor agitation or inhibition nearly every day 8. sleep disorders 9. appetite loss or increase every day / weight loss or weight gain **There is no subdivision by severity** **The diagnosis of depression can then be given, if there are five or more symptoms. One of the symptoms is 1) depressed mood or 2) loss of interest or pleasure .**
Clinical relevance	**Mild depression** Despite the patient's experience of mild depression, he can fulfill his professional obligations and continue to participate in social life. **Moderate Depression** In the case of moderate depression, performing everyday activities is much more difficult for the patient. **Severe Depression** The patient is no longer able to carry out his day-to-day activities.	**There is no subdivision by severity** The symptoms affect the person in social, occupational or other functional areas. The patient experiences this as a significant burden.
Symptom duration	**Mild or moderate depressive episode** Symptoms must at least two weeks before the diagnosis can be given. **Severe depression** Diagnosis can be given earlier than two weeks, but this depends on the severity of the disease	**There is no subdivision by severity** Symptoms must at least two weeks before the diagnosis can be given

2.1.4 The clinical picture of depression in DSM-5

Depressive disorders are listed in the DSM-5 in the “new depressive disorders” section, which is separated from the “bipolar disorders” section. In the DSM-IV, this division did not exist, all illnesses listed in these two chapters were listed under the term mood disorders. Chronic disease and dysthymic disorders are listed in DSM-5 under the heading of persistent depressive disorders (Ehret, Berking, 2013; Uher, Payne, Pavlova, Perlis, 2014). The depression criteria have hardly changed from DSM-IV to DSM-5. The diagnosis of major depression can be given in the event of one depressive episode or multiple depressive episodes over the life course. For the diagnosis, five to nine symptoms must be present over a period of two weeks. Depressed mood or anhedonia (loss of interest or pleasure) is a symptom that must be present. Some of the DSM-5 symptoms of depression need to be present almost every day, while others are not (see Table 2).

According to Uher et al. (2014), there are three changes in the DSM-V compared with the DSM-IV. The first change is the removal of the statement “that mood-incongruent delusions or hallucinations should not count toward the diagnosis of MDE/MDD”. The second change is that in the DSM-V the word *hopeless* was added. Thus far, only the word *sad* was listed here. The word "hopeless" deals with how a person perceives the future perspectives. The third change refers to the grief reaction, which was listed in DSM-IV but removed in DSM-V. The diagnosis of major depression should be well evaluated in bereavement because grief reactions may be confusingly similar to the symptoms of depressive illness (Ehret, Berking, 2013; Uher, Payne, Pavlova, Perlis, 2014). Ermann (2007) states that despite being similar, the two constructs of depressive illnesses and grief reactions are based on completely different experiences. A person suffering from depression often revolves around him-/herself mentally, feeling an inner emptiness and is confronted with self-blame. However, a person grieving for a loved one feels incredibly sad but this does not affect their self-esteem. This difference must be taken into account when making the diagnosis. For clarity, the criteria of DSM-IV-TR (Faumann, 2001) and DSM-5 (American Psychiatric Association, 1994) in Table 2 are contrasted.

Table 2.

Diagnostic criteria of depression DSM-5 and changes from DSM-IV-TR. Adapted to Uher, Payne, Pavlova, Perlis (2014); Faumann (2001); American Psychiatric Association (1994).

DSM-IV-TR	DSM-5
Symptoms 1. depressed mood most of the day and almost every day „indicated by either subjective report (e.g. feels sad or empty) or observation made by others (e.g. appears tearful))" 2. loss of interest or pleasure in all, or almost all, activities most of the day, nearly every day 3. Feelings of worthlessness or guilt feelings that are inappropriate and excessive nearly every day 4. suicidal thoughts or actions 5. concentrations problems or indecision nearly every day 6. increased fatigue nearly every day 7. psychomotor agitation or inhibition nearly every day 8. sleep disorders nearly every day 9. appetite loss or increase every day / weight loss or weight gain	**Symptoms** 1. depressed mood most of the day and almost every day „indicated by either subjective report (e.g. feels sad, empty, hopeless (added in DSM-5) or observation made by others (e.g. appears tearful))" 2. loss of interest or pleasure in all, or almost all, activities most of the day, nearly every day 3. Feelings of worthlessness or guilt feelings that are inappropriate and excessive nearly every day 4. suicidal thoughts or actions 5. concentrations problems or indecision nearly every day 6. increased fatigue nearly every day 7. psychomotor agitation or inhibition nearly every day 8. sleep disorders nearly every day 9. appetite loss or increase every day / weight loss or weight gain
The diagnosis of depression can then be given, if there are five or more symptoms. One of the symptoms is 1) depressed mood or 2) loss of interest or pleasure	The diagnosis of depression can then be given, if there are five or more symptoms. One of the symptoms is 1) depressed mood or 2) loss of interest or pleasure
The symptoms affect the person in social, occupational or other functional areas. The patient experiences this as a significant burden.	The symptoms affect the person in social, occupational or other functional areas. The patient experiences this as a significant burden.
Note: "Do not include symptoms that are clearly due to a general medical condition, or mood delusions or hallucinations (removed in DSM-5 → see column on the right)	Note: "Do not include symptoms that are clearly attributable to another medical condition"
"The symptoms are not better accounted for by bereavement (removed in DSM-5 → see column on the right) i.e., after the loss of a loved one, the symptoms persist for longer than 2 months or are characterized by marked functional impairment, morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms, or psychomotor retardation.	"Responses to a significant loss (e.g., bereavement, financial ruin, losses from a natural disaster, a serious medical illness or disability) may include the feelings of intense sadness, rumination about the loss, insomnia, poor appetite, and weight loss noted in Criterion A, which may resemble a depressive episode. Although such symptoms may be understandable or considered appropriate to the loss, the presence of major depressive episode in addition to the normal response to a significant loss should also be carefully considered. This decision inevitably requires the exercise of clinical judgment based on the individual's history and the cultural norms for the expression of distress in the context of loss."

2.1.5 Epidemiology of depression

Depression is one of the most common mental illnesses and the risk of developing a depressive disorder is steadily increasing (Will et al. 2008; Wittchen & Hoyer, 2006). This statement is in line with the information provided by the WHO (2017), which wrote in a report from 2017 that an increase of 18.4 percent was recorded between 2005 and 2015. A total of 322 million people suffer from depression, which corresponds with 4.4 of the world population (WHO, 2017). According to the WHO, the prevalence rates vary depending on the world region (Figure 1).

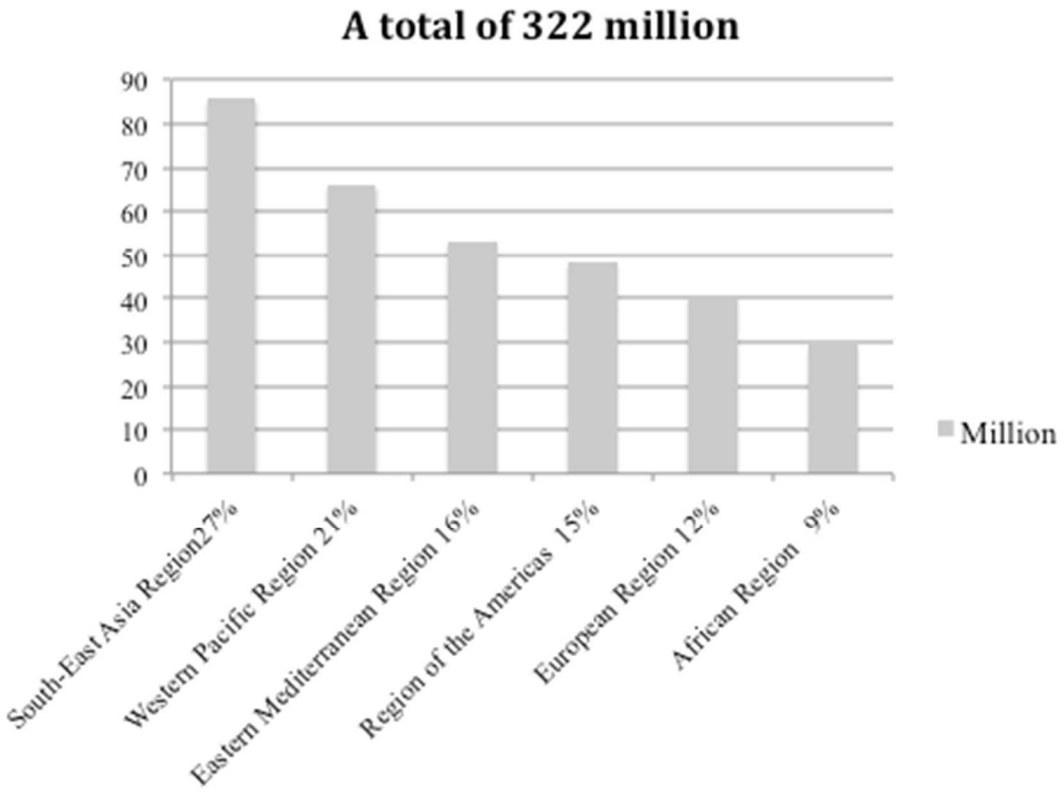

Figure 1. Cases of depressive disorders (millions) by WHO regions. Adapted to WHO (2017).

The prevalence rate in women between the ages of 55 and 74 years is 7.5%, for men it is 5.5% (WHO, 2017). The fact that women are more likely to develop depression than men is evident across different countries (Andrade et al., 2003). Albert (2015) attributes these differences to different types of stressors. Women are more prone to disagreements in interpersonal relationships, which are mainly processed internally. Men, on the other hand, are more sensitive to performance stress at work (Albert, 2015; Bartels et al., 2013; Kendler &

Gardner, 2014). In addition, women are more likely to undergo hormonal changes during their lifetime this can lead to *"premenstrual dysphoria, postpartum depression or postmenopausal depression"* (Albert, 2015). Another reason that may be responsible for this is developmental differences in the cerebral circulation (Albert, 2015; Gillies & McArthur, 2010). Women who have no stable relationship with a partner or friend, has three or even more children younger than 14 years or whose mother died before the age of eleven are more vulnerable to the onset of depression. Likewise, single unemployed women seem also be at increased risk (Brown & Harris, 1978; Will et al., 2008).

Approximately 12-26% of the adult population develops a moderate depression at some point in their lives (Will et al. 2008). The lifetime prevalence of depression, varies greatly, between 4% and 20%. These differences can be attributed to various factors such as *"study population, used diagnostic criteria, survey response rates or methodological procedures"* that can influence the study outcome. For example, depending on the country, it could be shown that the rate in Japan is very low at 3 percent, whereas in the US the rate is 16.9% (Andrade et al. 2003). The lifetime prevalence in Germany is 17.1 percent (Jacobi et al., 2004). Between 2008 and 2011, 7,988 people in Germany were examined for a recent depression. The age of the examined persons lay between 18 and 79 years. In adults, depressive symptoms were 8.1%. Here, again the prevalence was higher in women (10.2%) compared with men (6.2%). It could also be shown that a higher socioeconomic status is less strongly associated with depression. According to this study, the 12-month prevalence is 6.0% (Busch, Maske, Ryl et al., 2013). Although depression can occur at any age, the risk of developing depression seems to be highest between the ages of 30 and 40 years (Will et al., 2008). Andrade et al. (2003), on the other hand, indicates the average for onset of illness is between 20-25 years. According to Möller, Laux and Deister (2009), 10% of the people who present themselves in a general practice suffer from depression. In a cross-sectional epidemiological study from Beesdo-Baum, Knappe and Einsle (2018), it was examined how often depression is detected when patients present themselves in a general practice. For this purpose, 3,563 patients were examined on a fixed date in six different regions of Germany. For the investigation, the Depression Screening Questionnaire (DSQ) was used to raise mental and physical complaints. A medical diagnosis was made based on the ICD-10. According to the researchers' analysis, the ICD-10 depression reporting date prevalence was 14.3% according to DSQ and 10.7% after a doctor's diagnosis (Beesdo-Baum, Knappe, Einsle, 2018).

2.1.5 Types of depression measurement

Since the effects of depression can be very serious in terms of private and professional life, a quick and accurate diagnosis is particularly important. According to Wittchen and Hoyer (2006), the *"symptoms of depression, the severity of the disease and the temporal occurrence"* should be recorded reliably and validly. Accordingly, the measurement accuracy should be high and an existing depression should be detected and not something else. In order to guarantee this, various questionnaires as well as structured interviews have been developed in the past. It is impossible to list all existing depression questionnaires here, as there are over a hundred of them (Kronmüller & Mundt 2000 as cited in Bouman, 1993). According to Kronmüller and Mundt (2000), the most common depression questionnaires are the Beck-Depressions Inventar (BDI; Beck, Hautzinger, Bailer, Worall & Keller, 1995), Self-Rating-Depression Scale (SDS; Zung, 1965) or the Allgemeine Depressionsskala (ADS; Hautzinger & Bailer, 1993).

Due to the variety of questionnaires that are currently used in clinical practise, the question arises whether the results of the questionnaires are in fact all comparable. Questionnaires that are used to detect depression correlate on average by 0.69 (Kronmüller & Mundt, 2000 as cited in Bouman, 1993). Based on previous studies, Kronmüller and Mundt (2000) conclude that so far there is no questionnaire that can be fully recommended. Wittchen and Hoyer (2006) note that depression questionnaires often only capture the current severity of depression. However, according to the author, this is not always sufficient for the diagnosis of depression, as these do not allow differential-diagnostic conclusions. Therefore, the author advises using better structured or standardized clinical interviews. As an example of a structured clinical interview, he cites the SCID, which is based on the classification system of the DSM-IV. ("Structured Clinical Interview" for DSM-IV). It was shown that depression questionnaires and interviews correlate only moderately Kronmüller and Mundt, 2000. Both depression questionnaires and interviews seem to have disadvantages to minimize them and gain reliable diagnostic statements, the authors Katz's et al. (1995) suggestion is to use a test battery. In doing so, different survey instruments are to be combined with each other. For example, the SCID-I (Wittchen, Zaudig, Fydrich, 1997) could be used together with the HAM-D (Hamilton rating scale for depression; Hamilton, 1980). If the survey were based more on the classification system ICD-10, then the SCAN (Schedules for Clinical Assessment in Neuropsychiatry; Schützwohl, Kallert, Jurjanz, 2007) interview could be combined with

the Inventory of Depressive Symptomatology (IDS, Rush 1986, 1996), for example. However, the authors also note that the combination of different diagnostic tools often results in new methodological problems. This problem cannot be solved in a satisfactory manner at present.

