

## Journal Pre-proofs

Child maltreatment and hypothalamic-pituitary-adrenal axis functioning: A systematic review and meta-analysis

Selina Schär, Ines Mürner-Lavanchy, Stefanie J. Schmidt, Julian Koenig, Michael Kaess

PII: S0091-3022(22)00010-3  
DOI: <https://doi.org/10.1016/j.yfrne.2022.100987>  
Reference: YFRNE 100987

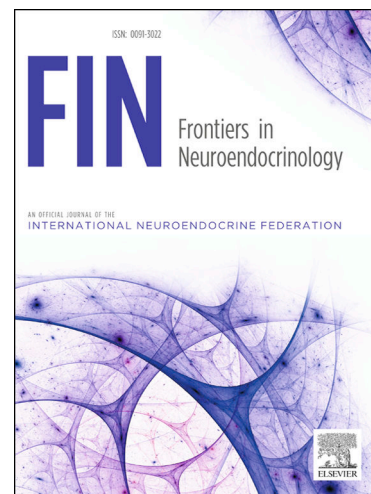
To appear in: *Frontiers in Neuroendocrinology*

Received Date: 17 October 2021  
Revised Date: 11 February 2022  
Accepted Date: 17 February 2022

Please cite this article as: S. Schär, I. Mürner-Lavanchy, S.J. Schmidt, J. Koenig, M. Kaess, Child maltreatment and hypothalamic-pituitary-adrenal axis functioning: A systematic review and meta-analysis, *Frontiers in Neuroendocrinology* (2022), doi: <https://doi.org/10.1016/j.yfrne.2022.100987>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier Inc.



# Child maltreatment and hypothalamic-pituitary-adrenal axis functioning:

## A systematic review and meta-analysis

Selina Schär<sup>a</sup>, Ines Mürner-Lavanchy<sup>a</sup>, Stefanie J. Schmidt<sup>b</sup>, Julian Koenig<sup>a,c,d,\*</sup>, Michael Kaess<sup>a,e,\*,#</sup>

\*: shared last-authorship

<sup>a</sup> *University Hospital of Child and Adolescent Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland*

<sup>b</sup> *Department of Clinical Psychology and Psychotherapy, University of Bern, Bern, Switzerland*

<sup>c</sup> *University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Cologne, Germany*

<sup>d</sup> *Section for Experimental Child and Adolescent Psychiatry, Department of Child and Adolescent Psychiatry, Centre for Psychosocial Medicine, Heidelberg University, Heidelberg, Germany*

<sup>e</sup> *Clinic for Child and Adolescent Psychiatry, Centre for Psychosocial Medicine, Heidelberg University, Heidelberg, Germany*

# Corresponding author at: University Hospital of Child and Adolescent Psychiatry and Psychotherapy, Bolligenstrasse 111, 3000, Bern 60, CH, Switzerland.

*E-mail address:* Michael.Kaess@upd.ch (M. Kaess).

**ABSTRACT**

Alterations in hypothalamic-pituitary-adrenal (HPA) axis and its effector hormone cortisol have been proposed as one possible mechanism linking child maltreatment experiences to health disparities. In this series of meta-analyses, we aimed to quantify the existing evidence on the effect of child maltreatment on various measures of HPA axis activity. The systematic literature search yielded 1,858 records, of which 87 studies ( $k = 132$ ) were included. Using random-effects models, we found evidence for blunted cortisol stress reactivity in individuals exposed to child maltreatment. In contrast, no overall differences were found in any of the other HPA axis activity measures (including measures of daily activity, cortisol assessed in the context of pharmacological challenges and cumulative measures of cortisol secretion). The impact of several moderators (e.g., sex, psychopathology, study quality), the role of methodological shortcomings of existing studies, as well as potential directions for future research are discussed.

**Keywords:** child maltreatment, hypothalamic-pituitary-adrenal axis, diurnal cortisol, cortisol awakening response, stress reactivity, dexamethasone suppression test, combined dexamethasone-corticotropin releasing hormone test, corticotropin-releasing hormone test, hair cortisol, urinary free cortisol, meta-analysis, systematic review

## 1. Introduction

Child maltreatment is a widespread phenomenon that affects the lives of millions of children worldwide (Stoltenborgh et al., 2015; Witt et al., 2017). Despite extensive research on the consequences of the experience of child maltreatment, surprising heterogeneity exists across studies with respect to its operational definition (Cicchetti and Toth, 2005; Leeb et al., 2008; Manly, 2005). Researchers, however, generally agree that child maltreatment involves both acts of commission - including physical, sexual, and emotional (psychological) abuse - as well as acts of omission (i.e., any form of neglect) by a parent or other caregiver that results in harm, potential for harm, or threat of harm to a child (usually interpreted as up to 18 years of age; Gilbert et al., 2009; Leeb et al., 2008). Typically, maltreated children experience multi-type maltreatment, suggesting that the different forms of maltreatment often co-occur (Herrenkohl and Herrenkohl, 2009; Vachon et al., 2015). While inconsistencies exist between studies in terms of its definition, extensive research including findings from prospective and retrospective cohort studies (e.g., Clark et al., 2010; Danese et al., 2009; Dube et al., 2001) as well twin studies (Kendler et al., 2000; Nelson et al., 2002), has shown that the experience of child maltreatment represents a profound, non-specific risk factor for the development of a broad variety of mental (e.g., Bonoldi et al., 2013; Carr et al., 2013; Dube et al., 2001; Infurna et al., 2016; Liu et al., 2018; Norman et al., 2012; Porter et al., 2020; Teicher and Samson, 2013) as well as physical disorders (e.g., Clemens et al., 2018; Hemmingsson et al., 2014). Importantly, most publications indicate a dose-dependent relationship between the experience of child maltreatment and the risk for health impairments, with those reporting more severe experiences or an increasing number of different types of child maltreatment, showing stronger associations (Clemens et al., 2018; Dube et al., 2001; Hemmingsson et al., 2014; Kendler et al., 2000; Norman et al., 2012).

One proposed mechanism by which child maltreatment might affect later disease risk involves epigenetic programming - a mechanism known to cause long-lasting changes in gene expression, especially in combination with specific risk alleles, thereby inducing long-lasting changes in biological functioning (i.e., biological embedding; Heim et al., 2019; Jacoby et al., 2016; Smith and Mill, 2011). It is assumed that particularly the expression of several genes relevant to stress regulation might be affected, ultimately leading to the development of a phenotype with core dysfunctions in circuits of the brain involved in the processing of stress and emotion regulation, and related changes in core outflow stress response systems, including - and this will present the focus of the current meta-analysis - the hypothalamic-pituitary-adrenals (HPA) axis (Chen and Baram, 2016; Heim et al., 2019; Jacoby et al., 2016; Koss and Gunnar, 2018; Strüber et al., 2014). In turn, these core dysfunctions might then increase the lifelong risk for the development of a wide range of adverse outcomes later in life (e.g., Heim et al., 2019).

As an important stress response system, HPA axis activity and associated cortisol secretion - i.e., *cortisol stress reactivity* - serves various functions, depending on the system under investigation (e.g., cardiovascular system, immune system), together being important for survival and restoring homeostasis (adapting to a homeostatic challenge; Nicolaides et al., 2015). Importantly, structures centrally involved in controlling stress-induced HPA axis activity (particularly anticipatory responses to social challenges) include limbic brain regions such as the hippocampus, the amygdala, and the prefrontal cortex (for a detailed review see Herman et al., 2003), structures also well-known to be affected by chronic stress such as the experience of child maltreatment (e.g., McEwen et al., 2016, 2015; Shirazi et al., 2015; Teicher et al., 2016). Interestingly, an aberrant cortisol stress response has been observed in patients with various mental health problems including borderline personality disorder (BPD), anxiety disorders, depression, and schizophrenia (e.g., Drews et al., 2019; Zorn et al., 2017).