2.1.6 Etiology of depression

According to Bramesfeld and Stoppe (2006), depression is a disease that can be caused by many factors, whereby it is also called a multicausal event. Möller et al. (2009) divide the factors underlying the depression into three groups, whereby genetic factors, psychological factors and neurobiological factors may play a role in the development of depression.

In their work entitled *"Genetic Epidemiology of Major Depression: Review and Meta-Analysis",* Sullivan et al. (2000) conclude that the heritability of depression is approximately between 31% and 42%. In severe depression or recurrent depression, the probability of disease could be quite higher. Children whose parents are already suffering from depression have a much higher genetic risk. In addition to this, mental vulnerability factors could be detected much more often (Schulte-Körnke et al., 2008). Schulte-Körnke et al. (2008) state that children of depressed parents are exposed to a rather *"passive, neglecting and repellent style of interaction"* and that these children perceive their parents as a *"model for negative self-attribution, dysfunctional schemes, and low self-efficacy expectations"*. In this context, the research team refers to Frye and Garber (2005) and Hammen et al. (2004). According to Wittchen and Hoyer (2006), there is both a passive and active gene-environment interaction. The passive gene-environment interaction is characterized by a greater vulnerability to uncontrollable life events. In the case of active gene-environment interaction, the probability that a depressive crisis event itself is caused is greater, such as a separation from a partner or a termination of an employment contract, etc.

Among the psychological factors, Möller et al. (2009) first mention critical life events. According to the authors, depressive persons more frequently report that such an event has preceded depression. Critical life stories are always associated with a stress response: if such a period of stress lasts longer, then this can lead to an overload, which is reacted to through marked exhaustion or withdrawal from the social environment, in turn leading to depression.

In terms of neurobiological factors, Möller et al. (2009) state that the imbalance between the neurotransmitters noradrenaline and serotonin is associated with depression. In persons suffering from depression, a deficit of these two neurotransmitters could be detected. Brain-

based examinations have shown that the disease of depression is accompanied by morphological changes as well as an activation disbalance of different brain regions. The prefrontal cortex as well as the hippocampus seem to show a diminished substance and negative stimuli lead to an overactivation of the amygdala. Among the neurobiological factors, seasonal differences can also frequently occur in autumn/winter and lead to depression. The circadian rhythm in depressed people is often disturbed and the sleep phases are lower and less pronounced in depressives.

Turning to neuroendocrinological research, depression is repeatedly associated and discussed with a thyroid disorder or dysregulation of the HPA axis (Möller et al., 2009).

2.2 Hypothalamic-pituitary-adrenal axis

The HPA axis is also referred to as the stress axis. This axis enables healthy people to deal with stressful situations and adapt to them (Verdu, Hayes, O'Mahony, 2016). The following section describes the anatomy and physiology of the HPA axis and discusses the associated circadian rhythm. Subsequently, the manifold possibilities of cortisol measurements are described.

2.2.1 Anatomy and physiology of the HPA axis

The HPA axis is a system of three existing hormones, the corticotropin-releasing hormone (CRH), the adrenocorticotropic hormones (ACTH) and the cortisol (Abel & Hautzinger, 2013). This finely tuned hormone cascade can be accomplished in several ways, by either internal stimuli such as pro-inflammatory cytokines or external stimuli such as stress (Goebel & Schedloski, 2008).

The hypothalamus responsible for the release of CRH is located in the diencephalon and can be understood as a mediator between the nervous system and the endocrine system. This brain structure is particularly significant because it participates in a variety of physical processes and ensures that the body functions does not fall out of balance. As a command center, it regulates our basic needs such as hunger and thirst, the temperature of our body or our sleep-wake cycle. In addition, this structure acts as a controlling organ for our nervous system or our behavior in social contexts. The latter happens through the regulation of emotions. Furthermore, the production of various hormones is also controlled by the hypothalamus (Bley, Centgraf, Cieslik, Haack, Hohloch, 2015).

The hypothalamus is connected to the pituitary gland, which is located in the base of the skull. This brain structure is divided into two distinct areas, namely the posterior and anterior pituitary lobes. For the function of the HPA axis, especially the anterior pituitary is important. In case of stress or physical exertion, the cortex and especially the limbic system stimulate the hypothalamus to release the CRH and the antidiuretic hormone (ADH) via serotonergic and cholinergic fibers (Rensing et al., 2006). These two hormones in turn stimulate the corticotropic cells of the pituitary gland to produce and release the adrenocorticotropic hormone (ACTH), which then enters the blood and is thus transported to the adrenal cortices (Bley, Centgraf, Cieslik, Haack, Hohloch, 2015; Hinson, Raven & Chew, 2018; Nicholson, 2007).

The adrenal cortices are located on the kidneys and can be divided into three different areas: the zona glomerulosa, zona reticularis and the zona fasciculata. The latter area is

responsible for the formation of glucocorticoids. One of the most important glucocorticoids is the hormone cortisol (Zeros-Kopp, 2007; Nicolson, 2008). In stress situations, this hormone is formed and ensured that the body has sufficient energy available. (Bley, Centgraf, Cieslik, Haack, Hohloch, 2015). The amount of cortisol in the blood is regulated by a complex interaction between the hypothalamus, pituitary and adrenal glands. Once released, the hormone cortisol enters into all parts of the body and back to the hypothalamus, where the amount of cortisol is controlled. The overshoot of a stress response is prevented by this negative feedback mechanism of the HPA axis (Hinson, Raven & Chew, 2018). An overview of this process is shown in Figure 3.

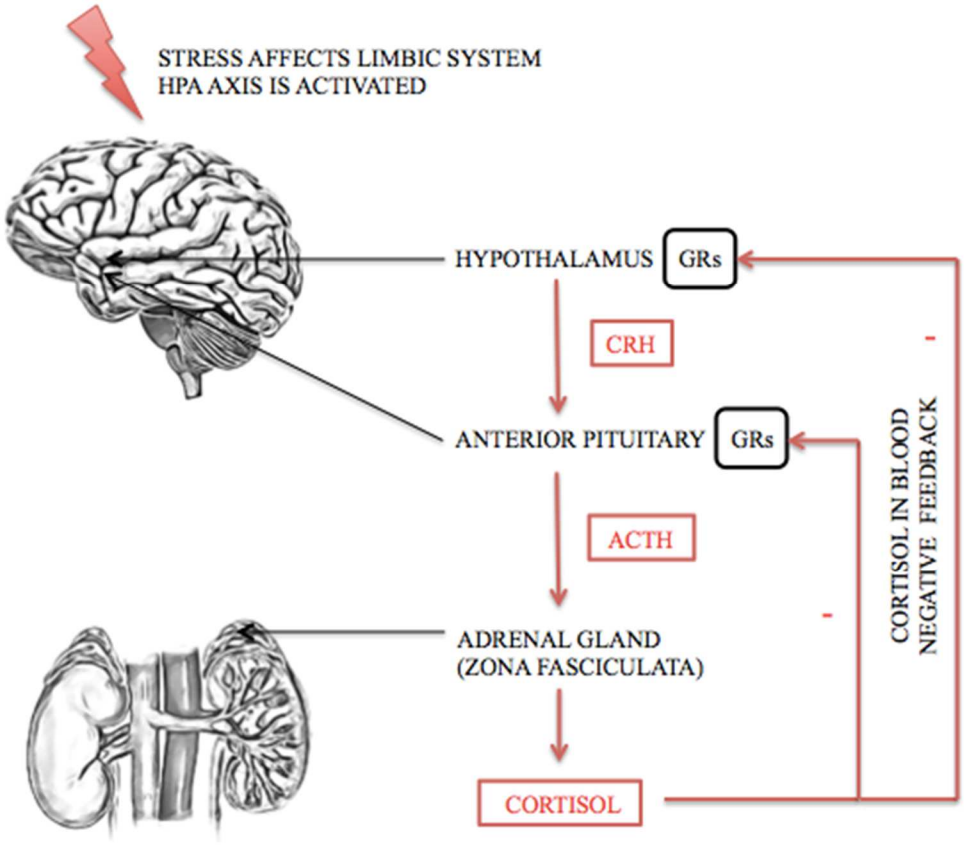

Figure 2. Interactions of the HPA axis. Adapted to Kulkarni, Gavrilidis & Worsley (2016).

2.2.2 Cortisol circadian rhythm

The release of the hormone cortisol is not only dependent on stress but is also subject to a circadian (Day-Night) rhythm. Between 6 o'clock and 9 o' clock in the morning the serum cortisol values reaches its maximum (Hinson, Raven & Chew, 2018). According to Nicolson (2008), the highs depend on the personal waking times. The cortisol production increases rapidly after awakening and reaches its peak about 30 to 45 minutes later. From this on, the concentration gradually decreases over the day (Hinson, Raven & Chew, 2018) until it reaches its lows between 4 and 8 pm (Froesch, 1981).

2.2.3 Types of cortisol measurement

Cortisol can be measured in several ways. Nicolson (2008) lists various types of measurement in the Handbook of Physiological Research Methods in Health Psychology discusses the pros and cons of each technique. According to the author, the measurement of cortisol in urine is one of the oldest methods. This type of measurement is particularly suitable for determining the cortisol level at night. For some diseases are mainly changes in the night cortisol but not in the daily cortisol can be observed (Nicolson, 2008). For example, this applies to patients with panic disorder (Abelson and Curtis, 1996) or with a post-traumatic stress disorder (Yehuda, 2002). The free urinary cortisol levels are not affected by the circadian rhythm, so the "relative free cortisol concentrations in the peripheral circulation in one day" can be measured well (Gatti et al. 2009). For research purposes, Nicolson (2008) argues that urine measurements are rather inappropriate because an acute stress response cannot be detected and the participation in the study as well as adherence to the measurement procedure by the participants are often impaired. Furthermore, cortisol is a molecular structure that is very similar to other steroid metabolites, which can lead to problems in the evaluation. This problem can be significantly improved by the use of solvents, which extracts certain steroids but they are not completely removed (Gatti et al. 2009; Murphy, 1999, 2002). Sauve et al. (2007) conducted a study comparing urine cortisol samples and hair cortisol samples, finding only a weak correlation. Sauve et al. (2007) states that "*urine represents cortisol secretion during one day, whereas hair cortisol levels represent levels during 1-2 months"*. Haircortisol samples have several factors that need to be considered. Hair cortisol levels are lower in dyed hair compared with untreated hair. Haircortisol samples can only be taken if there is sufficient hair in the area of the posterior vertex. The detection of short cortisol responses or day-to-day variation is not possible.

Another possibility to determine the cortisol concentration of subjects would be the use of the dexamethasone suppression test. Dexamethasone is a glucocorticoid. If it is administered from the outside, it suppresses the cortisol release in the body. This is achieved via the already-described negative feedback mechanism (Piper, 2007). The cortisol measurement before and after the test can be conducted via either a saliva test or a blood sample. As a diagnostic test care must be taken in the event of depression (Nicolson, 2008; Carroll et al., 1981). On the first day of the examination, 1 mg dexamethasone should be administered at 11 p.m. and on the second day at 4 p.m. In addition, Nicolson (2008) noted that the concentration of cortisol samples is dependent on the dexamethsone dose: when taking more dexamethasone, then cortisol suppression will take longer. The removal of several saliva samples leads to more reliable results. The necessity for precise adherence to this procedure may be one reason for the test having been rarely been used in the studies, so Nicolson (2008). In stress situations, this test is unsuitable due to its sensitivity because it increases the risk of false test results. In persons with malnutrition, obesity, sleep problems, drug and alcohol withdrawal or depression, suppression of plasma cortisol may be absent after the administration of dexamethasone (Berger 1984, Beuschlein & Reincke, 2006). Arane et al. (1995) conducted a study in which she investigated whether the administration of dexamethasone is successful in the treatment of depression. The results of the study have shown that the administration of dexamethasone over 4 days of treatment is more effective compared with a placebo for 14 days.

Measurements of cortisol levels in saliva are widely used today, due to the advantages this type of measurement entails. Cortisol is not bound in saliva, unlike in the blood, meaning that it is not bound to any carrier proteins such as corticosteroid-binding globulin (CBG). It is a non-invasive procedure that can be well used by the study participants at home can have a positive influence on participation in the study because it is stress-free. The fear of injection could be a criterion for refusing to study. In addition, the simple removal of the salivary cortisol makes it easier to obtain more samples from the participants. The cortisol samples can be obtained by either collecting the saliva in a plastic tube or using a salivette. The latter is more accurate in terms of the determination of total and free cortisol. Saliva cortisol measurements can be performed at any time of the day. A popular sampling time is shortly after waking up (Gatti et al., 2009; Nicolson 2008; Vining et al., 1983). Although a relationship between salivary cortisol samples and serum cortisol samples could be demonstrated (Sauve et al., 2007), there are some differences between these two types of measurements. Compared with the decrease in serum, the absolute level of free cortisol in

saliva is about 10% to 35% lower. It has been shown that cortisol response is weaker following awakening than in saliva. Serum cortisol returns more quickly to baseline after stress (Nicolson, 2008).

2.3 Neuroendocrinological changes in depression

2.3.1 HPA axis in depression

Pariante and Lightmann (2008) write in their review entitled "The HPA axis in major depression: classical theories and new developments":

"HPA axis hyperactivity is not a simple consequence or an epiphenomenon of depression, but on the contrary that is a risk factor predisposing to the development of depression, brought about by early life experiences programming molecular changes as well as by genetic liability".

According to Pariante (2006), there are two different corticosteroid receptors. The mineralocorticoid receptor (MR) and the glucocorticoid receptor (GR). These two receptors differ in their affinity; The MR is affine for endogenous corticosteroids, whereas the GR is affine for dexamethasone. According to Roth et al. (2018) mineral corticosteroids are particularly important for the regulation of the salt-water balance in the body. The mineralocorticoids are regulated via the renin-angiotensin system as well as via the calcium concentration in the plasma. Glucocorticoids, on the other hand, are important for the intermediary metabolism, and their regulation takes place via the hypothalamic pituitary system. Based on the existing research situation, it can be assumed that the GR-mediated negative feedback is impaired in depressive people, but the MR mediated negative feedback intact (Juruena et al., 2005; Pariante, 2006). According to Pariante and Lightman (2008), the GR plays an important role in depression. Spilled cortisol in the body can bind to those receptors that have their seat either outside the brain (e.g. in the pituitary) or in the brain itself (e.g. hippocampus or hypothalamus). This compound regulates HPA axis activity and inhibits cortisol release. It is assumed that this "GR-mediated feedback inhibition" in depression does not work as usual. The GR activity can be influenced by environmental influences e.g. through a previously experienced trauma or indirectly by inflammatory processes. According to Wang et al. (2004), via the P38 mitogen-activated protein kinase pro-inflammatory cytokines can reduce the GR function. Pariante and Lightman (2008) assume that the overactivation of the HPA axis in depressive individuals reflects previous critical stressful life events.

2.3.2 Cortisol secretion in depressed adults

In recent years, researchers have repeatedly been concerned with the question whether there is a dysregulation, of the hypothalamic-pituitary-adrenal axis in depressed people. Already in 1984, Claustrat and colleagues examined a total of eleven patients suffering from depression regarding their cortisol levels and compared them with the values of healthy volunteers. The evaluation of the results showed increased mean plasma cortisol levels in the group of persons with depression. Carroll et al. (1976) found that daily urinary free cortisol values are particularly high in patients with a depressive disorder. In people with neurotic depression, cortisol levels were not as high as in people with unipolar depression. The values in patients with unipolar depression were comparable to the values of people with Cushing's syndrome. According to a systematic review by Staufenbiel et al. (2013) an increased cortisol concentration could be detected in the hair of depressive people. Bhagwagar et al. (2005) also showed higher cortisol levels in depressive subjects. They used awakening cortisol and compared the values of those with depression to those of healthy ones. Cortisol levels were 25% higher in subjects with severe depression than control subjects. The studies listed here show that the cortisol levels in depressed people are changed compared with healthy people and that the same results are repeatedly obtained regardless of the type of cortisol measurement,.

In a meta-anaylsis, Burke et al. (2005) examined the stress response of the HPA axis in healthy and depressed people. In addition, the researchers were interested in how the response of the HPA axis behaves after the loss of the stressor. The evaluation of the results showed that depressives and healthy people differ in their stress response. The stress response of depressed is not as pronounced and the recovery is impaired. Depressed individuals thus the authors *"exhibit a relatively flat and unresponsive pattern of cortisol secretion"*. The researchers highlight that this pattern is consistent with previous work by Carroll and Mendels (1976) and Young et al. (1994), where a flattened daily activity (i.e. lower morning cortisol, higher afternoon cortisol) in terms of cortisol in depressives compared with healthy subjects could be shown.

2.3.3 Cortisol a potential biomarker for depression

Thus far, the diagnosis of a depressive disorder is solely based on clinical examination. Taking into account the diagnostic criteria, the resulting subjective assessment of the examiner determines whether a person is depressed or not (Gross, 1982). A subjective assessment is always error prone. A major challenge of the 21st century is to identify biological features that can be used to diagnose or predict a psychiatric disorder, such as a depressive disorder (Pratt & Hall, 2018). Such biological features are also referred to as biomarkers. The Biomarkers Definitions Working Group (2001) defines the term biomarker as follows:

"a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention."

In the past, risk factors could be identified that play a role in the onset of depression, such as age, gender, level of education, previous depressive episodes or familial predisposition (King et al., 2008). In the article *"Suggested Biomarkers for Major Depressive Disorders,"* Hacimuscular (2018) identifies potential biomarkers of depression including *"serum and plasma BDNF, serum IL-1, IL-6, TNF and peripheral RNA expressions, MDA, SOD , Leptin, ghrelin, 5-HT transporter mRNA in blood, hippocampus volume, HPA activity, cortisol response to DEX / CRH"*. Whether or not cortisol – as a biological regulator of HPA activity – can be a biological indicator of depression has not yet been conclusively clarified. However, a biological indicator could help to identify individuals at risk of depression and initiate early treatment. The fact that a variety of studies have shown that depression is associated with altered cortisol levels has led researchers to investigate whether cortisol is a potential biological risk factor regarding the onset of depression. The results of the existing studies are very different: some studies have shown that cortisol is a predictor of depression (e.g. McKinney et al 1997, Pinna et al., 2014, Harris et al., 2000, Herbert et al., 2012, Nabeta et al. 2014, etc.), whereas in other studies this could not be proven (e.g. Akbaraly et al. 2013, Vammen et al. 2014, Vinberg et al. 2014, etc.). It is legitimate to ask why an association between cortisol and depression could be found in some studies and not in others. It may be that these differences are due to the different methods of measuring cortisol or the different types of depression survey instruments used. Another reason could be the different measuring

times when determining the cortisol values. In addition, other factors such as age or gender, etc. could naturally play a role. In order to ascertain this, a detailed investigation of the available study results is necessary.

2.4 Research question

In this work, it will be examined whether aberrant cortisol levels are prognostic factors for the development of depression in the adult population And if yes, whether certain ways of cortisol measurements (salivary, urine, hair, serum) and times and modes of sample taking (awakening cortisol, morning cortisol, evening cortisol, cortisol profiles, DEX-CRH-rest) could be recommended for clinical practice in a pragmatic way.

3 Methodology

3.1 Selection of studies and data extraction

Designs of the primary studies were prospective to examine the predictive value of cortisol status for the development of depression in the cohorts of healthy adults. All studies of the peer-reviewed literature using different databases (Pubmed, PsycInfo, Medline and Embase) with the predefined search terms listed in Appendix A, grey literature, hand search and expert consultation were verified. Two independent reviewers (Daniela Stotz and Mareike Stieg) screened titles and abstracts of the studies for the relevance and suitability, while a third reviewer (Caroline Sievers) resolved disagreements. Data extraction was performed simultaneously by two researchers (Daniela Stotz and Susanne Kirchner). Studies were included according to the PRISMA four-phase flow diagram. Included studies were examined and selected according to study design, participants, intervention, data collection, data analysis, results and relevant contextual information. The results of the two researchers were compared and discussed afterwards to validate and standardize the data. The protocol for this study is published under prospero.

3.2 Inclusion criteria

Observational studies; prospective cohort studies with a minimum length of 6 weeks; studies from 1992 to 2015; studies in which depression is measured at two time points.

3.3 Exclusion criteria

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

3.4 Study population

The study population comprised 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 were stable for at least 6 months without taking antidepressants. The intervention is the cortisol assessment as a prognostic test. The outcome of this review is the occurrence of depression or symptoms of depression. It was defined according to ICD-10 criteria (first appeared 1992, released by the WHO), DSM-IV criteria (first appeared 1994; released by the American Psychiatric Association in 2013) or DSM-V (first appeared 2015), which are internationally established diagnostic frames. The outcome variable was measured by standardized depression rating scales. Both self-report depression scales such as the BDI or the Center for Epidemiologic Studies Depression *Scale* (CES-D) and depression scales completed by the researcher such as the Hamilton Depression Scale (HAM-D) or the Structural Clinical Interview (SCID-I) were allowed. Predictors and outcomes are listed in Table 3.

Table 3.
Predictors and Outcomes.

Hormone of interest	**Neuroendocrine axis**	**Context of interest**	**Disease of interest**
Salivary cortisol Urinary cortisol Hair cortisol Serum cortisol CRH Test DEX-CRH Test	Hypothalamic-pituitary-adrenal axis	Mental health	Major Depressive disorder Dysthymic Disorder Depressive symptoms Adjustment disorder Suicide in Depression

3.5 Assessment of study quality

Quality of primary studies were assessed by using the quality appraisal tool "Newcastle Ottawa Scale" (Appendix B) for observational studies and "Jadad Scale" for interventional studies (Appendix C).

3.6 Newcaste-Ottawa Scale

This scale evaluates the studies in terms of three different aspects: selection of study groups, comparability of study groups and outcome of interest. The assessment is based on a star rating system. For the selection of study groups, the maximum number of stars is four, the

comparability of the studies received one star and the outcome a total of three stars. Therefore, a maximum of eight stars can be awarded**.**

3.7 Jadad Scale

This scale was used to assess the quality of interventional studies. This is used to assess the studies in terms of three different domains: randomization, double blinding and description of withdrawals and dropouts. For the description of the randomization, one point is awarded if it is mentioned and two if it is described and appropriate. For the description of double blinding, one point is awarded for mentioning and two if it is described and appropriate. The description of withdrawals and dropouts is rated with one point if described. Thus, maximum number of points to be achieved is five.

3.8 Search of studies

By combining key words (Cortisol OR DEX/CRH OR "HPA axis") AND (Depression OR Depressive Disorder OR Adjustment Disorder OR Suicide) we found 10,673 appropriate articles. After filtering duplicates, 4,619 articles in English language remained. We included all studies where the correlation between cortisol levels and depression has been examined and which met all inclusion criteria and none of the exclusion criteria. Overall we included 41 studies in our systematic review. Of these 17 were observational studies and 24 interventional studies (Figure 4).

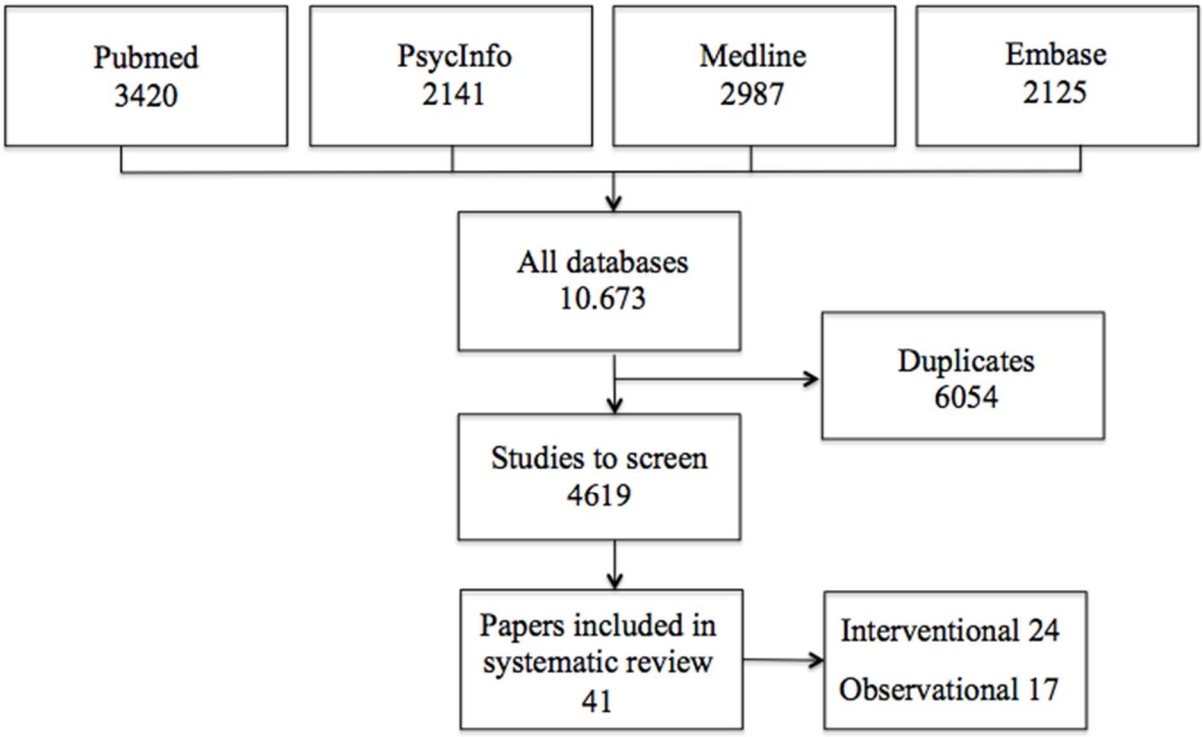

Figure 3. Information PRISMA flow chart during study search and selection.

4 Results

4.1 Patient characteristics

Table 4 shows information about the first authors' location, population, study design, total number of participants, sex, age at baseline, follow-up and funding of the 41 studies included. The total number of participants of all 41 studies was 20,006, their age ranging from 18 to 96, and the follow-up time laid between one month and nine years. Eleven studies were from the UK, eight from the US, four from Sweden, three each from Denmark and the Netherlands, two each from Canada, Japan and Germany and one each from Italy, France, Switzerland, Spain, Australia and Belgium.

Table 4.
Patient characteristics.