Aside from its central role as a stress response system, HPA axis activity and associated cortisol secretion, in addition to stressors, is also triggered by other regulatory control factors, including circadian signals (for a comprehensive overview see Spencer and Deak, 2017). Accordingly, in addition to the well-known cortisol stress response, several other measures of HPA axis activity can be distinguished. These include the *cortisol awakening response (CAR)*; Clow et al., 2010; Fries et al., 2009; Spencer et al., 2018), *diurnal cortisol (DC)*; Adam and Kumari, 2009; Segerstrom et al., 2014; Spiga et al., 2014), cortisol assessed following pharmacological challenge tests (i.e., the *dexamethasone suppression test (DST)*, the *combined dexamethasone-corticotropin releasing hormone test (Dex-CRH)*, and the *corticotropin-releasing hormone test (CRH)*; Carroll, 1981; Gold et al., 1986; Watson et al., 2006), as well as cumulative measures of cortisol secretion including *24-hour urinary free cortisol (24-hour UFC)*; Deutschbein et al., 2011; Moore et al., 1985) and *hair cortisol concentrations (HCC)*; for an overview see Meyer and Novak, 2012; Stalder and Kirschbaum, 2012). Interestingly, similar to findings related to the cortisol stress response, alterations in these other HPA axis activity measures have been associated with various mental and physical health problems as well, although substantial inconsistencies exist among findings (e.g., Adam et al., 2017; Berger et al., 2016; Chida and Steptoe, 2009; Leistner and Menke, 2018; Stalder and Kirschbaum, 2012).

To summarize, a growing number of studies indicate associations between the experience of child maltreatment and epigenetic changes in key genes involved in stress regulation. Interestingly, epigenetic changes have been found in genes important for the regulation of HPA axis activity, such as in the GR gene, the FKBP5 gene and the CRH gene (Hoffmann and Spengler, 2012; Klengel et al., 2013; Palma-Gudiel et al., 2015; Turecki and Meaney, 2016). Together with findings of altered brain morphology, functioning and connectivity, especially in brain regions involved in the regulation of HPA axis activity (e.g., McCrory et al., 2010; Teicher et al., 2016), long-lasting changes in the regulation of this system in individuals exposed to child maltreatment are suggested. If additionally, the manifold biological effects of

cortisol are considered (e.g., Sapolsky et al., 2000), a dysregulation in the release of this hormone might increase the susceptibility for a wide variety of negative health outcomes, particularly in combination with other stressful experiences later in life. As discussed above, alterations in several measures of HPA axis activity have indeed been linked to various health conditions including psychopathology. Accordingly, the association between child maltreatment and alterations in cortisol secretion has been widely investigated, including various measures of HPA axis activity. However, findings are inconclusive, with inconsistencies reported (similarly with findings related to psychopathology), specifically regarding the direction of alteration (hyper- versus hyosecretion). To date, three meta-analyses have investigated the association between adverse childhood experiences and aberrant cortisol secretion (Bernard et al., 2017; Bunea et al., 2017; Fogelman and Canli, 2018), whereby two of them focused on early life adversity (ELA) in general and one specifically on child maltreatment. With respect to cortisol stress reactivity, the meta-analysis conducted by Bunea and colleagues (2017), revealed a blunted cortisol stress response to social stressors in those with a history of ELA. In contrast to these findings, the two other meta-analyses failed to show alterations in circadian rhythms including measures such as the diurnal slope (DSL) as well as the CAR (Bernard et al., 2017; Fogelman and Canli, 2018).

When examining the relationship between child maltreatment and HPA axis activity, careful control of several potential moderators may be of particular importance. As mentioned previously, various mental disorders have been linked to alterations in HPA axis functioning (e.g., Chida and Steptoe, 2009; Stalder et al., 2017; Zorn et al., 2017). The meta-analysis conducted by Zorn et al. (2017), for instance, was able to show that women with major depressive disorder, women with an anxiety disorder and patients with schizophrenia show a blunted cortisol stress response to psychosocial stressors compared to healthy control participants. The meta-analysis conducted by Drews et al. (2019) similarly found evidence for an overall attenuated cortisol stress response in patients with borderline personality disorder.

Due to the close association between the experience of child maltreatment and psychopathology, these studies, however, cannot rule out the possibility that the observed endocrinological alterations were actually caused by childhood adversity and may have been present even before the development of the respective mental health conditions. Moreover, beyond the epigenetic changes found in genes important for the regulation of HPA axis activity that are likely caused by the experience of chronic stress such as the experience of child maltreatment, specific polymorphisms of respective genes associated with differences in cortisol secretion independent of child maltreatment have also been found to be related to an increased risk of developing mental disorders (Fan et al., 2021; Ising et al., 2008; Mahon et al., 2013). Accordingly, psychopathology may interfere with or moderate the relationship between child maltreatment and cortisol, and should therefore be considered carefully. In addition, several lines of evidence suggest that women are generally more prone to mental disorders, particularly stress-related disorders, and that sex hormones might account for some of these findings (Li and Graham, 2017). Interestingly, marked sex differences have also been reported with respect to HPA axis functioning (Foley and Kirschbaum, 2010; Stalder et al., 2016; Stalder and Kirschbaum, 2012; Zänkert et al., 2019), indicating that in order to accurately capture the relationship between child maltreatment and cortisol, the influence of sex should be considered as well. Furthermore, there is some support that hormonal activity is elevated at stressor onset (e.g., at the time of child maltreatment) and reduces with passing time (e.g., with increasing age; Miller et al., 2007), implying that the age of the population studied may likewise be important. Finally, ethnic differences in HPA axis functioning have been found (Boileau et al., 2019), keeping in mind, of course, that minorities are generally exposed to more adversity (e.g., O'Connor et al., 2020), which in turn may explain this association.

A number of variables related to the measurement of child maltreatment may also be important when investigating the association between child maltreatment and HPA axis functioning. These include the assessment modality employed (e.g., informant versus self-reports), the



various approaches to defining the presence of child maltreatment (e.g., records, cut-offs, specifications), the age at maltreatment onset, as well as the chronicity of the maltreatment experiences. For instance, a nationally-representative birth cohort study, the Environmental Risk (E-Risk) Longitudinal Twin Study, demonstrated that retrospective self-report data of child maltreatment were more strongly associated with adult psychopathology than prospective informant reports (Newbury et al., 2018). Chronic maltreatment starting early in life is generally associated with poorer neurocognitive functioning and worse psychopathology (Cowell et al., 2015; Jaffee and Maikovich-Fong, 2011; Kaplow and Widom, 2007). In addition, findings suggest that chronic exposure to stress hormones can impact brain structures differently, depending on the timing and duration of the exposure (Lupien et al., 2009). Thus, effects on HPA axis functioning might vary depending on the age of first child maltreatment experience and/or the chronicity of these experiences.

Furthermore, studies vary widely regarding the assessment of cortisol. Depending on the HPA axis activity measure of interest, some of these variations may account for additional variability in the child maltreatment cortisol relationship. Variables that have been associated with different cortisol findings include, for instance, sample type (blood versus saliva; Spencer and Deak, 2017), slope type and whether morning samples were collected at awakening for diurnal cortisol (Adam et al., 2017; Ryan et al., 2016), as well as the type of stressor in the context of the cortisol stress reactivity (social-evaluative versus other; e.g., Dickerson and Kemeny, 2004; Zänkert et al., 2019), and differences might emerge depending on how well a task can elicit a cortisol response. In addition, the variation in dosage of dexamethasone in pharmacological stimulation tests (0.5 mg versus 1.0 mg) might account for variability as well (Leistner and Menke, 2018).

Finally, several aspects of methodological quality need to be considered when attempting to quantify the relationship between the experience of child maltreatment and HPA axis functioning. Besides matching the child maltreatment and the control group with respect to age,

sex and psychopathology, and ensuring that no one from the control group was exposed to child maltreatment experiences, these include instructions for sampling and collection, the day and timing of sampling as well as controlling for specific state covariates such as being sick or experiencing any current stress at the time of testing, with these methodological variables differing to some extent between the various HPA axis activity measures. For a comprehensive overview and corresponding references see Appendix B Tables B1-B7. Finally, there are certain disease states (e.g., addictions, endocrine diseases), various drugs (e.g., glucocorticoids, psychoactive medications) and sex hormone-dependent variables (e.g., menstrual cycle, oral contraceptive intake, pregnancy) that can strongly influence cortisol levels (e.g., Foley and Kirschbaum, 2010; Granger et al., 2009; Kudielka et al., 2012; Kudielka and Wüst, 2010; Locatelli et al., 2009; Stalder et al., 2016; Zänkert et al., 2019). Considering that participants with experiences of child maltreatment are more likely to suffer from medical conditions, are at increased risk for substance abuse and more often experience unintended teenage pregnancy, (e.g., Hughes et al., 2017), factors related to altered cortisol secretion (e.g. Foley and Kirschbaum, 2010; Stalder et al., 2016; Stalder and Kirschbaum, 2012), careful assessment and matching between groups (i.e., maltreated versus control group) on these variables is of major importance.