First author and year of publication	First author location	Population	Design	Total No. of Participants	Sex (% male)	Age at baseline mean age SD, (range)	Follow up months	Funding
Akbaraly et al. (2013)	UK	Adults	Observation	3145	76.5	60.6 ± 5.9	60	British Medical Research Council; British Heart Foundation; British Health and Safety Executive; British Department of Health, National Heart, Lung and Blood Institute; National Institute of Aging, Agency for Health Care Policy and Research, Economic and Social Research Council professorial fellowship
Berry et al. (2012)	Italy	Elderly patients	Intervention	19	31.6	85 (70-96)	5	Instituto Superiore di Sanita and Nando Peretti Foundation
Bigelow et al. (2012)	Canada	Mothers with postpartum depression	Intervention	90	0	29.4 ± 5.1	3	Nova Scotia Health Research Foundation
Duggal et al. (2013)	UK	Hip fracture patients Healthy controls	Observation	151	27.8	Group 1: 83.8 ± 7.48 Group 2: 84.0 ± 8.62 Group 3: 74.9 ± 5.69	6	Not mentioned
Duggal et al. (2014)	UK	same cohort as Duggal et al. (2013) thus same results	Observation					
Duggal et al. (2015)	UK	same cohort as Duggal et al. (2013) thus same results	Observation					
Edström et al. (2010)	Sweden	Patients with dermatological conditions healthy volunteers	Intervention	99	Not mentioned	(27-59)	1,5	Edvard Welander Foundation and Finsen Foundation
Eriksson et al. (2007)	Sweden	Patients with irritable bowel syndrome	Intervention	42	14.3	(21-67)	3, 6	Not mentioned
Filaire et al. (2001)	France	Soccer players	Intervention	17	100	23.7 ± 2.2	3, 4 and 8	Not mentioned
Gaab et al. (2006)	Switzerland	Healthy economic students	Intervention randomized	28	60.7	Group 1: 22.5 ± 1.7 Group 2: 24.3 ± 4.5	2	Not mentioned
Geoffrey et al. 2013	UK	Men and women from a 1958 British Birth Cohort	Observation	5403	47.7	45	60	Canadian Institutes of Health Research, MRC Career Development Award in Biostatistics, Medical Reserach Council, Department of Health's NIHR Biomedical Research Centre, Medical Reserach Council; Canada Research Charisn PRogram
González-Bono et al. (2002)	Spain	Professional basketball players	Intervention	20	100	Group 1: 21.91 ± 1.07 Group 2: 21.78 ± 1.52	2, 4	Spanish Committee for Scientific and Technical Research; Spanish Superior Council of Sports
Grant (2011)	Australia	Individuals participating in a reality TV show	Intervention	8	50	40.5 (26-63)	6	Not mentioned

Grynderup et al. (2013)	Denmark	Public employess from the municipal and hospital sector	Observation	4467	22.3	515 < 35 682 (35 - 44) 1076 (45 - 54) 585 ≥ 55	24	Danish Work Environment Research Fund, the Lundbeck Foundation
Harris et al. (2000)	UK	Adult women	Observation	116	0	38.5 ± 7.06 (23-58)	13	Medical research council
Herbert et al. 2012	UK	Menopausal women	Observation	279	0	36.8	12, 18	Medical Research Concil Wellcome Trust
Houtveen & Doornen (2008)	Netherlands	Individuals with medically unexplaned symptoms	Observation	128	24.2	Group 1: 37.27 ± 8.34 Group 2: 34.55 ± 8.72 Group 3: 38.33 ± 8.94 Group 4: 36.24 ± 7.70	12	Netherlands Organization for Scientific Research
Jayadevappa et al. (2007)	USA	African American patients ≥ 55 years of age with congestive heart failure	Intervention randomized	23	39.1	Group 1: 64.4 ± 5.7 Group 2: 63.8 ± 8.9	3 and 6	National Institue of Health-National Center for Complementary and Alternative Medicine
Kalmijn et al. (1998)	Netherlands	Healthy elderly participants from the Rotterdam Study	Observation	189	50	67.3 ± 5.7	22,8	Not mentioned
Kuningas et al. (2007)	Netherlands	Elderly aged 85 and older	Observation	563	34%	85	50	IOP grant (Innovative Orientated Research) from the Dutch Ministry of Economic Affairs, Centre for Medical Systems Biology (CMSB), Marie Curie Fellowship of the European Community program EUROGENDIS
Lautenbacher et al. (2010)	Germany	Young male patients with congenital malformation of thorax	Intervention	84	100	21.1 ± 4.5 (16-37)	6	Deutsche Forschungsgemeinschaft
Lieberman et al. (2008)	USA	Female recruits	Intervention	51	0	19.7 ± 2.1	3	Not mentioned
Lieberman et al. (2012)	USA	Female recruits	Intervention	35	0	19.3 ± 1.7	3	U.S. Army Medical Research and Materiel Command
Limm et al.(2014)	Germany	Lower or middle managment employees	Intervention	154	99	Group 1: 40.67 ± 7.62 Group 2: 41.06 ± 7.86	12	Federal Ministry of Education and Research
Lynch et al. (2010)	UK	Students	Intervention	16	25	Group 1: 34.30 ± 12.24 (21-60) Group 2: 28.83 ± 11.05 (19-46)	2,5	Not mentioned
Matousek et al. (2011)	Canada	Women breast cancer after completion of their medical treatment for breast cancer	Intervention	33	0	55.9 ± 10.8 (28-72)	2	Not mentioned
Mc Grady (1994)	USA	Patients with a diagnosis of essential hypertension	Intervention Randomised Control Trial	101	38.6	Group 1: 48 Group 2: 49	10	Hypertension Control Program of the City of Toledo Health Department, from the Ohio Department of Health.
McKinney et al. 1997	USA	Healthy adults	Intervention Randomised Control Trial	28	14.2	(23-45)	1.5, 3	Not mentioned

Nabeta et al. (2014)	Japan	Elderly healthy people	Observation	68	35	Men 72.8 ± 4.9 Female 73.1+-5.4	36	Japan Foundation for Aging and Health. In part by a Japanese Ministry of Education and Science Grant
Pinna et al. (2014)	USA	Older men and women	Intervention	95	33	69.25 ± 10.12 (49-88)	1 and 3	Summa-Kent State Center for the Treatment and Study of Traumatic Stress (DLD). In part, by NIMH grants.
Roberts et al. (2007)	UK	Healthy women (registered nurses)	Intervention	71	0	43 ± 7.1	3	Not mentioned
Romanowska et al. (2011)	Sweden	Leaders and their subordinates	Intervention (randomized, researchers were blinded)	231	not mentioned	Group 1: 51 Group 2: 47	12, 18	Swedish Research Council, Swedbank Research Foundation
Saxton et al. (2014)	UK	Women treated for breast cancer 3 to 18 months previously	Intervention Randomised Control Trial	85	0	Group 1: 55.8 ± 10.0 Group 2: 55.3 ± 8.8	6	American Institute for Cancer Research
Sjödin et al. (2014)	Sweden	Employees at a preschool	Intervention	89	13.5	41.9 ± 9.9	12	AFA Insurance, Stockholm Sweden
T'Sjoen et al. (2005)	Belgium	Elderly aged 70 and older	Observation	236	100	75.3	36	Fund for Scientific Research
Thorn et al. (2010)	UK	Participants with self-assessed seasonal affective disorder and age- and sex-matched healthy controls	Observation	52	34.6	50 ± 12	6	Bial Foundation Portugal
Trueba et al. (2013)	USA	Healthy college students	Intervention	41	17.1	20 (18-21)	1,5	Not mentioned
Vammen (2014)	Denmark	Public sector employees	Observation	3536	22.2	19-66	24	Danish Environment Fund
Vinberg et al. (2014)	Denmark	Healthy monozygotic and dizygotic twins with or without a cotwin history of affective disorder	Observation	234	35	43.9 ± 13.3	108	Danish council for Independent Research and Lundbeck Foundation for their economical support for the study
Williams et al. (2010)	USA	Caregivers of patients with Alzheimer's disease or related dementia	Intervention randomized	116	40.4	60.5 ± 13.4	2, 3, 6	National Institue of Aging
Yoshihara et al. (2014)	Japan	Healthy women	Observation	99	0	Group 1: 36,79 ± 6.43 (25-46) Group 2: 33.84 ± 7.33 (22-49) Group 3: 34.43 ± 8.16 (22-49)	3	Meiji Yasuda Life Foundation of Healthand Welfare, Japan Yoga Therapy Society Research Grant

4.2 Observational studies

In Table 5, all included observational studies are listed. In total, seventeen observational studies were evaluated. Although cortisol and depression values were measured in all observational studies, seven studies did not report their association. In five studies no correlation could be found and in another five studies the outcome was dependent on other factors, such as the time of cortisol collection or gender.

A positive correlation could be found in four studies. Two of these studies are identical with respect to all parameters (Harris et al. 2000; Herbert et al. 2012). In these two studies it could be shown that salivary cortisol levels taken in the morning were positively associated with depression. Herbert et al. (2012) found to be an exact U-shaped relationship between adjusted salivary morning cortisol levels at baseline and the probability of depression onset during the follow-up. In addition, Harris et al. (2000) also provided that cortisol levels taken in the evening were not associated with depression. It is important to mention that both research teams only examined women. Herbert et al. (2012) even investigated menopausal women. The salivary cortisol level was measured on several consecutive days. No statement was made as how long the participants were already awake at the time of taking the salivary cortisol samples. In the following the study parameters: Harris et al. 2000: salivary, 8 o'clock in the morning and evening, SCAN interview, mean age 38.5, follow-up 13 months; Herbert et al.: salivary, 8 o'clock in the morning, SCAN interview, mean age 36.8, follow-up 12 as well as 18 months.

These results are in line with the results of two other studies. In one study, a positive association was found for women when the salivary cortisol samples were determined between 1 and 4 pm (Nabeta et al., 2014) and in the other study when the salivary cortisol samples were taken three hours after awakening (Geoffrey et al., 2013). Cortisol levels were determined in a single day. For men, Nabeta et al. (2014) found no association and Geoffroy et al. (2013) prove a negative correlation. In men with a low cortisol level at the age of 45, depressive symptoms are more likely at the age of 50 years. These studies differ regarding the depression diagnostic tool used, the age and the follow-up time: Nabeta et al. 2014: BDI, mean age 73, 36 months; Geoffrey et al. 2013: MHI-5, mean age 45, 60 months.

Grynderup et al. (2013) also found a negative correlation. The research team studied a mixed cohort, in which men and women were not considered independently. Concerning the type of cortisol measurement, time of measurement and the type of depression measurement the study is similar to that of Harris et al. (2000) and Herbert et al. (2012). The exact time for the morning salivary cortisol determination was 30 min after waking up and all samples

collected up two hours later were accepted. The determination of the cortisol level took place in a single day. For cortisol levels after awakening and in the evening, the research team could not find an association. However, the risk of depression increases by increasing daily mean cortisol concentration and by increasing morning-to-evening slope. Grynderup et al. (2013) refers in his article to the Harris et al. (2000) study and notes the differences between his and the Harris study to several factors. The participants of Grynderup et al. (2013) were more educated, all were employees had a much less history of depression and were not selected because they were likely to develop depression. What Grynderup et al. (2013) does not address is that Harris et al. (2000) studied only women. In the following the study parameters of Grynderup et al. (2013): salivary; morning, evening, profile; SCAN interview, age between 35 and 55 years, follow-up 24 months.

In five studies, no correlation could be proven. Salivary cortisol-awakening samples were collected in two of these five studies (Vammen et al. 2014, Vinberg et al. 2014). In the remaining three studies, the researchers used all serum samples (Akbaraly et al., 2013, Kalmijn et al. 1998, Kuningas et al., 2007), while the dexamethasone test was used in one study (Kalmijn et al., 1998). In all of these studies, mixed cohorts were studied. In the following the exact study parameters: Vammen et al. (2014): salivary, awakening and evening on a single day, SCAN interview, age between 109 and 66, follow-up 24 months; Vinberg et al. (2014): salivary, awakening and evening, SCAN interview, mean age 44 years, follow-up 108 months; Akbaraly et al. (2013): serum, awakening, profile, CES-D, mean age 60.6 years, follow-up 60 months; Kalmijn et al. (1998): serum and dexamethasone, CES-D; mean age 67.3 years, follow-up 22.8 months; Kunings et al. (2007): serum, morning, GDS-15, mean age 85 years; follow-up 50 months.)

4.2.1 Quality of observational studies

The assessment of the individual studies is listed in Table 3. As the evaluation shows, the ratings of most studies range from four to six stars out of the maximum of eight stars according to the Newcastle-Ottawa Scale. Only one study by Houtveen and Doornen et al. (2008) was rated with three stars.

Table 5.
Observational studies.

First author and year of publication	Total No. of Participants	Follow up months	Predictor	Time point of predictor measure	Unit of measurement	Depression assessment method	Association description	Quaity score Newcastle Ottawa Scale
Akbaraly et al. (2013)	3145	60	Serum	Awakening Profile	nmol/l/h	CES-D ≥ 16 0-9 none or minimal 10-16 mild 17-24 moderate > 24 moderate to severe	o	Selection: **, Comparability: *, Outcome: **, Score: 5
Duggal et al. (2013)	101	6	Serum	Morning	mg/ml	GDS ≥ 6 HADS ≥ 8	NR	Selection: **, Comparability: *, Outcome: ***, Score: 6
Duggal et al. (2014) Duggal et al. (2015)								
Geoffrey et al. 2013	9377	60	Salivary	Morning Noon	nmols/l	MHI-S ≤ 52 (das ist für uns relevant) CIS-R	Women + Men -	Selection: **, Comparability: -, Outcome: **, Score: 4
Grynderup et al. (2013)	4467	24	Salivary	Morning Evening Profile	nmol/l	CMDQ (six items - score was 3 or higher on three or more of the six items on the subscale for depression) SCAN-Interview	daily mean cortisol concentration and difference between morning and evening cortisol concentration - evening and morning cortisol concentrations	Selection: **, Comparabilit: - , Outcome: ***; Score 5
Harris et al. (2000)	17	13	Salivary	Morning Evening	ng/ml	SCAN-Interview	Morning + Evening o	Selection: **, Comparability: -, Outcome: **, Score: 4
Herbert et al. 2012	279	12 and 18	Salivary	Morning	ng/ml	SCAN-Interview	U-shaped Morning cortisol +	Selection: **, Comparability: -, Outcome: ***, Score: 5

Houtveen & Doornen (2008)	128	12	Salivary	Awakening Profile	nmol/L	SCL-90-R	NR	Selection: *, Comparability: -, Outcome: **, Score: 3
Kalmijn et al. (1998)	189	22,8	Serum Dexamethasone	Morning	nmol/l	CES-D: ≥ 16 0-9 none or minimal 10-16 mild 17-24 moderate > 24 moderate to severe HADS: ≥ 8	o	Selection: ***, Comparability: -, Outcome: **; Score: 5
Kuningas et al. (2007)	563	50	Serum	Morning	µmol/l	GDS-15	o	Selection: **, Comparability: -, Outcome: **, Score: 4
Nabeta et al. (2014)	68	36	Salivary	Noon	µg/dL	BDI 0-13 no depression 14-19 mild depression 20-28 moderate depression 29-63 severe depression	Women + Men o	Selection: ***, Comparability: *, Outcome: **, Score: 6
T'Sjoen et al. (2005)	238	3	Serum	Morning	µg/dl	GDS ≥ 11	NR	Selection: **, Comparability: -, Outcome: **, Score: 4
Thorn et al. (2010)	52	6	Salivary	Awakening Morning Afternoon Evening	nmol/l	HADS: ≥ 8	NR	Selection: *, Comparability: **, Outcome: ***, Score: 6
Vammen (2014)	3536	24	Salivary	Awakening Evening	nmol/l	CMDQ subscale SCL-DEP6 ≥ 3 SCAN-Interview	o	Selection: **, Comparability: **, Outcome: **, Score: 6
Vinberg et al. (2014)	234	108	Salivary	Awakening Evening	ng/ml	SCAN-Interview	o	Selection: ***, Comparability: -, Outcome: *; Score: 4

Yoshihara et al. (2014)	99	3	Urinary	Morning	µg/gCre.	POMS		NR	Selection: **, Comparability: *, Outcome: **; Score: 5

NR = not reported, o = no association, - = negative association, + = positive association

4.3 Interventional studies

In Table 6, all included interventional studies are listed. We identified 24 studies in which cortisol and depression values were determined. In nineteen studies, cortisol was not examined as a predictor of depression, although all relevant values were collected. In five of the remaining studies, the association between cortisol and depression was investigated and reported. Of the total of five studies in which a relationship has been reported, there are no studies consistent with all parameters such as type of cortisol measurement, time of taking blood, and diagnostic instrument of depression. A positive correlation could be found in two studies (McKinney et al., 1997, Pinna et al., 2014). Mc Kinney et al. (1997) investigated healthy adults in a randomized trial of Bonny Method of Guided Imagery and Music sessions and their effects on mood and cortisol. The research team concluded that a pre-test to the follow-up decrease in cortisol was significantly associated with a decrease in mood disturbance. Pinna et al. (2014) examined 95 people who had a total replacement of their knee joint and found that higher cortisol levels are associated with a greater risk of depression at one and three months following surgery. Apparently, this effect was stronger for men than women at one month following surgery but did not differ between genders at three months post-surgery. The selected study parameters differ and were as follows: Pinna et al., 2014 urinary, profile, CES-D, mean age 69.25; follow-up 1 and 3 months and Mc Kinney et al. 1997: serum, morning, POMS, age 23-45, follow-up 1.5 and 3 months.

In three studies, no association between cortisol and depression was found. Filaire et al. (2001) studied male soccer players and stated that no statistically relevant correlations were found regarding hormonal states and mood states. However, the researchers do not indicate which correlations were calculated exactly. Thus, an exact evaluation of the study is not possible. The study parameter was as follows: salivary; awakening, noon, afternoon; POMS, mean age 23.7, follow-up 3, 4 and 8 months.

Matousek et al. (2011) conducted a study with female participants of a mindfulness-based stress reduction program having undergoing breast cancer treatment. The participants were assessed for stress, depressive symptoms and somatic (e.g. gastrointestinal, respiratory, pain) and psychosocial symptoms (e.g. difficulty in relaxing, sexual difficulties”). The researchers found that the area under the curve (AUC) cortisol-awakening response is negatively correlated with the change in somatic symptoms. They found no association between the AUC cortisol response after awakening and the depressive symptoms. The study parameter was as follows: salivary; awakening, CES-D, mean age 55.9, follow-up 2 months.

Liebermann et al. (2012) who examined also only women could find no association between cortisol and depression. The women participated in an intense twelve-week physical and mental training. In addition to a variety of biomarkers, the predictive relationship between cortisol and depression was also investigated. They found that eleven biomarkers did not predict at least two Profile of Mood States (POMS) questionnaire parameters, one of them cortisol. Liebermann was the only one who used plasma cortisol samples. In the following the study parameters: plasma, morning, POMS, mean age 19,3, follow-up 3 months).

4.3.1 Quality of Interventional studies

The assessment of the individual studies is listed in Table 4. As the evaluation shows, the ratings of most studies range from zero to four out of a maximum of five points.

Table 6.
Interventional studies.

First author and year of publication	Total No. of Participants	Follow up months	Predictor	Time point of predictor measure	Unit of measurement	Depression assessment method	Association description	Quaity score Jadad
Berry et al. (2012)	19	5	Salivary	Morning	µg/dl	GDS-15 > 5	NR	1
Bigelow et al. (2012)	90	3	Salivary	Morning	µg/dl	EPDS >13 CES-D ≧16	NR	1
Edström et al. (2010)	99	1,5	Serum, Plasma, Urinary	Morning	NR	MADRS < 20 non-pathological 20-40 mild depression ≥ 40 severe depression	NR	1
Eriksson et al. (2007)	42	6	Salivary	Morning Noon Afternoon Evening	nmol/l	SCL-90 The higher the score is, the more the symptoms are. A total score and subscales of specific symtpoms were used	NR	1
Filaire et al. (2001)	17	3, 5, 8	Salivary	Awakening Noon Afternoon	nmol/l(-1)	POMS Recommended cut-off scores have not been established although higher scores indicate increased mood disturbance	o	0
Gaab et al. (2006)	28	2	Salivary	Awakening Profile	nmol/l	ADS-K short version of a revised German translation of the Center for Epidemiologic Studies Depression Scale CES-D, it assesses depressive symptoms in non-clinical populations)	NR	3
González-Bono et al. (2002)	20	4	Plasma	Morning	µmol · l^{-1}	Not mentioned	NR	0
Grant (2011)	8	6	Not mentioned	Not mentioned	Not mentioned	Depression, Anxiety and Stress Scale (DASS) Normal 0-9 Mild 10-13 Moderate 14-20 Severe 21-27 Extremely Severe 28+	NR	0

Jayadevappa et al. (2007)	23	3 and 6	Serum	Not mentioned	µg/dl	CES-D: ≥ 16 0-9 none or minimal 10-16 mild 17-24 moderate > 24 moderate to severe	NR	3
Lautenbacher et al. (2010)	84	6	Salivary Dexamethasone 0,5mg	Awakening Morning Cortisol suppression	Morning cortisol (ng/ml min) Cortsiol suppression ng/ml min)	CES-D: ≥ 16 0-9 none or minimal 10-16 mild 17-24 moderate > 24 moderate to severe	NR	1
Lieberman et al. (2008)	51	3 and 25	Serum	Morning	µg/dL	POMS	NR	1
Liebermann et al. (2012)	35	3	Plasma	Morning	NR	POMS	o	1
Limm et al. (2011)	154	12	Salivary	Awakening Morning Area under the curve Profile	NR	HADS: ≥ 8	NR	2
Lynch et al. (2011)	16	2,5	Salivary	Awakening Daily profile	NR	HADS: ≥ 8	NR	1
Matousek et al. (2011)	33	2	Salivary	Awakening	nmol/L	CES-D ≥ 16 0-9 none or minimal 10-16 mild 17-24 moderate > 24 moderate to severe	o	0
Mc Grady (1994)	101	10 months after treatment	Plasma Urinary	Morning Profile	Plasma cortisol (micrograms %) Urinary cortisol (micrograms / gram creatinine)	BDI 0-13 no depression 14-19 mild depression 20-28 moderate depression 29-63 severe depression	NR	2
McKinney et al. 1997	28	1,5 and 3	Serum	Morning	µg/dL	POMS	+	2
Pinna et al. (2014)	95	1 and 3	Urinary	Profile	µg/dL	CES-D > 15 (elderly populations)	+	0

Roberts et al. (2007)	71	3	Salivary	Profile	nM	HADS: ≥ 8	NR	0
Saxton et al. (2014)	85	6	Salivary	Profile	µg/dL	BDI 0-13 no depression 14-19 mild depression 20-28 moderate depression 29-63 severe depression	NR	3
Sjödin et al. (2014)	89	12	Salivary	Awakening Morning Evening	nmol/L	MDI 20 -24 mild depression 25 - 29 moderate depression 30 or more severe depression	NR	0
Trueba et al. (2013)	41	1,5	Salivary	NR	pg/ml	HADS: ≥ 8	NR	0
Romanowska et al. (2011)	231	18	Serum	Morning	nmol/l	SCL-90 The higher the score is, the more the symptoms are. A total score and subscales of specific symtpoms were used	NR	4
Williams et al. (2010)	116	2, 3, 6	Salivary	Awakening Post Awakening Evening	NR	CES-D ≥ 16 0-9 none or minimal 10-16 mild 17-24 moderate > 24 moderate to severe	NR	2

NR = not reported, o = no association, - = negative association, + = positive association

5 Discussion

In this systematic review, 41 observational and interventional studies were evaluated regarding the question whether aberrant cortisol levels are a prognostic factor for the development of depression in the adult population. The picture that emerges from the interventional studies is inconclusive. Only two out of the total 24 interventional studies reported a positive association between cortisol and depression (McKinney et al. 1997; Pinna et al., 2014); and no association could be demonstrated in three studies (Fialire et al., 2001; Liebermann et al., 2012; Matousek et al., 2011). Due to the small number of studies in which an association has been reported and the fact that there are almost no interventional studies that can be compared regarding the parameters used (methods of cortisol measurement, timepoints and conditions, screening instrument for depression), a statement is impossible here.

The observational studies provided a different picture, whereby only ten of the total seventeen observational studies examined reported a correlation. In five studies, no correlation could be found (Vammen et al., 2014; Vinberg et al., 2014; Akbaraly et al., 1998; Kuningas et al., 2007). And in one study, the risk of depression has been shown to increase by increasing daily mean cortisol concentration and morning-to-evening slope (Grynderup et al.). A total of four observational studies reported a positive association. Higher morning cortisol levels in women tend to be a predictor of depression (Harris et al., 2000; Herbert et al., 2012; Nabeta et al., 2014; Geoffroy et al., 2013). In men, the findings are unclear. In one study, no association was found between cortisol and depression (Nabeta et al., 2014), whereas in another study it could be shown that men with lower morning cortisol levels are at risk of developing depression (Geoffroy et al., 2013). In healthy men and women, the cortisol levels are usually indistinguishable (Bangasser et al., 2014; Uhart et al., 2006; Kirschbaum et al. 1999). In women who are suffering from depression, higher cortisol levels can be detected compared with men (Bangassser et al., 2014, Young et al., 1995; Young et al. 2004). This difference is most pronounced after stressful and other negative live events (Bangasser et al., 2014; Peeters et al., 2003; Chopra et al. 2009). Overall, women have a more susceptible HPA axis than men, which slowly returns to baseline after stress (Seidler, Freyberger & Maercker, 2015). Cortisol release in women is cycle-dependent and may be influenced by the use of a contraceptive. In women in the folic phase and in those who take an oral contraceptive, a lower cortisol release could be detected (Kudielka & Kirschbaum, 2005). Whether taking contraceptives has an effect on the cortisol level or not depends on the type of cortisol

measurement used. Thus, after a three-month contraceptive use a higher basal and ACTH-stimulated serum cortisol was found while the basal and stimulated salivary cortisol was not affected (Šimunková et al., 2008). The majority of the studies listed here did not mention whether female participants took oral contraceptive or in which phase of the cycle the women were at the time of the examination. This minimizes the informative value of the studies.

A main task of this systematic review was to ascertain which methods of cortisol measurement, (salivary, urine, hair and serum), which timepoints and conditions (awakening cortisol, cortisol profiles, DEX-CRH tests) and which screening instrument for depression should be used to clarify whether different cortisol levels predict depression. The examination of the types of cortisol measurement shows that serum measurement and measurement of salivary cortisol has been used most frequently. However, only in one study an association between serum cortisol and depression could be demonstrated (Mc Kinney et al., 1997). Whereas in studies in which salivary cortisol was determined more frequently, a positive or negative association could be detected. As already mentioned, salivary cortisol samples offer some advantages compared with serum cortisol samples: obtaining specimens is very easy, it is a non-invasive procedure, it is not bound to any carrier proteins, etc. (Nicolson, 2008). For reliable results, salivary cortisol values should be measured on at least three consecutive days (Segerstorm et al., 2014). Not all of the studies listed in our systematic review have followed the recommendation. For example, in some studies the measurements took place over four days or even longer (Harris et al., 2000; Herbert et al., 2012) and in other studies, cortisol levels were measured on a single day (Geoffroy et al., 2013; Nabeta et al., 2014).

Regarding the time of cortisol determination, the evaluation of the results suggests that morning cortisol is most predictive in determining the association between cortisol and depression. However, it should be noted that the studies we examined in the course of preparing this systematic review did not clearly differentiate between awakening cortisol and morning cortisol. For example, Grynderup et al. (2013) stated in his study that they measured morning cortisol of the participants. The subjects should take the cortisol sample 30 minutes after awakening. However, the researchers accepted all samples collected within 2 hours of awakening. In two other studies, the cortisol samples were taken at eight o'clock in the morning. However, no statement was made as how long the participants were already awake at this time (Harris et al., 2000; Herbert et al.2012). Geoffroy et al. (2013) took the samples 45 min after awakening and 3 hours later and they found a positive association for the late morning cortisol for women and for men a negative association. For the cortisol-awakening response, although the researchers found no association. The level of cortisol-awakening

response is dependent on several factors. Pain, exhaustion, socioeconomic factors, social status or material living standards make the awakening cortisol response lower (Kumari et al., 2009; Fabian et al., 2009; Wright et al., 2005; Ranjit et al. 2005) whereas acute stress leads to an increased awakening cortisol response (Stalder et al. 2016). According to Pruessner et al. (1999) perceived stress correlates with increased cortisol levels while burnout correlates with low cortisol. The latter could be related to the pronounced exhaustion, but this has not yet been sufficiently clarified. Kudielka and Kirschbaum (2003) found that the age and health status seem to have an impact on cortisol-awakening response. Furthermore, the feelings of the previous day, such as loneliness, sadness, threat and lack of control, may also increase the cortisol-awakening response the next day (Adam et al., 2006). Whether it comes to an awakening response and how this fails depends not only on these factors or the awakening alone but especially the timepoint of awakening. A study with nurses shows that the cortisol-awakening response reaches the peak 60 minutes after getting up when the getting up time is between 4 and 5:30 in the morning. However, if the time of getting up is between 6 and 9 o'clock in the morning or between 11 and 14 o'clock, the peak is already reached after 45 minutes. The fact that the cortisol-awakening response can be influenced by so many factors suggests the question of whether this type of measurement is appropriate for determining the association of cortisol and depression. For our systematic review, we found two observational studies (Vammen et al., 2014; Vinberg et al. 2014) and two interventional studies (Filaire et al., 2001; Matousek et al., 2011), in which the cortisol-awakening response and the association between cortisol and depression were determined. In all four studies, no correlation could be found. In addition to the studies from our systematic review, we were able to find two studies in which it was investigated whether the cortisol-awakening response in adolescents is a risk factor for the development of depression. In one study, this could not be confirmed (Carnegie et al., 2014), whereas in the other study, the researchers came to the conclusion that the magnitude of the cortisol-awakening response is predictive for the onset of depression (Adam et al. 2010). In the latter study, the researchers note that their study differed from the previous studies in which no association could be found. Adam et al. (2010) took the cortisol values of their participants not at a specific time but the personal waking times. As already mentioned, in the studies of our systematic review the personal waking time was not sufficiently considered. And a clear separation between the cortisol-awakening response and the morning cortisol levels has not occurred. Based previous data showing by how many factors the cortisol-awakening response can be influenced, this should be taken into account in further investigations. This would increase the validity in terms of determining the

association between cortisol and depression. In any case, the salivary morning cortisol or the salivary awakening cortisol response was most studied as evidenced by the evaluation of our studies.