Thus, in this comprehensive systematic review and meta-analysis, we aimed to quantify existing evidence on the effect of child maltreatment on cortisol metabolism including all of the previously mentioned measures of HPA axis activity. In contrast to existing meta-analyses, we were particularly interested in the potential influence of psychopathology in interfering or moderating the effect of child maltreatment on changes in cortisol secretion. Accordingly, psychopathology, and especially, the matching of the groups (psychopathology versus no psychopathology) was recorded thoroughly. Furthermore, we were interested in a range of other factors likely to moderate the effect of child maltreatment on cortisol regulation such as age at the time of study participation, sex, ethnicity, child maltreatment assessment method (informant

versus self-report), child maltreatment grouping method (i.e., cut-off scores, child protective services (CPS) records, other), age at the time of child maltreatment onset, chronicity of the child maltreatment experiences, as well as variables related to the assessment of the corresponding HPA axis activity measure. We considered different indices for each HPA axis activity measures, at least including one index of total cortisol production and one index reflecting change in cortisol over time (sensitivity of the system; Pruessner et al., 2003). In contrast to the existing meta-analyses, the present investigation sought to determine whether aberrant cortisol secretion patterns following child maltreatment can be observed in both of these largely independent components of HPA axis activity. Finally, a comprehensive quality assessment based on expert guidelines was developed for each activity measure, and several elements of methodological quality and their potential moderating role were investigated.

## 2.0 Methods

### 2.1 Systematic literature search

Articles were identified by searching the electronic databases Pubmed, PsycINFO, Web of Science (WOS) and the Cumulative Index to Nursing and Allied Health Literature (CINAHL) from their inception to June 2018. The search consisted of titles and abstracts and used the following search string: (“maltreatment” OR “neglect” OR “emotional abuse” OR “sexual abuse” OR “physical abuse” OR “childhood trauma”) AND (“hypothalamic-pituitary-adrenal” OR “HPA” OR “cortisol” OR “adrenocorticotrophic hormone” OR “ACTH”). Moreover, the reference lists of the prior meta-analyses on HPA axis functioning and child maltreatment (or ELA) were checked, as were studies proposed by authors who were contacted in the context of data collection.

### 2.2 Selection criteria

Studies were included if they: (1) involved human participants of all ages; (2) reported on at least one measure of child maltreatment (e.g., self-report [questionnaire, interview], report by

an outside source [CPS record, parental report] or self-identification), whereby child maltreatment was defined according to the definition provided by the Centers for Disease Control and Prevention (*“any act or series of acts of commission or omission by a parent or other caregiver that results in harm, potential for harm, or threat of harm to a child”*; Leeb et al., 2008, p. 11) including the subtypes emotional, physical, sexual abuse, as well as neglect; (3) reported measuring cortisol levels, either as indicator of daily activity (DC, CAR), in response to a stressor (cortisol stress reactivity) or pharmacological challenges (DST, Dex-CRH, CRH) or as a cumulative index (24-hour UFC or HCC). All assessment methods, i.e., saliva, blood (serum, plasma), urine and hair, were eligible. Additionally, several preconditions were formulated for the various measures of HPA axis activity: With respect to DC at least two sampling time points, one cortisol assessment in the morning (best with reference to awakening) and one in late afternoon/evening, had to be available (Segerstrom et al., 2014). In case of the CAR, only studies that collected at least two samples, with the first sample anchored to awakening and a second sample between +30 min or +45 min post-awakening, were included (Stalder et al., 2016). With regard to the cortisol stress reactivity, studies needed to collect at least one cortisol baseline measure (before being introduced to a stressor) as well as a sample between 20-40 min post-stressor onset to capture the peak of the cortisol stress response (Dickerson and Kemeny, 2004). UFC had to be collected over a period of 24 hours (Moore et al., 1985) and studies assessing HCC needed to focus on the first 3cm hair segment (Meyer and Novak, 2012). No restrictions were applied to studies involving pharmacological stimulation tests. Articles were excluded for one or more of the following reasons: (1) evaluated a non-human sample; (2) were not written in English or German; (3) did not contain primary data (e.g., systematic reviews, meta-analysis, book chapters); or (4) were not peer reviewed (e.g., dissertations, master-theses, conference abstracts). Additionally, studies that included (5) participants with substance abuse (e.g., alcohol, cocaine); (6) individuals who suffered from a medical condition (e.g., endocrine disorder, chronic fatigue, chronic pain); or studies including

(7) pregnant women or women in the postpartum period (up to 6 months) were excluded, given the well-known effects of these factors on HPA axis activity (e.g., Stalder et al., 2016; Zänkert et al., 2019). Studies were first screened based on their title and abstract and then further examined in full text in case of suitability.

### 2.3 Data extraction

As mentioned earlier, for each of the HPA axis activity measures (exception: 24-hour UFC and HCC) several outcome indices were defined including at least one measure for total cortisol production and one measure reflecting changes in cortisol over time. For DC these included waking (morning) and bedtime levels (for total cortisol) as well as the delta between bedtime and waking cortisol samples (reflecting change over time; DSL). For the CAR, the following assessment time points or indices were extracted: waking cortisol, peak cortisol (expected between 30 to 45 min post awakening), end cortisol (assessed 60 min post awakening), peak reactivity (delta between the peak and the waking cortisol sample) as well as the area under the curve with respect to the ground ( $AUC_g$ ; reflecting total cortisol production) and the area under the curve with respect to increase ( $AUC_i$ ; reflecting changes over time; for more details regarding these two formulas see Pruessner et al., 2003). Similarly, the following assessment time points or indices were extracted for the cortisol stress reactivity, the CRH- and the Dex-CRH-test: baseline cortisol (before being introduced to a stressor; before CRH injection), peak levels, recovery levels (last sample assessed), peak reactivity (defined as delta between peak and baseline levels),  $AUC_g$  and  $AUC_i$ . In some of the studies, the timing of the peak differed between the child maltreatment and the control group and the values at the individual peak times were extracted for each group. Whenever a peak occurred in only one of the groups, the value at that peak time was extracted also for the other group. There were also studies in which neither group showed a cortisol response following the perception of a stressor. In this case the values were extracted at the time a response would have been expected (around 30 min post stressor onset). For these few studies, however, no peak reactivity values were extracted or

included in this meta-analysis. For the DST, cortisol assessed in the morning prior administration of dexamethasone (which was normally administered at 11pm) and cortisol measured the next day, as well as the delta between these two measurement time points were extracted. As we were interested in obtaining all of the defined outcome indices from each study, the authors of all studies containing missing information were contacted with a data request. The data request consisted of an excel file containing all the variables of interest. Since we assumed that the requested additional calculations were relatively time-consuming, one option was to send us the raw data, with which we calculated the desired indices. If the data could not be provided, whenever possible, means and standard deviations were extracted from tables or text. If those data were not available, data were extracted from figures using a web-based digitizer (Rohatgi, 2012). Two independent reviewers extracted the data from the corresponding figures and the mean value of both extractions was calculated. If not clearly stated in the text or in the subheading of the figures, we assumed that they represented means and standard errors. Standard deviations were calculated from standard errors or from confidence intervals using the RevMan Calculator provided by the Cochrane group (Drahotka and Beller, n.d.). If none of these data sources were available, the study was excluded. In case of multiple publications based on the same cohort, the study with the largest sample size or the one that provided extractable data was included. Whenever possible, non-transformed (raw) cortisol data were extracted. In the case of more than two groups described in the paper, we extracted the data from those two groups that best matched in terms of psychopathology, or if the data of two groups could be combined, weighted means and standard deviations were calculated using the StatsToDo software (<https://www.statstodo.com/index.php>). Whenever a study included both a clinical and a healthy control group and appropriate measures of child maltreatment were taken in both groups, data were requested or extracted separately for these two subgroups. If the experience of child maltreatment was assessed, but grouping was based on other criteria, the authors were asked to (re)group participants based on the presence or

absence of child maltreatment, or in case of cut-off scores in high versus low child maltreatment groups (see Appendix A for details on the respective (re)grouping method of the studies in question). Finally, if a measure did not just assess child maltreatment but other traumatic experiences as well, such as it is the case for the Early Trauma Inventory (ETI; Bremner et al., 2000), authors were requested to group participants including only the subscales which assessed child maltreatment. However, this was not always possible, which is why some of the included studies did not focus on child maltreatment only, but on ELA in general. The few studies to which this applies are marked accordingly. We always asked authors to provide the data including only those without missing cortisol or child maltreatment assessments. Therefore, the data presented in this meta-analysis might not completely correspond to the data displayed in the original studies.