Based on the data in our systemic review, it seems likely that the choice of diagnostic tool for depression is not critical in determining the association between cortisol and depression. In the four observational studies that reported a positive association for women, different diagnostic tools were used. In two of these studies the SCAN interview was used (Harris et al., 2000; Herbert et al., 2012) in one study the Mental Health Inventory MHI-5, a brief questionnaire was used (Geoffroy et al., 2013) and in the other the BDI, a self-reporting questionnaire (Nabeta et al., 2014). Our assumption is supported by three additional studies not included in our systematic review in which adolescents were examined. In all three studies, researchers concluded that higher salivary cortisol levels increase the risk of depression. However, these studies differed regarding the diagnostic instruments used. In two studies, the Mood and Feeling Questionnaire was used (Goodyer et al., 2000; Halligan et al., 2007) whereas one study used a semi-structured psychiatry interview, the Schedule for Affective Disorders (Goodyer et al., 2010).

As in all systematic reviews, the statements in our work depend on the quality of the studies available to us. As the study quality evaluation using the Newcastle-Ottawa Scale and the Jadad Scale shows, the quality of the observational studies was rather mediocre and the quality of the interventional studies was mostly between zero and two, and only three studies received a three and one study a four. Of course, this quality of the studies significantly reduces the informative value. As our systematic review shows, cortisol is a very susceptible hormone that can be affected by various factors. According to Nicolson (2008), these may be age, gender, ethnicity, somatic changes or illnesses, medication. Moreover, daily changes regarding the eating behavior, smoking or the individual sleep pattern can also affect the cortisol values. Such factors were only partially controlled or were not controlled at all in the studies included in this systematic review. Hence, we cannot judge the extent to which the results were influenced by such factors. As already mentioned, the examination of the individual studies showed that different measurement methods were used concerning measurements and tools for assessing depression. Similarly, the timepoints in time for measuring cortisol varied.

6 Conclusion

In conclusion, our systematic review provides evidence that cortisol measurements may be a future candidate to be used as additional indicator in the prediction of depression in women. In men, the current insufficient research situation does not allow specific statements. However, the reliability of the results depends on a number of factors. Salivary cortisol appears to be the best method in terms of assessing the risk for the development of depression. The use is easy and our systematic review has shown that this type of measurement provides robust results. Regarding the timepoint, the awakening response and the morning cortisol are physiologically the most promising timepoints. However, we could not find any results for the awakening cortisol response in respect to the association between cortisol and depression. Most probably because awakening cortisol response is influenced by many factors such as age, stress, etc. Morning cortisol could be a more stable parameter, although this cannot be conclusively stated based on the existing data. The diagnostic tool for depression diagnosis was not crucial in the investigating studies. For a better comparison with current studies, we recommend using depression diagnostic tools that have been successfully used in determining an association between cortisol and depression, i.e. self-rating tools like the Mental Health Inventory, the BDI, Mood and Feeling Questionnaire or interviews like Schedule for Affective Disorders or Schedules for Clinical Assessment in Neuropsychiatry. For a reliable statement, further studies should be carried out, in which the waking times of the participants are considered and stated and in which the samples are collected on three consecutive days. Since cortisol is a very sensitive parameter, factors such as gender, age, illnesses, stress, medication contraceptive, eating behavior, smoking and individual sleep should be controlled to allow for a general recommendation and prospective use of cortisol as a predictor of depression.

List of Figures

List of Tables

Reference List

Abel, U., Hautzinger M. (2013). *Kognitive Verhaltenstherapie bei Depressionen im Kindes- und Jugendalter* (p.14). Berlin: Springer.

Abelson, J.L., & Curtis, G.C. (1996). Hypothalamic-pituitary-adrenal axis activity in panic disorder. 24-hour secretion of corticotropin and cortisol. *Arch Gen Psychiatry, 53*, 323–31.

Adam, E.K., Hawkley, L.C., Kudielka, B.M., Cacioppo, J.T. (2006). Day-to-day dynamics of experience-cortisol associations in a population-based sample of older adults. *Proceedings of National Academy of Science USA, 103*, 17058-17063.

Adam, E. K., Doane, L. D., Zinbarg, R. E., Mineka, S., Craske, M. G., Griffith, J. W. (2010). Prospective prediction of major depressive disorder from cortisol awakening responses in adolescence. *Psychoneuroendocrinology, 35(6)*, 921–931.

Akbaraly, T. N., Kumari, M., Head, J., Ritchie, K., Ancelin, M. L., Tabak, A. G., … Kivimaki, M. (2013). Glycemia, insulin resistance, insulin secretion, and risk of depressive symptoms in middle age. Diabetes Care, 36(4), 928–934.

Albert, P.R. (2015) Why is depression more prevalent in women? *J Psychiatry Neurosci, 40*, 219–221.

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: American Psychiatric Association.

Andrade, L., Caraveo-Anduaga, J.J., Berglund, P., Bijl, R.V., De Graaf, R., Vollebergh, W., Wittchen, H.-U. (2003). The epidemiology of major depressive episodes. *Int J Methods Psychiatr Res.*, *12*, 3-21.

Arana, G.W., Santos, A., Laraia, M.T. (1995). Dexamethasone for the treatment of depression: a randomised, placebo-controled, double blind trial. *Am J Psychiat, 152,* 265–267.

Bhagwagar, Z., Hafizi, S., Cowen, P.J. (2005). Increased salivary cortisol after waking in depression. *Psychopharmacology,182*, 54 –7.

Bangasser, D. A., & Valentino, R. J. (2014). Sex differences in stress-related psychiatric disorders: neurobiological perspectives. *Frontiers in neuroendocrinology, 35(3)*, 303-19.

Bartels, M., Cacioppo, J. T., van Beijsterveldt, T. C., & Boomsma, D. I. (2013). Exploring the association between well-being and psychopathology in adolescents. *Behavior Genetics, 43*, 177–190.

Beck, A.T., Hautzinger, M., Bailer, M., Worall, H., Keller F. (1995). *Beck-Depressions-Inventar (BDI)* (2th ed.). Bern: Huber.

Beesdo-Baum, K., Knappe, S., Einsle, F., Knothe, L., Wieder, G., Venz, J., Bergmann, A. (2018). Wie häufig werden Patienten mit depressiven Störungen in der hausärztlichen Praxis erkannt? [How often are patients with depressive disorders recognized in primary care practice]. *Bundesgesundheitsblatt, 61(1)*, 52-64.

Berger, M., Klein, H.E. (1984). Dexamethasone suppression test: a biologic marker of endogenous depression? *Eur Arch Psychiatry Neurol Sci, 234(2)*, 137-46.

Berry, A., Borgi, M., Terranova, L., Chiarotti, F. Alleva, E., Cirulli, F. (2012). Developing effective animal-assisted intervention programs involving visiting dogs for institutionalized geriatric patients: a pilot study. *Psychogeriatrics, 12*, 143-150

Beuschlein, G., & Reincke, M. (2006). Nebennierenrinde [adrenal cortex]. In W. Siegenthaler, H.E. Blum (Eds.), *Klinische Pathophysiologie* (9th ed., p.333). Stuttgart: Thieme.

Bhagwagar, Z., Hafizi, S., & Cowen, P.J. (2005). Increased salivary cortisol after waking in depression. *Psychopharmacology, 182(1)*, 54–7.

Bigelow, A., Power, M., MacLellan-Peters, J., Alex, M., Mc Donald, C. (2012). Effect of mother/infant skin-to-skin contact on post partum depressive symptoms and maternal physiological stress. *JOGNN, 41*, 369-382.

Biomarkers Definitions Working Group (2001). Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther., 69*, 89-95.

Bley, C.-H., Centgraf M., Cieslik, A., Haack J., Hohloch L. (2015). *I care, Anatomie Physiologie* (p.265). Stuttgart: Thieme.

Bouman, T.K. (1993). Einschätzung von Stimmungsstörung [Assessment of mood disorders]. In F. A. Albersnagel, P.M.G. Emmelkamp, R.H. Van den Hoffdakker (Eds.) *Depression* (pp. 45-62). Göttingen: Verlag für angewandte Psychologie.

Bramesfeld, A.; Stoppe, G. (2006) Einführung [Introduction]. In G.Stoppe, A.Bramesfeld, F.-W. Schwartz (Eds.), *Volkskrankheit Depression* (p.1). Heidelberg: Springer.

Brown, G.T. & Harris, T. (1978). *Social origin of depression. A study of psychiatric disorders in women* (p.240). London: Tavistock.

Burke, H. M., Davis, M. C., Otte, C., Mohr, D. C. (2005) Depression and cortisol responses to psychological stress: a meta-analysis. *Psychoneuroendocrinology, 30*, 846–856.

Carnegie R., Araya R., Ben-Shlomo Y., Glover V., O'Connor T.G., O'Donnell K.J., Pearson R., Lewis G. (2014). Cortisol awakening response and subsequent depression: prospective longitudinal study. *Brit J Psychiat., 204*, 137-143.

Carroll, B.J., Curtis, G.C., Davies, B.M., Mendels, J., Sugerman, A.A. (1976). Urinary free cortisol excretion in depression. *Psychol Med, 6*, 43-50.

Carroll, B.J., Mendels, J. (1976). Neuroendocrine regulation in affective disorders. In E.J. Sachar (Ed.) *Hormones, Behavior and Psychopathology.* New York: Raven Press.

Carroll, B. J., Feinberg, M., Greden, J. F., Tarika, J., Albala, A. A., Haskett, R. F., Young, E. (1981). A specific laboratory test for the diagnosis of melancholia. Standardization, validation, and clinical utility. *Archives of General Psychiatry, 38(1)*, 15–22.

Chopra, K.K., Ravindran, A., Kennedy, S.H., Mackenzie, B., Matthews, S., Anisman, H., Levitan, R.D. (2009). Sex differences in hormonal responses to a social stressor in chronic major depression. *Psychoneuroendocrinology, 34*, 1235-1241.

Claustrat, B., Chazot, G., Brun, J., Jordan, D., Sassolas, G. (1984). A chronobiological study of melatonin and cortisol secretion in depressed subjects: plasma melatonin, a biochemical marker in major depression. *Biol Psychiatry, 19,* 1215–1228.

Dilling, H., Mombour W., Schmidt, M.H. (2014). *Internationala Klassifikation psychischer Störungen: ICD-10 Kapitel V(F) klinisch-diagnostische Leitlinien* (pp.156-186). [International classification of psychic Interference: ICD-10 Chapter V (F) Clinical Diagnostic Guidelines]. Bern: Huber.

Duggal, N.A., Upton, J., Phillips, A.C., Hampson, P., Lord, J.M. (2013). Depressive symptoms are associated with reduced neutrophil function in hip fracture patients. *Brain Behav. Immun., 33*, 173-182.

Duggal, N.A., Beswetherick, A., Upton, J., Hampson, P., Phillips, A.C., Lord, J.M. (2014). Depressive symptoms in hip fracture patients are associated with reduced monocyte superoxide production. *Exp. Gerontol., 54*, 27–34.

Duggal, N.A., Upton, J., Phillips, A.C., Hampson, P., Lord, J.M. (2015). NK cell immunesenescence is increased by psychological but not physical stress in older adults associated with raised cortisol and reduced perforin expression. *Age (Dordr), 37*, 9748.

Edwards, S., Evans, P., Hucklebridge, F., Clow, A. (2001). Association between time of awakening and diurnal cortisol secretory activity. *Psychoneuroendocrinology, 26*, 613–622.

Edström, D.W., Linder, J., Wennersten, G., Brismar, K., Ros, A.M. (2010). Phototherapy with ultraviolet radiation: a study of hormone parameters and psychological effects. *J Eur Acad Dermatol Venereol., 24*, 403-409.

Ehret, A., & Berking, M. (2013). DSM-IV und DSM-5 was hat sich tatsächlich verändert? [DSM-IV and DSM-5 what has actually changed?]. *Verhaltenstherapie 23,* 258-266.

Ellenbogen, M.A., Hodgins, S., Linnen, A.M., Ostiguy, C.S. (2011). Elevated daytime cortisol levels: a biomarker of subsequent mahor affective disorder? *J. Affect. Disord., 132*, 265-269.

Eriksson, E. M., Moller, I. E., Soderberg, R. H., Eriksson, H. T., Kurlberg, G. K. (2007). Body awareness therapy: A new strategy for relief of symptoms in irritable bowel syndrome patients. *World Journal of Gastroenterology, 13(23)*, 3206–3214.

Ermann, M. (2007). *Psychosomatische Medizin und Psychotherapie.* Ein Lehrbuch auf psychoanalytischer Grundlage (2th ed., p.200) [Psychosomatic medicine and psychotherapy. A textbook on a psychoanalytical basis]. Stuttgart: Kohlhammer.

Fabian, L.A., McGuire, L., Page, G.G., Goodin, B.R., Edwards, R.R. (2009). The association of the cortisol awakening response with experimental pain ratings *Psychoneuroendocrinology, 34 (8)*, 1247-1251.

Fang, V.S., Tricou, B.J., Robertson, A., Meltzer, H.Y. (1981). Plasma ACTH and cortisol levels in depressed patients: relation to dexamethasone suppression test. *Life Sci, 29*, 931-938.

Faumann, M.A. (2001). *Study Guide to DSM-IV-TR* (p.174-175). Washington DC: American Psychiatric Association.

Federenko, I., Wüst, S., Hellhammer, D.I., Dechoux, R., Kumsta, R., Kirschbaum, C. (2004). Free cortisol awakening responses are influenced by awakening time. *Psychoneuroendocrinology, 29*, 174-184.

Filaire, E., Bernain, X., Sagnol, M., Lac, G. (2001). Preliminary results on mood state, salivary testosterone: cortisol ratio and team performance in a professional soccer team. *Eur J Appl Physiol, 86 (2)*, 179-84.

Froesch, E.R. (1981) Innere Sekretion [inner secretion]. In A.A. Bühlmann, E.R. Froesch. *Pathophysiologie* (4th ed., p.223). Berlin Heidelberg New York: Springer.

Frye, A. A., & Garber, J. (2005). The relations among maternal depression, maternal criticism, and adolescents' externalizing and internalizing symptoms. *Journal of Abnormal Child Psychology, 33*, 1–11.