#### 2.4 Coding of study characteristics for moderator analyses

The following details were extracted from each study (1) identifying features (i.e., authors, year of publication, journal), (2) participant characteristics (i.e., sample size, age-range, average age, sex ratio, ethnicity [or race: percentage of Caucasians, non-Caucasians], assessment of psychopathology [clinical sample, healthy controls, mixed, not assessed, as well as the percentage of participants meeting criteria for a current mental disorder]), (3) trauma related information (i.e., measurement method [self-report, informant report, mixed], instrument used, grouping method [cut-off scores, record, other], type of child maltreatment [emotional, physical, sexual abuse, neglect], average age of first child maltreatment report and average duration of child maltreatment), (4) cortisol related information (i.e., type of sample [blood, saliva, urine, hair], measurement unit [e.g., nmol/l, µg/dl], time points of sampling, number of samples, reliability of measure [sampling over one day, two-days, more], minutes to peak, duration of stressor, type of stressor [social-evaluative, other], whether a cortisol response was observed [yes, both groups, only in one group, no], dose of the respective stimulant in case of the pharmacological stimulation tests) and (5) data related information (source of data [paper,

provided] and whether the data were (re)grouped or not). Some of these variables were viewed as potential moderators that might account for variability in the child maltreatment cortisol relationship (see moderator analyses).

## 2.5 Risk of bias in individual studies

In order to quantify the risk of bias for each individual study and to examine the potential moderating role of several elements of methodological quality, a quality assessment tool was developed. This quality assessment tool covered the following three key domains: (1) variables associated with the selection of participants (including the measurement of child maltreatment and the matching of the two groups with respect to age, gender and psychopathology), (2) variables associated with the measurement of HPA axis activity and related to the (3) assessment of important confounders. These quality criteria, particularly those related to the assessment of cortisol and associated confounders, were developed based on expert guidelines and differ to some degree between the various HPA axis activity measures (see Appendix B for corresponding references). The risk of bias assessment for each HPA axis activity measure was conducted by two independent reviewers, with disagreements being resolved through discussion. In case of (re)grouped data or missing statistics (e.g., t-test or fisher's exact test), corresponding group comparisons were conducted based on available means and standard deviations using QuickCals from GraphPad (<https://www.graphpad.com/quickcalcs/>). As the data of some articles were (re)grouped, the information with respect to some quality items was no longer available at the group level. In this case, a conservative approach was followed and the corresponding point was not awarded (marked accordingly in the corresponding tables of Appendix B). In certain cases, the assessment of a quality item was not meaningful, e.g., scoring the matching between the child maltreatment and the control group with respect to oral contraceptive intake in an all-male sample, and in those cases, corresponding items were coded as NA (not applicable). For each of the three quality domains (selection of participants, appropriate assessment of the corresponding HPA axis activity measure, appropriate control for



confounders) a score derived from the mean of all associated items (excluding the NA items) multiplied by 100, was calculated. In addition to these domain-specific scores, we also calculated an overall total score. These scores were then used in corresponding meta-regression analyses.

## 2.6 Statistical analyses

All analyses were run using R and R studio (version 3.6.2 (2019-12-12), packages: meta, metafor, dmetar) and were guided by the online book “A Hands-on Guide” from Harrer et al., (2019). Effect sizes for the primary studies were estimated using the Hedge’s  $g$  coefficient, corrected for small sample sizes (Hedges, 1982). In order to calculate the overall effect, random-effects models for the different HPA axis activity measures and the various outcome indices were performed, applying the Restricted Maximum-Likelihood (REML) method to estimate the variance of the distribution of the true effect sizes ( $\tau^2$ ; Veroniki et al., 2016). Between-study heterogeneity was evaluated focusing on the Cochran’s  $Q$  statistics (with a  $p < 0.05$  indicating the presence of statistical heterogeneity), the Higgins’s and Thompson’s  $I^2$  measure (with  $I^2$ : 25% = low heterogeneity, 50% = moderate heterogeneity, 75% = high heterogeneity) and the prediction interval (Higgins, 2003; Higgins and Thompson, 2002; IntHout et al., 2016). By means of the `find.outliers` (meta package) and `inf.analysis` function (Leave-One-Out-method; dmetar package), studies with extreme effect sizes (outlier studies) and studies exerting a high impact on the overall result (potential influential studies) were identified and excluded in the context of corresponding sensitivity analyses (Harrer et al., 2019; Viechtbauer and Cheung, 2010). Additionally, meta-regression and subgroup analyses (mixed/fixed-effects model) were conducted to examine the influence of several predefined moderator variables. For some studies, cortisol data were available for a lower number of participants than reported in the original paper, with information on the various moderator variables only available for the original sample. Despite this, these original values were included in corresponding moderator analyses. To our best knowledge, we marked this in the

tables describing the characteristics of the included studies. In case of substantial between-study heterogeneity ( $I^2 > 50\%$ ), meta-regression and subgroup analyses were based on the sensitivity model excluding outlier studies. Finally, in order to evaluate the presence of publication bias, funnel plots were visually inspected and the Egger's test for funnel plot asymmetry was performed (Egger et al., 1997; Peters et al., 2008).

Journal Pre-proofs

### 3.0 Results

#### 3.1 Search results

The literature search yielded a total of  $N = 1,858$  records of which  $n = 575$  duplicates were removed. Screening of reference lists of existing meta-analyses on HPA axis functioning and child maltreatment (or ELA), as well as studies proposed by authors who were contacted in the context of data collection, yielded an additional  $n = 9$  studies. After title and abstract screening,  $n = 1,025$  articles were discarded because they did not meet inclusion criteria. The remaining  $n = 267$  studies were assessed in full-text. Of these, another  $n = 120$  publications were excluded for the following reasons (1) no appropriate HPA axis measure ( $n = 52$ ), (2) all participants experienced child maltreatment ( $n = 8$ ), (3) unusual measure of child maltreatment ( $n = 6$ ), (4) intervention study with no baseline assessment ( $n = 3$ ) and (5) samples used in multiple studies ( $n = 51$ ). Additionally,  $n = 60$  articles had to be excluded due to missing relevant statistics, leaving a total of  $n = 87$  independent studies included in this series of meta-analyses (for full process of study selection see Fig. 1). Of the  $n = 87$  studies,  $n = 14$  studies included two subgroups, one study contained three subgroups and  $n = 18$  articles collected data on more than one HPA axis activity measure (with DC and CAR most frequently jointly assessed), leaving a total of  $k = 132$  group comparisons. Since some studies collected data on various HPA axis activity measures, it was possible that an effect size (e.g., cortisol measured at awakening) was included in two different random-effects models relating to two different outcome indices (e.g., morning cortisol in the context of DC and awakening cortisol in the context of the CAR). With respect to the various HPA axis activity measures,  $n = 23$  studies reported on DC ( $k = 26$ ),  $n = 22$  on the CAR ( $k = 27$ ),  $n = 35$  on cortisol stress reactivity ( $k = 39$ ) and  $n = 19$  studies assessed cortisol following pharmacological challenges (DST:  $n = 11$  ( $k = 17$ ); Dex-CRH test:  $n = 8$  ( $k = 10$ )). Only two studies examined cortisol after the CRH test, which is why these two studies were combined with the data reported for the Dex-CRH test. With respect to the cumulative measures,  $n = 8$  studies reported on HCC ( $k = 9$ ) and  $n = 4$  studies on 24-hour UFC. Overall,

data from  $n = 41$  studies were provided by the respective authors, of which  $n = 23$  data sets ( $k = 29$ ) were (re)grouped for the purpose of this meta-analysis (see Appendix A). For three studies including large sample sizes (Hibel et al., 2019; Lovallo et al., 2019; Vreeburg et al., 2009) from which we obtained data, the publications that best described the respective samples and not those that appeared in the initial literature search were chosen as references.

< INSERT FIGURE 1 ABOUT HERE >

## 3.2 Synthesis of results

### 3.2.1 Diurnal cortisol

#### 3.2.1.1 Included studies

In total, our systematic search strategy identified  $n = 40$  studies that assessed waking (or morning) and evening cortisol to measure some aspects of the circadian rhythm of cortisol secretion. Of these,  $n = 23$  studies, including  $k = 26$  comparisons involving a total of  $n = 5,248$  participants were retained for quantitative synthesis. Owing to a lack of statistical information, data from the remaining  $n = 17$  studies that were eligible for inclusion could not be considered. The mean age of the total sample was 26.89 ( $SD = 14.99$ ) years, the majority of studies included predominantly female subjects (with the percentage of females ranging between 33.1-100.0%,  $M = 65.4%$ ,  $SD = 24.0%$ ;  $k = 5$  comparisons with a purely female sample) and the percentage of Non-Caucasians ranged between 0.0-81.7% ( $M = 30.0%$ ,  $SD = 30.7%$ ;  $k = 14$  not reporting on ethnicity). Most studies ( $k = 16$ ) were conducted with adults only, with fewer studies involving children or adolescents ( $k = 10$ ). With respect to psychopathology,  $k = 8$  comparisons included healthy participants,  $k = 5$  involved participants all fulfilling diagnostic criteria for a mental disorder,  $k = 6$  comparisons comprised participants with at least some fulfilling the diagnostic criteria for a mental disorder (range: 17.5-96.0%; with  $k = 2$  not matched in terms of psychopathology) and  $k = 7$  did not report on psychopathology at all. The majority of studies used self-reports to assess the presence of child maltreatment ( $k = 17$ ) and  $k = 9$  comparisons relied on informant reports. The assessment of child maltreatment and the grouping of

participants into a child maltreatment and a control group varied across the studies. This refers both to the instruments used as well as to the grouping procedure applied (e.g., cut-off scores, specific definitions, presence of records). Three studies ( $k = 4$ ) not only focused on child maltreatment but also included participants with other types of ELA (Carrion et al., 2002; Faravelli et al., 2010; Faravelli et al., 2017). In terms of reliability, the fewest studies assessed cortisol over more than two days ( $k = 5$ ). In total, we received data from 13 studies ( $k = 16$ ), of which the respective authors of six studies ( $k = 8$ ) regrouped or grouped their data based on the available assessment of child maltreatment (or, in case of raw data, the (re)grouping was performed by us). For further details on the characteristics of the included studies, see Table 1.

#### 3.2.1.2 Risk of bias assessment

Studies received an average total score of 51.9/100.0 ( $SD = 12.7$ , range: 27.6-76.7). With respect to the selection of participants ( $M = 61.0/100.0$ ,  $SD = 15.5$ ), the majority of studies used an established instrument to assess the experience of child maltreatment and matched their participants with respect to age, sex and psychopathology (assessed with a gold-standard diagnostic tool). However, less than half of the comparisons ( $k < 13$ ) assured that all participants in the child maltreatment group were exposed to child maltreatment, while none of the participants in the control group were, and only four studies ( $k = 5$ ) used two different sources to establish the presence of child maltreatment. Relating to the appropriate assessment of DC ( $M = 53.5/100.0$ ,  $SD = 17.4$ ), most studies did report on clear sampling instructions (including prohibitions of certain behaviors before sampling as well as clear information about how to collect, where to place and how to return samples) and provided details on their test protocol as well as on missing data and/or handling of outliers. However, only few studies provided clear instructions regarding the day of sampling ( $k = 9$ ), assured that the time of awakening did not differ between the groups ( $k = 7$ ), assessed sampling time adherence ( $k = 7$ ), rescheduled sampling if participants were sick ( $k = 5$ ), reported on batch analysis ( $k = 8$ ) and assured that participants were not under any current extraordinary stress ( $k = 2$ ). Finally, as shown by the

relatively low scores related to the control of confounding variables ( $M = 40.4/100.0$ ,  $SD = 24.6$ ), less than half of the comparisons ( $k \leq 13$ ) excluded participants with a medical condition or participants working night shifts, assessed smoking, menstrual cycle, oral contraceptive and medication use (especially medications affecting the central nervous system (CNS)) and thus assured that participants did not differ in these respects and only  $k = 2$  comparisons assured that participants did not differ with respect to other ELA or adult adversity. It should be noted, however, that several studies would have assessed some of the variables of interest, but since the data of six studies ( $k = 8$ ) were (re)grouped, the corresponding information at the group level was no longer available for all of these studies. For details on individual scoring results of the primary studies as well as a summary of the average risk of bias scores see Appendix B Table B8 or Table B1 for individual quality items.

#### 3.2.1.3 Meta-analysis

The results of the meta-analyses for the three indices of circadian activity (morning, evening, DSL) revealed no overall differences for morning (Hedges'  $g = -0.02$ , 95% CI [-0.11; 0.06],  $p = 0.586$ ) and DSL cortisol (Hedges'  $g = 0.00$ , 95% CI [-0.11; 0.11],  $p = 0.987$ ) between the child maltreatment and the control sample (this also held true for the corresponding sensitivity analyses; for further details see Table 2). In contrast, participants in the child maltreatment group had slightly elevated evening cortisol levels compared to their respective control group (Hedges'  $g = 0.10$ , 95% CI [0.03; 0.18],  $p = 0.008$ ). For corresponding forest plots see Appendix C Figs. 1.1 to 1.3. Between-study heterogeneity was in the low range for all outcome indices ( $I^2 < 30\%$ ,  $Q$ -statistics all  $p > 0.05$ ) and visual inspection of traditional and counter-enhanced funnel plots as well as Egger's regression test of funnel plot asymmetry implied absence of small-study bias (all  $p > 0.05$ ; for funnel plots see Appendix C Figs. 1.1 to 1.3).

#### 3.2.1.4 Meta-regression and subgroup analyses

We conducted a number of pre-defined meta-regression and subgroup analyses. The summary results for each outcome index and each moderator examined are shown in Appendix D Table

D1. In the following section, results for moderators found to significantly influence the main effects are outlined. Despite low heterogeneity in the effect size estimates between studies reporting on morning cortisol ( $I^2 = 29.1\%$ ), the following two continuous moderators influenced the main effect: (1) age at the time of study participation and (2) the sub-domain “appropriate measure of cortisol in the context of DC” of the quality assessment. With respect to the mean age of study participants, studies including older-aged samples reported a tendency for higher morning cortisol ( $\beta = 0.006$ , 95% CI [0.001; 0.010],  $p = 0.014$ ,  $R^2 = 73.04\%$ ; when comparing the child maltreatment group to the control group) compared to younger samples. Concerning the assessment of cortisol, studies with higher quality scores were associated with lower morning cortisol ( $\beta = -0.005$ , 95% CI [-0.010; -0.001],  $p = 0.012$ ,  $R^2 = 85.49\%$ ) in the child maltreatment group compared to the control group. In addition, forming subgroups of studies using informant reports and those relying on self-report data to assess the presence of child maltreatment revealed overall reduced morning cortisol in those studies applying informant reports (Hedge’s  $g = -0.114$ , CI [-0.204; -0.024]), whereas a tendency for increased morning cortisol in the child maltreatment compared to the control group was observed in studies relying on self-report information (Hedge’s  $g = 0.060$ , 95% CI [-0.026; 0.146];  $Q_1 = 7.48$ ,  $p = 0.006$ ). Since the majority of studies relying on informant reports used the presence of records to group participants, the corresponding subgroup comparison of the different grouping methods applied (records, cut-off scores, other grouping approaches (mainly specifications)) also reached significance ( $Q_2 = 8.32$ ,  $p = 0.016$ ). Finally, the subgroup comparison between studies where original data were extracted and those that (re)grouped their data for this meta-analysis also explained some of the between-study heterogeneity, with studies where original data could be extracted implying overall reduced morning cortisol (Hedge’s  $g = -0.096$ , 95% CI [-0.177; -0.016]) and those with (re)grouped data pointing to slightly increased morning cortisol levels in the child maltreatment group (Hedge’s  $g = 0.087$ , 95% CI [-0.014; 0.189];  $Q_1 = 7.72$ ,  $p = 0.005$ ). However, since the majority of studies ( $k = 8$  of  $k = 9$ ) that relied on informant reports

to group participants also belonged to the original data subgroup, interpretation of these findings should be done with caution. With respect to evening cortisol ( $I^2 = 2.0\%$ ), studies focusing on other types of ELA showed larger positive effect size estimates than studies focusing on child maltreatment only ( $Q_I = 12.24, p < 0.001$ ). Further, studies including original data showed larger positive effect size estimates than studies which provided (re)grouped data ( $Q_I = 6.47, p = 0.011$ ). It should be noted that by excluding those three studies ( $k = 4$ ) focusing not only on child maltreatment experiences (but also including participants with loss experiences), the initial model on evening cortisol became insignificant ( $p = 0.098$ ). Finally, with respect to DSL cortisol, no moderator was identified that significantly influenced between-study heterogeneity.