Gaab, J., Sonderegger, L., Scherrer, S., Ehlert, U. (2006). Psychoneuroendocrine effects of cognitive-behavioral stress management in a naturalistic setting – a randomized controlled trial. *Psychoneuroendocrinology, 31(4)*, 428-438.

Gatti, R., Antonelli, G., Prearo, M., Spinella, P., Cappellin, E., De Palo, E.F. (2009). Cortisol assays and diagnostic laboratory procedures in human biological fluids. *Clin. Biochem. 42*, 1205-1217.

Geoffroy, M. C., Hertzman, C., Li, L., Power, C. (2013). Prospective association of morning salivary cortisol with depressive symptoms in mid-life: A life-course study. *PLoS ONE, 8(11)*, 1.

Gerber, M., Kalak, N., Elliot, C., Holsboer-Trachsler, E., Pühse, U., Brand, S. (2013). Both hair cortisol levels and perceived stress predict increased symptoms of depression: an exploratory study in young adults. *Neuropsychobiology, 68(2)*, 100–109.

Gierk, B. & Wahl, I. (2013). Depressionsfragebögen für den alltäglichen klinischen Alltag [Depression questionnaires for the everyday clinical routine]. In M. Borca, M. Broda, V. Köllner, H. Schauenburg, W. Senf, B. Stein. B. Willms, *Psychotherapie im Daialog. Systemische Therapie, Verhaltenstherapie, Humanistische Therapie. Diagnostik und Evaluation* (1th ed.). Thieme. Retrieved from https://books.google.de/books?id=FBfZHHtqw1gC&pg=PA37&dq=Depression+Frageb%

C3%B6gen&hl=de&sa=X&ved=0ahUKEwiFvMiei_jgAhURJVAKHT3lB5sQ6AEIOTAD#v=onepage&q=Depression%20Frageb%C3%B6gen&f=false.
Gillies, G.E., McArthur, S. (2010). Estrogen actions in the brain and the basis for differential action in men and women: a case for sex-specific medicines. *Pharmacol Rev., 62*,155–98.
Giltay, E.J., Enter, D., Zitman, F.G., Penninx, B.W.J.H., van Pelt, J., Spinhoven, P., Roelofs, K. (2012). Salivary testosterone: associations with depression, anxiety disorders, and antidepressant use in a large cohort study. *Journal of psychosomatic research, 72(3)*, 205-13.
Goebel, M.U., & Schedloswki M. (2008). Immunologische Erkrankungen: Rhema, Lupus erythematodes und HIV-Infektion [Immunological disorders: Rhema, lupus erythematosus and HIV infection]. In U. Ehlert (Ed.) *Verhaltensmedizin* (p.457). Heidelberg: Springer.
González-Bono, E., Salvador, A., Serrano, M.A., Moya-Albiol, L., Martínez-Sanchis, S. (2002). Effects of Training Volume on Hormones and Mood in Basketball Players. *Int J Stress Management, 9(4)*, 263-273.
Goodyer, I.M., Herbert, J., Tamplin, A., Altham, P.M.E. (2000). Recent life events, cortisol, dehydroepiandrosterone and the onset of major depression in high-risk adolescents. *Br. J. Psychiatry, 177*, 499–504
Goodyer, I., Croudace, T., Dudbridge, F., Ban, M., Herbert, J. (2010). Polymorphisms in BDNF (Val66Met) and 5-HTTLPR, morning cortisol and subsequent depression in at-risk adolescents. *British Journal of Psychiatry, 197*, 365–371.
Grant, A. M. (2011). Reality TV gets positive: Psychological reflections on Making Australia Happy. *International Coaching Psychology Review, 6(2)*, 229-239.
Gross, J. (1982). Der protrahierte Suizid. In C. Riemer (Ed.) *Suizid Ergebnisse und Therapie* (p.52) Berlin: Springer.
Gruenberg, A.M., Goldstein, R.D., Pincus, H.A. (2005). Classification of depression: Research and diagnostic criteria: DSM-IV and ICD-10. In Licinio J, Wong M.-L., (Eds.) *Biology of depression. From novel insights to therapeutic strategies* (pp1-12). Weinheim: Wiley-VCH.
Grynderup, M.B., Kolstad, H.A., Mikkelsen, S., Andersen, J.H., Bonde, J.P., Buttenschøn, H.N., Hansen, A.M. (2013). A two-year follow-up study of salivary cortisol concentration and the risk of depression. *Psychoneuroendocrinology, 38(10)*, 2042–50.
Hacimusalar, Y, Eşel, E. (2018) Suggested biomarkers for major depressive disorder. Noro *Psikiyatr Ars, 55(3)*, 280-290.
Halbreich, U., Asnis, G.M., Shindledecker, R., Zumoff, B., Nathan, R.S. (1985). Cortisol Secretion in Endogenous Depression I. Basal Plasma Levels. *Arch Gen Psychiatry, 42(9)*, 904–908.
Halligan, S.L., Herbert, J., Goodyer, I., Murray, L. (2007). Disturbances in morning cortisol excretion in association with maternal postnatal depression predict subsequent depressive symptomatology in adolescents. *Biol Psychiat, 62*,40–46.
Hamilton, M. (1980). Rating depressive patients. *Journal of Clinical Psychiatry, 41*, 21-24.
Hammen, C., Brennan, P. A. & Shih, J. H. (2004). Family discord and stress predictors of depression and other disorders in adolescent children of depressed and nondepressed women. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 994-1002.
Harris, T.O., Borsanyi, S., Messari, S., Stanford, K., Cleary, S.E., Shiers, H.M., Brown, G.W., Herbert, J. (2000). Morning cortisol as a risk factor for subsequent major depressive disorder in adult women. *The British Journal of Psychiatry, 177(6)*, 505–510.
Hautzinger M., & Bailer M. (1993). *Allgemeines Depressions-Skala* [General depression scale]. Weinheim: Beltz Test.

Hegerl M, Althaus D, Reiners H (2005). *Das Rätsel Depression: Eine Krankheit wird entschlüsselt* (p.9) [The mystery of depression: becoming a disease decrypts]. Beck: München.

Herbert, J., Ban, M., Brown, G.W., et al., 2012. Interaction between the BDNF Val/66/Met polymorphism and individual levels of morning cortisol as a predictor of depression in adult women. *Br J Psychiatry, 201*, 313–319.

Hinson J., Raven P., Chew S. (2018). *Organsysteme verstehen: Endokrines System. Integrative Grundlagen und Fälle* (p.64) [Understanding Organ Systems: Endocrine System. Integrative basics and cases]. München: Elsevier.

Holsboer, F. (2000). The Corticosteroid Receptor Hypothesis of Depression. *Neuropsychopharmacology, 23(5)*, 477–501.

Houtveen, J. H., & Van Doornen, L. J. P. (2008). Negative affect and 24-hour ambulatory physiological recordings as predictors of spontaneous improvement of medically unexplained symptoms. *Scand J Psychol, 49(6)*, 591-601.

Jacobi, F., Wittchen, H. U., Holting, C., Hofler, M., Pfister, H., Muller, N., Lieb, R. (2004). Prevalence, co-morbidity and correlates of mental disorders in the general population: results from the German Health Interview and Examination Survey (GHS). *Psychol Med, 34(4)*, 597-611.

Jayadevappa, R., Johnson, J.C., Bloom, B.S., Nidich, S., Desai, S., Chhatre, S., Raziano, D.B., Schneider, R. (2007). Effectiveness of transcendental meditation on functional capacity and quality of life of African Americans with congestive heart failure: A randomized control study. *Ethnicity and Disease, 17(1)*, 72-77.

Juruena, M. F., Papadopoulos, A., Cleare, A., Poon, L., Lightman, S., Pariante, C. M. (2005). Treatment resistant depression: different response to dexamethasone and prednisolone test. ECNP Hot Topic Award. *Eur Neuropsychopharmacol, 15*, 414.

Kalmijn, S., Launer, L.J., Stolk, R.P., de Jong, F.H., Pols, H.A., Hofman, A., Breteler, M.M., Lamberts, S.W. (1998). A prospective study on cortisol, dehydroepiandrosterone sulfate, and cognitive function in the elderly. *J. Clin. Endocrinol. Metab., 83*, 3487–3492.

Katz, R., Shaw, B.F., Vallis, T.M., Kaiser A.S. (1995). The assessment of severity and symptom patterns in depression. In: Beckham E.E., Leber W.R. (eds.) *Handbook of depression (pp.61-85).* New York: Guilford.

Kendler, K.S., Gardner, C.O. (2014). Sex differences in the pathways to major depression: a study of opposite-sex twin pairs. *Am J Psychiatry, 171*, 426-35.

King, M., Walker, C., Levy, G., Bottomley, C., Royston, P., Weich, S., Nazareth, I. (2008). Development and validation of an international risk prediction algorithm for episodes of major depression in general practice attendees: the PREDICT study. *Arch Gen Psychiatry, 65*, 1368–76.

Kirschbaum, C., Kudielka, B.M., Gaab, J., Schommer, N.C., Hellhammer, D.H. (1999). Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamuspituitary-adrenal axis. *Psychosom Med, 61,* 154-162.

Kölch, M., Fegert, J.M., Freyberger, H.J. (2011). Affektive Störungen [Affective disorders]. In J.M. Fegert, A. Streeck-Fischer, H.J. Freyberger (Eds.), *Kompendium Adoleszenzpsychiatrie.* Stuttgart: Schattauer.

Kronmüller, K.T. & Mundt, C. (2000). Depressive Episoden [depressive episodes]. In H. Helmchen, F. Henn, H. Lauter, N. Sartorius (Eds.). *Psychiatrie der Gegenwart. Schizophrene und affektive Störungen* (5th ed, pp 293-294). Heidelberg: Springer.

Kudielka, B.M., Kirschbaum, C. (2003). Awakening cortisol responses are influenced by health status and awakening time but not by menstrual cycle phase. *Psychoneuroendocrinology, 28*, 35–47.

Kudielka, B.M., Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. *Biol. Psychol., 69*, 113–132.

Kulkarni, J., Gavrilidis, E., Worsley, R. (2016). Hormones and Schizophrenia. In M. V. Pletnikov, J. L. Waddington. Modeling the Psychopathological Dimensions of Schizophrenia. London: Academic Press.

Kumari, M., Badrick, E., Chandola, T., Adam, E.K., Stafford, M., Marmot, M.G., et al. (2009). Cortisol secretion and fatigue: Associations in a community based cohort. *Psychoneuroendocrinology, 34(10)*, 1476–1485.

Kuningas, M., de Rijk, R.H. , Westendorp, R.G.J. , Jolles J. , Eline Slagboom, P. , van Heemst D. (2007). Mental performance in old age dependent on cortisol and genetic variance in the mineralocorticoid and glucocorticoid receptors. *Neuropsychopharmacology, 32*,1295–1301.

Lautenbacher, S., Huber, C., Schofer, D., et al. (2010). Attentional and emotional mechanisms related to pain as predictors of chronic postoperative pain: a comparison with other psychological and physiological predictors. *Pain, 151*, 722–731.

Lieberman, H., Kellogg, M., Bathalon, G. (2008). Female marine recruit training: mood, body composition, and biochemical changes. *Medicine and Science in Sports and Exercise, 40(11)*, 671-676.

Lieberman, H. R., Kellogg, M. D., Kramer, F. M., Bathalon, G. P., Lesher, L. L. (2012). Lipid and other plasma markers are associated with anxiety, depression, and fatigue. *Health Psychology, 31*, 210–216.

Limm, H., Gündel, H., Heinmüller, M., Marten-Mittag, B., Nater, U.M., Siegrist, J. (2011). Stress management interventions in the workplace improve stress reactivity: a randomised controlled trial. *Occup Environ Med., 68*, 126-133.

Lynch, S., Gander, M. L., Kohls, N., Kudielka, B., Walach, H. (2011). Mindfulness-based coping with university life: A nonrandomized wait-list controlled pilot evaluation. *Stress and Health, 27*, 365–375.

Matousek, R.H., Pruessner, J.C., Dobkin, P.L. (2011). Changes in the cortisol awakening response (CAR) following participation in mindfulness-based stress reduction in women who completed treatment for breast cancer. *Complement Ther Clin Pract, 17(2)*, 65–70.

McGrady, A. (1994). Effects of group relaxation training and thermal biofeedback on blood pressure and related physiological and psychological variables in essential hypertension. *Biofeedback & Self Regulation, 19(1)*, 51–66.

McKinney, C. H., Antoni, M. H., Kumar, M., Tims, F. C., and McCabe, P. (1997). Effects of guided imagery and music (GIM) therapy on mood and cortisol in healthy adults. *Health Psychol., 16*, 1–12.

Murphy, B. E. P. (1999). Lack of specificity of urinary free cortisol determinations: why does it continue? *J Clin Endocrinol Metab, 84*, 2258-9.

Murphy, B. E. P. (2002). Urinary free cortisol determinations: what they measure. *Endocrinologist, 12*, 143–50.

Möller H.-J., Laux, G., Deister A. (2009). *Psychiatrie und Psychotherapie* (4th ed., pp.76-108). Stuttgart: Thieme.

Nabeta, H., Mizoguchi, Y., Matsushima, J., Imamura, J., Watanabe, I., Tateishi, T., Monji, A. (2014). Association of Salivary Cortisol Levels and Later Depressive State in Elderly People Living in A Rural Community: A 3-Year Follow-up Study. *Journal of Affective Disorders, 158,* 85-89.

Nicolson, N.A. (2008). Measurement of cortisol. In: Luecken, L.J., Gallo, L.C. (Eds.), *Handbook of Physiological Research Methods in Health Psychology* (pp.37-74). London: Sage Publications.

O'Toole, M.T. (2017). *Mosby's Pocket Dictionary of Medicine, Nursing & Health Professions* (8th ed., p.381) Louis, Missouri: Elsevier.

Pariante, C.M. (2003). Depression, Stress and the Adrenal axis. *Journal of Neuroendocrinology, 15*, 811-812.

Pariante, C.M. (2006) The glucocorticoid receptor: part of the solution or part of the problem? J. *Psychopharmacol., 20*, 79–84.

Pariante, C.M., Lightman, S.L. (2008). The HPA axis in major depression: classical theories and new developments. *Trends Neurosci, 31*, 464-468.

Peeters, F., Nicholson, N.A., Berkhof, J. (2003). Cortisol responses to daily events in major depressive disorder. *Psychosom. Med., 65*, 836–841.