### 3.2.2 Cortisol awakening response

#### 3.2.2.1 Included studies

A total of  $n = 22$  studies, comprising  $k = 27$  comparisons, with an overall sample size of  $n = 3,545$  participants assessed cortisol in response to awakening. The mean age of the total sample was 27.36 ( $SD = 10.45$ ) years ( $k = 8$  comparison involved children/adolescents and  $k = 19$  including adults only) and the majority of studies included mainly female subjects (with the percentage of females ranging between 0.0-100.0%,  $M = 65.4\%$ ,  $SD = 30.0\%$ ;  $k = 7$  comprised a purely female sample). With respect to ethnicity, most studies included samples composed predominantly of an ethnic majority group with percentages of Non-Caucasians ranging between 0.0-100.0% ( $M = 45.7\%$ ,  $SD = 42.2\%$ ; with  $k = 20$  not reporting on ethnicity). Concerning psychopathology,  $k = 9$  comparisons included healthy participants,  $k = 6$  involved participants all fulfilling diagnostic criteria for a mental disorder,  $k = 6$  comparisons comprised participants with at least some fulfilling the diagnostic criteria for a mental disorder (with  $k = 4$  not matched in terms of psychopathology) and  $k = 6$  did not report on psychopathology at all. The majority of studies ( $n = 18, k = 23$ ) employed self-reports to assess the presence of child maltreatment and only  $k = 4$  relied on informant reports. Along with the use of different instruments, the grouping of participants into a child maltreatment and a control group varied,



however, with the majority of studies using specific cut-off scores ( $n = 13$ ,  $k = 17$ ). Only one of the included studies (Klaus et al., 2018) not only focused on child maltreatment but also included participants with other types of ELA including death of a close friend or relative, parental separation or divorce, major illnesses or injuries or other traumatic experiences. Three studies ( $k = 4$ ) assessed the CAR over more than two days, and there were several studies ( $n = 8$ ,  $k = 9$ ) with cortisol sampled at only two time points (i.e., awakening and 30 min post awakening or 45 min post awakening). In  $n = 4$  studies peak cortisol values were not observed at the same assessment time points for both groups. Finally, data from 12 studies ( $k = 15$ ) were provided by the respective authors, of which  $k = 10$  comparisons contained (re)grouped data. For further details on the characteristics of the included studies, see Table 3.

#### 3.2.2.2 Risk of bias assessment

Studies which assessed cortisol in response to awakening received an average total score of 52.8/100.0 ( $SD = 11.7$ , range: 35.7-76.7). With respect to the selection of participants ( $M = 65.7/100.0$ ,  $SD = 14.9$ ), the majority of comparisons ( $k > 13$ ) ensured that all participants in the child maltreatment group were exposed to maltreatment, while none of the participants in the control group was, used an established instrument to assess the experience of child maltreatment and matched their participants with respect to age, sex and psychopathology (assessed with a gold-standard diagnostic tool;  $k = 12$  in case of self-reports). However, only one study used two different sources to establish the presence of child maltreatment. Concerning the appropriate assessment of cortisol in the context of the CAR ( $M = 56.8/100.0$ ,  $SD = 12.0$ ), most studies reported on clear sampling ( $k = 22$ ) and collection ( $k = 27$ ) instructions, provided information on the day of sampling ( $k = 14$ ), collected at least three samples (with one sample between 30 and 45 min post awakening,  $k = 18$ ) over at least two days ( $k = 14$ ), provided information about how samples were collected, stored or analyzed ( $k = 27$ ) and reported on outliers or missing data ( $k = 22$ ). However, less than half of the comparisons ( $k < 14$ ) assessed the time of awakening (thus ensuring that the two groups did not differ in this respect;  $k = 13$ ), assessed

sampling time adherence ( $k = 8$ ), reported whether sampling was rescheduled if participants were sick ( $k = 7$ ), reported on batch analyses ( $k = 9$ ) and only  $k = 3$  comparisons ensured that participants were not under any current extraordinary stress, or whether sampling was rescheduled if participants experienced any stressor during the day of collection. Finally, many of the studies failed to control for several important confounding variables ( $M = 38.0/100.0$ ,  $SD = 22.5$ ). For instance, less than half of the comparisons ( $k < 14$ ) reported whether participants were excluded if pregnant or working night shifts, assessed smoking, menstrual cycle, oral contraceptive and medication use (especially medications affecting the CNS), thus ensuring that participants did not differ in these respects, and only  $k = 2$  comparisons ensured that participants did not differ with respect to other ELA or adult adversity. Again, several studies would have assessed some of the variables of interest, but since the data of eight studies ( $k = 10$ ) were (re)grouped, the corresponding information at the group level was no longer available for some of these studies. For details on individual scoring results of the primary studies as well as a summary of the average risk of bias scores, see Appendix B Table B9 or Table B2 for individual quality items.

### 3.2.2.3 Meta-analysis

The pooled effect estimates for the different CAR indices (including corresponding sensitivity analyses) are displayed in Table 4. As shown in the corresponding lines, none of the examined indices suggested a difference ( $p > 0.05$ ) in cortisol assessed in response to awakening when comparing the child maltreatment and the control group. For corresponding forest plots see Appendix C Figs. 2.1 to 2.6. Between-study heterogeneity was in the moderate to high range for some of the outcome indices ( $I^2 = 41.8\% - 73.1\%$ ), exceeding the level of significance ( $Q$ -statistics all  $p < 0.05$ ) for peak, delta,  $AUC_g$  and  $AUC_i$  cortisol, suggesting that other variables differing between the included studies might be of importance as well. Visual inspection of traditional and counter-enhanced funnel plots as well as Egger's regression test of funnel plot

asymmetry implied absence of small-study bias for all outcome indices examined (all  $p > 0.05$ ; for funnel plots see Appendix C Figs. 2.1 to 2.6).

#### 3.2.2.4 Meta-regression and subgroup analyses

The summary results for the pre-defined meta-regression and subgroup analyses for each outcome index and each moderator examined are shown in Appendix D Table D2. In the following, the results for moderators found to significantly influence the main effects are outlined. The subgroup comparison between studies where original data were extracted and those that (re)grouped their data for this meta-analysis explained some of the between-study heterogeneity for awakening cortisol and AUC<sub>g</sub> cortisol, with studies where original data could be extracted demonstrating overall reduced morning cortisol (Hedge's  $g = -0.169$ , 95% CI [-0.323; -0.015] and Hedge's  $g = -0.611$ , 95% CI [-0.868; -0.355], respectively) and those with (re)grouped data pointing to slightly increased morning cortisol levels in the child maltreatment group (Hedge's  $g = 0.054$ , 95% CI [-0.036; 0.145] and Hedge's  $g = 0.069$ , 95% CI [-0.025; 0.164], respectively;  $Q_I = 6.02$ ,  $p = 0.014$  and  $Q_I = 23.80$ ,  $p < 0.001$ , respectively). Since, as noted before, there is a relatively large overlap between studies reporting on morning cortisol assessed in the context of DC as well as on cortisol assessed in response to awakening ( $k = 14$ ), this finding was to be expected. Furthermore, with respect to awakening cortisol, age seems to explain some of the between-study variance ( $R^2 = 70.27\%$ ), but in contrast to morning cortisol (DC), does not represent a significant moderator ( $p = 0.338$ ). For delta cortisol, we identified the proportion of women in the sample as a significant continuous moderator (for peak cortisol:  $p = 0.071$ , for 60 min post awakening cortisol:  $p = 0.096$  and for AUC<sub>g</sub> cortisol:  $p = 0.058$ ), with an increase in the proportion of females being associated with lower cortisol when comparing the child maltreatment and the control sample ( $\beta = -0.005$ , 95% CI [-0.010; -0.000],  $p = 0.040$ ,  $R^2 = 0.0\%$ ). Finally, with respect to AUC<sub>g</sub> cortisol the sub-domain "appropriate measure of confounders" of the quality assessment explained some of the variance in the effect estimates, with studies with higher quality scores being associated with lower AUC<sub>g</sub> cortisol ( $\beta$

= -0.012, 95% CI [-0.019; -0.005],  $p < 0.001$ ,  $R^2 = 95.17\%$ ) in the child maltreatment group compared to the control group (for awakening cortisol:  $p = 0.086$ ,  $R^2 = 64.43\%$ ).