Pinna, K.L., Johnson, D.M., Delahanty, D.L. (2014). PTSD, comorbid depression, and the cortisol waking response in victims of intimate partner violence: preliminary evidence. *Anxiety Stress Coping 27*, 253–69.

Piper W. (2007). *Innere Medizin* [internal medicine] (p.564). Heidelberg Springer.

Piwowarska, J., Chimiak, A., Matsumoto, H., Dziklińska, A., Radziwoń-Zaleska, M., Szelenberger, W., Pachecka, J. (2012). Serum cortisol concentration in patients with major depression after treatment with fluoxetine. *Psychiatry Research, 198*, 407–411.

Pratt, J., Hall, J. (2018). Biomarker in Neuropsychiatry: A Prospect for the Twenty-First Century? In J. Pratt, J. Hall (Eds.) *Biormarkers in Psychiatry* (p.3). Switzerland: Springer.

Pruessner, J., Hellhammer, D.H., Kirschbaum, C. (1999). Burnout, perceived stress and salivary cortisol upon awakening. *Psychosomatic Medicine, 61*, 197–204.

Ranjit, N., Young, E.A., Kaplan, G.A. (2005). Material hardship alters the diurnal rhythm of salivary cortisol. *Int J Epidemiol, 34*, 1138 –1143.

Rensing, R., Koch, M., Rippe, B., & Rippe, V. (2006): *Mensch im Stress. Psyche, Körper, Moleküle* (p.95) [human in stress. Psyche, body, molecules]. München: Elsevier

Roberts, C., Troop, N., Connan, F., Treasure, J., Campbell, I. C. (2007). The effects of stress on body weight: Biological and psychological predictors of change in BMI. *Obesity, 15*, 3045-3055.

Romanowska, J., Larsson, G., Eriksson, M., Wikström, B., Westerlund, H., Theorell, T. (2011). Health effects on leaders and coworkers of art-based leadership development program. *Psychotherapy and Psychosomatics, 80(2)*, 78-87.

Roth E., Flich K., Huber J. (2018). Endokrines System [endocrine system]. In N. Bachl, H. Löllgen, H. Tschan, H. Wackertage, B. Wessner (Eds.). *Molekulare Sport und Leistungsphysiologie* (p.237). Wien: Springer.

Rush, A.J., Giles, D.E., Schlesser, M.A., Fulton, C.L., Weissenburger, J.E. and Burns, C.T. (1986). The Inventory of Depressive Symptomatology (IDS): Preliminary findings. *Psychiatry Research, 18*, 65-87.

Rush, A.J., Gullion, C.M., Basco, M.R., Jarrett, R.B. and Trivedi, M.H (1996). The Inventory of Depressive Symptomatology (IDS): Psychometric properties. *Psychological Medicine, 26*, 477-486.

Sauve, B., Koren, G., Walsh, G., Tokmakejian, S., Van Uum, S.H. (2007). Measurement of cortisol in human hair as a biomarker of systemic exposure. *Clin. Invest. Med. 30*, E183-E191.

Saxton, J.M., Scott, E.J., Daley, A.J., Woodroofe, M., Mutrie, N., Crank, H., Powers, H.J., Coleman, R.E. (2014). Effects of an exercise and hypocaloric healthy eating intervention on indices of psychological health status, hypothalamic-pituitary-adrenal axis regulation and immune function after early-stage breast cancer: a randomised controlled trial. *Breast Cancer Res, 16(2)*, R39.

Schulte-Körne, G., Allgaier, K. (2008). Genetik depressiver Störungen [genetic of depressive disorders]. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie, 36*, 27-43.

Schützwohl M, Kallert T, Jurjanz L. (2007). Dimensionale Diagnostik mit der deutschsprachigen Version der Schedules for Clinical Assessment in Neuropsychiatry 2.1. *Nervenarzt, 78*, 304–313.

Scott, L.V., Dinan, T.G. (1998). Urinary free cortisol excretion in chronic fatigue syndrome, major depression and in healthy volunteers. *J Affect Disord, 47(1-3)*, 49-54.

Segerstrom, S.C., Boggero, I.A., Smith, G.T., Sephton, S.E. (2014). Variability and reliability of diurnal cortisol in younger and older adults: implications for design decisions. *Psychoneuroendocrinology 49*, 299-309.

Seidler, G.H., Freyberger, H.J., Maercker, A. (2015). *Handbuch der Psychotraumatologie* (2th ed., p.114) [Handbook of psychotraumatology]. Stuttgart: Klett-Cotta.

Sievers, C., Auer, M.K., Klotsche, J., Athanasoulia, A.P., Schneider, H.J., Nauck, M., Grabe, H.J. (2014) IGF-I levels and depressive disorders: results from the Study of Health in Pomerania (SHIP). *European Neuropsychopharmacology, 24(6)*, 890–896.

Simunkova, K., Starka, L., Hill, M., Kriz, L., Hampl, R., Vondra, K. (2008). Comparison of total and salivary cortisol in a low-dose ACTH (Synacthen) test: Influence of three-month oral contraceptives administration to healthy women. *Physiol. Res., 57*, 193–199.

Sjödin, F., Kjellberg, A., Knutsson, A., Landström, U., Lindberg, L. (2014). Measures against preschool noise and its adverse effects on the personnel: An intervention study. *Int. Arch. Occup. Environ. Health, 87*, 95–110.

Stalder, T., Kirschbaum, C., Kudielka, B.M., Adam, E.K., Pruessner J.C., Wüst, S., Clow, A. (2016). Assessment of the cortisol awakening response: Expert consensus guidelines. *Psychoneuroendocrinology, 63*, 414.

Staufenbiel S, Penninx BW, Spijker AT, Elzinga B, van Rossum EF. (2013). Hair cortisol, stress exposure, and mental health in humans: a systematic review. *Psychoneuroendocrinology, 38*, 1220–38.

Sullivan, P. F., Neale, M. C., Kendler, K. S. (2000). Genetic epidemiology of major depression: review and metaanalysis. *Am. J. Psychiatry 157*, 1552–1562.

Thorn, L., Evans, P., Cannon, A., Hucklebridge, F., Clow, A. (2011). Seasonal differences in the diurnal pattern of cortisol secretion in healthy participants and those with self-assessed seasonal affective disorder. *Psychoneuroendocrinology, 36(6)*, 816-823.

Trueba, A. F., Smith, N. B., Auchus, R. J. and Ritz, T. (2013). Academic exam stress and depressive mood are associated with reductions in exhaled nitric oxide in healthy individuals. *Biological Psychology, 93*, 206-212.

T'Sjoen, G.G., De Vos, S., Goemaere, S., Van Pottelbergh, I., Dierick, M., Van Heeringen, C., Kaufmann J.-M. (2005). Sex steroid level, androgen receptor polymorphism, and depressive symptoms in healthy elderly men. *Journal of the American Geriatrics Society, 53*, 636–642.

Uhart, M., Chong, R., Oswald, L., Lin, P.I., Wand, G., (2006). Gender differences in hypothalamic-pituitary-adrenal (HPA) axis reactivity. *Psychoneuroendocrinology, 31*, 642–652.

Uher, R., Payne, J.L., Pavlova, B., Perlis, R.H. (2014). Major depressive disorder in DSM-5: implications for clinical practice and research of changes from DSM-IV. *Depress Anxiety, 31*, 459–471.

Vammen, M.A., Mikkelsen, S., Hansen, A.M., Grynderup, M.B., Andersen, J.H., Bonde J.P.,... Thomsen, J.F. (2014). Salivary cortisol and depression in public sector employees: Cross-sectional and short term follow-up findings. *Psychoneuroendocrinology, 41*, 63–74.

Velders, F.P., Kuningas, M., Kumari, M., Dekker, M.J., Uitterlinden, A.G., Kirschbaum, C., et al. Tiemeier, H. (2011). Genetics of cortisol secretion and depressive symptoms: a candidate gene and genome wide association approach. *Psychoneuroendocrinology, 36(7)*, 1053–61.

Verdu, E.F., Hayes, C.L., & O'Mahony, S.M. (2016). Importance of the Microbiota in Early Life and Influence on Future Health. In N. Hyland, & C. Stanton (Eds.), *The Gut Brain Axis. Dietary, Probiotic, and Prebiotic Interventions on the Microbiotica* (p.170). London: Elsevier.

Vinberg, M., Miskowiak, K., Kessing, L.V. (2014). Serotonin transporter genotype, salivary cortisol, neuroticism and life events: impact on subsequent psychopathology in healthy

twins at high and low risk for affective disorder. *Prog Neuropsychopharmacol Biol Psychiatry, 48*, 193–198.

Vining, R.F., McGinley, R.A., Maksvytis, J.J., Ho, K.Y. (1983). Salivary cortisol: a better measure of adrenal cortical function than serum cortisol. *Ann. Clin. Biochem., 20 (6)*, 329-335.

Wang, X., Wu, H., Miller, A.H. (2004). Interleukin-1 alpha-induced activation of p38 mitogen-activated kinase inhibits glucocorticoid receptor function. *Mol Psychiatry*, *9*, 65–75.

Will, H., Grabenstedt, Y.; Völkl, G., Banck, G. (2008). Depression Psychodynamik und Therapie [Depression psychodynamics and therapy] (9th ed., pp.47-49). In W. Mertens (Ed.), *Psychoanalytische Krankheitslehre*. Stuttgart: Kohlhammer.

Williams, V.P., Bishop-Fitzpatrick, L., Lane, J.D., Gwyther, L.P., Ballard, E.L., Vendittelli, A.P., Williams, R.B. (2010). Video-based coping skills to reduce health risk and improve psychological and physical well-being in Alzheimer's disease family caregivers. *Psychosom Med., 72*, 897-904.

Wittchen, H.-U., Zaudig, M., Fydrich, T. (1997). *Strukturiertes Klinisches Interview für DSM-IV* [Structured Clinical Interview for DSM-IV]. Göttingen: Hogrefe.

Wittchen H.-U., Hoyer J. (2006). *Klinische Psychologie und Psychotherapie* [clinical psychology and psychotherapy] (2th ed., pp.732-760). Heidelberg: Springer.

Wittchen, H.U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol, 21*, 655–679.

World Helath Organization (1993). The International Classification of Diseases-10th Revision. Classification of mental and behavioral disorders. Geneva: World Health Organization.

World Health Organization (2017). Depression and other common mental disorders: global health estimates. Geneva: World Health Organization.

Wright, C.E., Steptoe, A. (2005). Subjective socioeconomic position, gender and cortisol responses to waking in an elderly population. *Psychoneuroendocrinology, 30*, 582–590.

Yehuda, R. (2002). Current status of cortisol findings in post-traumatic stress disorder. Psychiatr. *Clin. North Am., 25*, 341–368.

Yoshihara, K., Hiramoto, T., Oka, T., Kubo, C., Sudo, N. (2014). Effect of 12 weeks of yoga training on the somatization, psychological symptoms, and stress-related biomarkers of healthy women. *Biopsychosoc Med, 8*, 1-9.

Young, E.A., & Altemus, M. (2004). Puberty, ovarian steroids, and stress. *Annals of the New York Academy of Sciences, 1021*, 124-133.

Young, E.A., Haskett, R.F., Grunhaus, L., Pande, A., Weinberg, V.M., Waston, S.J., Akil, H. (1994). Increased evening activation of the hypothalamic-pituitary-adrenal axis in depressed patients. *Archives of General Psychiatry 51*, 701–707.

Young, E.A. (1995). Glucocorticoid cascade hypothesis revisited: role of gonadal steroids. *Depression, 3*, 20–27.

Zeros-Kopp J. (2007). *Anatomie, Biologie und Physiologie* [anatomy, biology and physiology] (p.426). New York: Thieme.

Zung, W.W., Richards, C.B., Short, M.J. (1965). Self-rating depression scale in an outpatient clinic. Further validation of the SDS. *Arch Gen Psychiat, 13*, 508-515.

The protocol for this study is published under prospero: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=23286.

Index of Appendices

Appendix A. Detailed search syntax, PubMed, PsychInfo, Medline, Embase.

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)

Appendix B. Newcastle - Ottawa Quality Assessment Scale Cohort studies.

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort
   a) truly representative of the average _____ (describe) in the community ★
   b) somewhat representative of the average _____ in the community ★
   c) selected group of users eg nurses, volunteers
   d) no description of the derivation of the cohort
2) Selection of the non exposed cohort
   a) drawn from the same community as the exposed cohort ★
   b) drawn from a different source
   c) no description of the derivation of the non exposed cohort
3) Ascertainment of exposure
   a) secure record (eg surgical records) ★
   b) structured interview ★
   c) written self report
   d) no description
4) Demonstration that outcome of interest was not present at start of study
   a) yes ★
   b) no

Comparability

1) Comparability of cohorts on the basis of the design or analysis
   a) study controls for _____ (select the most important factor) ★
   b) study controls for any additional factor ★ (This criteria could be modified to indicate specific control for a second important factor.)

Outcome

1) Assessment of outcome
   a) independent blind assessment ★
   b) record linkage ★
   c) self report
   d) no description
2) Was follow-up long enough for outcomes to occur
   a) yes (select an adequate follow up period for outcome of interest) ★
   b) no
3) Adequacy of follow up of cohorts
   a) complete follow up - all subjects accounted for ★
   b) subjects lost to follow up unlikely to introduce bias - small number lost - > ____ % (select an adequate %) follow up, or description provided of those lost) ★
   c) follow up rate < ____% (select an adequate %) and no description of those lost
   d) no statement

Appendix C. Jadad Score Calculation.

Item	Score
Was the study described as randomized (this includes such words as "randomly," "random," and "randomization")?	0/1
Was the method used to generate the sequence of randomization described and was it appropriate (e.g., table of random numbers, computer-generated)?	0/1
Was the study described as double-blind?	0/1
Was the method of double-blinding described and was it appropriate (e.g., identical placebo, active placebo, dummy)?	0/1
Was there a description of withdrawals and dropouts?	0/1
Deduct 1 point if the method used to generate the sequence of randomization was described but was inappropriate (e.g., patients were allocated alternately or according to date of birth or hospital number).	0/−1
Deduct 1 point if the study was described as double-blind but the method of blinding was inappropriate (e.g., comparison of tablet vs. injection with no double dummy).	0/−1

Eidesstattliche Versicherung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Thema **„Are aberrant cortisol levels prognostic factors for the development of depression in the adult population? A systematic review“** selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 22.06.2019

_____

Daniela Stotz