### 3.2.3 Cortisol stress reactivity

#### 3.2.3.1 Included studies

In total, our systematic search strategy identified  $n = 73$  studies that measured cortisol in the context of a stressor. Of these,  $n = 35$  publications ( $k = 39$  comparisons) were included. Owing to a lack of statistical information, the data of  $n = 22$  studies that were eligible for inclusion could not be considered. The total sample of the  $k = 39$  comparisons consisted of  $n = 4,284$  (range: 17-699) participants with a mean age of 25.57 ( $SD = 12.33$ ) years and an average of 66.1% females ( $SD = 28.5\%$ , range: 0.0-100.0%;  $k = 2$  studies contained a purely male sample and  $k = 13$  a purely female sample).  $K = 10$  comparisons involved samples consisting of children and/or adolescents only,  $k = 26$  comprised exclusively adult participants and  $k = 3$  studies included both adolescent and adult subjects. Eleven studies ( $k = 12$ ) did not report on percentages of Non-Caucasians, while the percentage of Non-Caucasians in the remaining studies ranged between 0.0-88.7% ( $M = 35.8\%$ ,  $SD = 27.9\%$ ). With respect to psychopathology,  $k = 10$  comparisons included healthy participants,  $k = 6$  involved participants all fulfilling diagnostic criteria for a mental disorder and  $k = 10$  comparisons comprised participants with at least some fulfilling the diagnostic criteria for a mental disorder (with  $k = 6$  not matched in terms of psychopathology; however, in two of these studies, the authors were able to show that the presence of the specific mental disorder did not affect the cortisol data). Finally,  $k = 13$  comparisons did not report on psychopathology at all. Various instruments to assess child maltreatment were applied with  $n = 7$  studies relying on informant reports,  $n = 25$  ( $k = 29$ ) on self-report data and three studies using both information sources. The most frequently used self-report was the Childhood Trauma Questionnaire (CTQ;  $n = 15$ ,  $k = 17$ ), and accordingly, cut-off scores were mostly used to group study participants in these studies. Nevertheless, several other instruments were also employed, resulting again in various grouping approaches. It should

be noted that five studies did not focus on child maltreatment only (Hengesch et al., 2018; Ivanov et al., 2011; Kaiser et al., 2018; Otte et al., 2005; Ouellet-Morin et al., 2011) but also included participants with other ELA experiences. By far, the most frequently applied stress task was the Trier Social Stress Test (TSST) or the TSST-C ( $n = 18$ ,  $k = 21$ ) and the majority of studies contained some social-evaluative aspects ( $k = 29$ ; for an overview of the different tasks applied in the various studies see Appendix E). The average duration of the stressors used was about 19.02 ( $SD = 16.94$ ) min ( $k = 27$  between 10 and 20 min). In  $n = 3$  studies no cortisol response following the onset of the corresponding stressor was observed. Interestingly, these studies all applied stressors that did not contain any social-evaluative challenges. In  $k = 29$  comparisons a cortisol response was observed in both groups (with different peak times found for  $k = 5$  comparisons), and finally in  $k = 7$  comparisons, the response was observed only in one but not in the other group ( $k = 6$  only in the control sample,  $k = 1$  only in the child maltreatment sample). On average, the time between the onset of a stressor and peak cortisol levels being reached was 29.84 ( $SD = 15.98$ ) min, with the majority of studies reporting that the peak was reached between 20-40 min post stressor onset ( $k = 26$ ). The vast majority of studies used saliva samples to assess cortisol. Baseline, peak and recovery data were reported by most publications, with considerably fewer studies reporting on  $AUC_i$  or  $AUC_g$  indices. Finally, the data of  $n = 19$  ( $k = 20$ ) studies were provided by corresponding authors, with the data of  $k = 13$  comparisons being (re)grouped. For further details on the characteristics of the included studies, see Table 5.

### 3.2.3.2 Risk of bias assessment

Studies assessing cortisol in the context of a stressor received an average total score of 58.1/100.0 ( $SD = 12.4$ , range: 32.1-78.6). With respect to the selection of participants ( $M = 66.7/100.0$ ,  $SD = 15.8$ ) less than half of the included studies ensured that all participants in the child maltreatment group were exposed to child maltreatment, while none of the participants in the control group were ( $k = 18$ ), used at least two different sources of information to establish

the presence of child maltreatment ( $k = 8$ ) and ensured that participants were matched with respect to psychopathology assessed with corresponding self-report questionnaires ( $k = 19$ ). Most studies however, employed an established measure to assess child maltreatment, matched their participants with respect to age, sex and psychopathology (assessed with a gold-standard diagnostic tool). Regarding the appropriate measurement of cortisol in the context of a stressor ( $M = 61.7/100.0$ ,  $SD = 13.4$ ), less than half of the included studies reported on whether sampling was rescheduled if participants were sick ( $k = 9$ ), ensured that all women were tested during a specific period of their menstrual cycle ( $k = 13$ ), reported on whether samples were analyzed in one batch ( $k = 7$ ) and only  $k = 5$  comparisons included measures attempting to ensure that none of the participants were under any current stress at the time of testing, or if testing was rescheduled if participants experienced any stressor during the respective day. Finally, concerning the appropriate control of potential confounders ( $M = 44.3/100.0$ ,  $SD = 24.0$ ), less than half of the included studies made efforts to exclude participants with any medical condition ( $k = 17$ ) known to influence HPA axis functioning, ensured that the groups did not differ with respect to smoking ( $k = 14$ ), clearly stated whether pregnant women were excluded ( $k = 14$ ), ensured that participants did not differ with respect to the intake of medications known to influence the CNS ( $k = 10$ ), and finally, only  $k = 5$  comparisons took measures to ensure the two groups did not differ with respect to other types of ELA or adult adversity. The detailed quality ratings for the individual studies as well as the detailed description of the individual quality items can be found in Appendix B Table B10 or Table B3.

### 3.2.3.3 Meta-analysis

The pooled effect estimates for the different indices are displayed in Table 6. The results of the sensitivity analyses (where appropriate) are also presented. As shown in the corresponding lines of Table 6, the results of the meta-analyses on baseline cortisol showed no significant overall differences in cortisol levels assessed prior to the onset of the respective stress task between the child maltreatment and the control sample (holding true for the sensitivity analyses). In contrast,

the release of cortisol following the perception of a stressor – expressed as peak, recovery, delta and  $AUC_i$  cortisol – was lower in the child maltreatment group compared to the control sample (with all pooled effect estimates being in the small range), indicating a blunted cortisol stress reactivity (see Appendix C Figs. 3.1 to 3.6 for corresponding forest plots). For  $AUC_g$ , the pooled effect estimate was not statistically significant ( $p = 0.081$ ). However, when excluding one outlier study (Ivanov et al., 2011), significance was also reached for this outcome index ( $p = 0.021$ ). Between-study heterogeneity was in the moderate to high range for some of the outcome indices ( $I^2 > 50\%$ ), exceeding the level of significance ( $Q$ -statistics all  $p < 0.05$ ) for all but  $AUC_i$  cortisol. Visual inspection of traditional and counter-enhanced funnel plots as well as Egger's regression test of funnel plot asymmetry revealed the absence of small-study bias for baseline, peak,  $AUC_g$  and  $AUC_i$  cortisol levels (all  $p > 0.05$ ). However, the Egger's regression test of funnel plot asymmetry reached significance for delta as well as recovery cortisol levels ( $p < 0.05$ ), suggesting the presence of small-study bias (see Appendix C Figs. 3.1 to 3.6 for corresponding funnel plots).

#### 3.2.3.4 Meta-regression and subgroup analyses

We conducted a number of pre-defined meta-regression and subgroup analyses focusing on peak, delta, recovery,  $AUC_g$  and  $AUC_i$  cortisol. The summary results for each outcome index and each moderator examined are shown in Appendix D Table D3. In the following section, the results for moderators found to significantly influence the main effects are outlined. For delta, recovery and  $AUC_i$  cortisol, we identified the proportion of women in the sample as a continuous moderator (for  $AUC_i$ :  $p = 0.052$ ), with an increase in the proportion of females being associated with lower cortisol secretion following the perception of a stressor, when comparing the child maltreatment and the control sample. Additionally, for delta cortisol, the proportion of participants fulfilling diagnostic criteria for a mental disorder significantly moderated the summary effect, with an increase of the proportion being associated with a stronger blunting of the cortisol stress response ( $\beta = -0.006$ , 95% CI [-0.010; -0.002],  $p = 0.007$ ,  $R^2 = 99.20\%$ ). This

finding, however, should be interpreted with caution, as only two of the studies that included a purely clinical sample (Schalinski et al., 2015; Suzuki et al., 2014) reported on delta cortisol. In addition, and in contrast to the other studies involving a clinical sample (exception Rao & Morris, 2015), these two studies reported relatively strong negative effects. Nevertheless, despite considerable heterogeneity between the studies, all outcome indices showed stronger effects for studies including purely clinical samples and markedly weaker effects for those studies that involved healthy subjects only (see results subgroup analyses). Furthermore, stronger effects were found for studies that observed a cortisol response in just one of the groups (holding true for all outcome indices) compared to studies that found a response in both groups and those that found no response in either of the groups, with the subgroup comparison reaching significance for delta ( $Q_2 = 4.53, p = 0.033$ ) and  $AUC_i$  ( $Q_2 = 12.33, p = 0.002$ ) cortisol. Comparing studies focusing on child maltreatment experiences only to those involving participants with other types of ELA as well showed that the few studies that also considered other types of ELA overall yielded greater negative effect estimates for all outcome indices, but significant for delta cortisol only ( $Q_1 = 3.95, p = 0.047$ ). However, it should be noted that heterogeneity within these studies varied substantially between the different outcome indices and thus depended highly on the included studies. Finally, again depending on the outcome index investigated (and thus on the studies included), the different sub-domains of the quality assessment appeared to explain part of the variance in the effect estimates between studies, although this effect was only significant for  $AUC_i$  cortisol (and only for the subdomain: selection of participants:  $\beta = 0.009, 95\% \text{ CI } [0.002; 0.016], p = 0.011, R^2 = 77.45\%$ ). In general, there was a tendency that a higher study quality was associated with a smaller negative difference in cortisol secretion between the child maltreatment and the control group. As an additional note, although subgroup comparisons between studies with (re)grouped data to those with original data could not explain significant heterogeneity between studies for any outcome indices, those studies with (re)grouped data still showed considerably less pronounced effects.



### 3.2.4 Pharmacological challenge tests

#### 3.2.4.1 Dexamethasone suppression test

##### 3.2.4.1.1 Included studies

Eleven articles, containing  $k = 17$  comparisons, involving a total of  $n = 2,222$  participants (range: 16-1,112) assessed cortisol in the context of the DST. Of these,  $k = 16$  reported on baseline cortisol levels (cortisol assessed before dexamethasone administration; pre-DST),  $k = 17$  on cortisol assessed following the administration of dexamethasone (post-DST) and  $k = 9$  contained information on delta values (post-DST cortisol minus pre-DST cortisol). The included studies mainly consisted of adults, with only one study involving adolescents. The average age was 33.32 (SD = 8.32) years and studies ranged from 45.6-100.0% ( $M = 72.6%$ ,  $SD = 18.6%$ ) in terms of the proportion of women ( $k = 3$  studies with a purely female sample). Five studies ( $k = 8$ ) did not report on the percentage of Non-Caucasians, while the percentage of Non-Caucasians in the remaining studies ranged between 0.0-100.0% ( $M = 53.4%$ ,  $SD = 41.2%$ ). Three out of the  $k = 17$  comparisons involved healthy participants and  $k = 14$  included participants in whom the proportion of people suffering from a mental illness ranged from 13.2-100.0% ( $k = 9$  studies involved purely clinical samples and  $k = 3$  involved participants where the child maltreatment and the control sample were not matched in terms of psychopathology). Various instruments to assess child maltreatment were applied, all relying on self-report information. The most common self-report used was the CTQ ( $n = 5$ ,  $k = 9$ ). It should be noted that the child maltreatment sample of the study from Faravelli et al. (2010) did not only consist of participants with child maltreatment experiences, but also included several participants with loss experiences. Approximately half of the studies used established cut-off values to group participants in the corresponding child maltreatment and control groups, with the others mostly applying specific definitions. All but three studies ( $k = 4$ ) used 0.5 mg of dexamethasone and the data of four studies ( $k = 8$ ) were re-grouped for this meta-analysis (see Table 7 for more details).

#### 3.2.4.1.2 Risk of bias assessment

Studies assessing cortisol in the context of the DST received an average score of 72.1/100.0 ( $SD = 19.0$ ) for selection of participants, 61.3/100.0 ( $SD = 12.1$ ) for appropriate assessment of cortisol and 44.7/100.0 ( $SD = 20.3$ ) for adequate controlling for confounders, resulting in an average overall score of 58.1/100.0 ( $SD = 13.3$ , range: 40.0-76.0). The detailed quality ratings for the individual studies as well as the detailed description of the individual quality items can be found in Appendix B Table B11 or Table B4. None of the studies included used two different sources to establish the presence of child maltreatment, reported whether cortisol was analyzed in one batch, and whether participants were excluded when working night shifts, and only one study assessed whether exposure and control groups differed in relation to the experience of other traumatic events during childhood or adulthood. Moreover, less than half of the comparisons reported whether sampling was postponed when participants were sick ( $k = 7$ ), whether dexamethasone intake was checked ( $k = 7$ ), whether participants differed in smoking ( $k = 6$ ), intake of oral contraceptives ( $k = 7$ ) and their use of medication (with CNS effect;  $k = 4$ ).

#### 3.2.4.1.3 Meta-analysis

The results of the meta-analyses yielded no significant overall differences in pre-DST cortisol (Hedges'  $g = 0.07$ , 95% CI [-0.09; 0.23],  $p = 0.402$ ), post-DST cortisol (Hedges'  $g = 0.01$ , 95% CI [-0.18; 0.20],  $p = 0.936$ ) nor in delta cortisol values (Hedges'  $g = -0.13$ , 95% CI [-0.31; 0.06],  $p = 0.178$ ) between participants with child maltreatment experiences and participants without corresponding experiences. Between-study heterogeneity was in the low to moderate range (all  $I^2 < 45\%$ ), reaching significance for pre- ( $Q_{15} = 25.21$ ,  $p = 0.047$ ) and post-DST ( $Q_{16} = 28.64$ ,  $p = 0.027$ ) cortisol (see Appendix C Figs. 4.1.1 to 4.1.3 for corresponding forest plots). Visual inspection of traditional and counter-enhanced funnel plots as well as the results of the Egger's regression tests (all  $p > 0.316$ ) implied absence of small-study bias (for funnel plots see Appendix C Figs. 4.1.1 to 4.1.3). Corresponding sensitivity analyses excluding studies with

extreme effect sizes as well as influential studies did not change the overall results (see Table 8 for related statistics).

#### 3.2.4.1.4 Meta-regression and subgroup analyses

We conducted a number of pre-defined meta-regression and subgroup analyses focusing on post-DST cortisol only. The summary results for each moderator examined are shown in Appendix D Table D4. The only moderator that significantly influenced the main effect was the proportion of women in the respective samples ( $\beta = -0.014$ , 95% CI [-0.022; -0.005],  $p = 0.003$ ,  $R^2 = 95.16\%$ ), with an increase of the proportion being associated with lower cortisol levels following the administration of dexamethasone (increased cortisol suppression) when comparing the child maltreatment sample with participants from the control sample. Since only one study was included that focused on ELA in general, the significant subgroup result of different pooled effect estimates for studies focusing on child maltreatment only and the study including also other childhood adversities has to be interpreted with caution. None of the methodological quality criteria significantly influenced the pooled effect estimate.

#### 3.2.4.2 Combined dexamethasone-corticotropin releasing hormone test

##### 3.2.4.2.1 Included studies

In total, our search strategy identified  $n = 21$  studies that measured the responsivity of the pituitary to CRH. Of these,  $n = 6$  studies consisting of  $k = 8$  comparisons reporting on cortisol in the context of the Dex-CRH test and  $n = 2$  studies reporting on cortisol in the context of the CRH test, were included ( $k = 10$ ). Of the included studies,  $k = 4$  comparisons reported on baseline cortisol (cortisol assessed after the administration of dexamethasone, before CRH injection),  $k = 6$  on peak (after the CRH injection) and delta cortisol (peak minus baseline),  $k = 9$  on  $AUC_g$  and  $k = 6$  on  $AUC_i$  cortisol. There was only one study with available recovery data (in all other studies, the peak value corresponded to the last measurement time point). The total sample of the  $k = 10$  comparisons consisted of  $n = 561$  participants (range: 21-230) with a mean

age of 31.19 ( $SD = 12.70$ ) years and an average of 60.0% females ( $SD = 39.6\%$ , range: 0.0-100.0%;  $k = 2$  comparisons contained a purely male sample and  $k = 4$  a purely female sample). Studies reporting on Dex-CRH cortisol consisted of adult samples only, whereas the two studies focusing on the CRH test were conducted with children or adolescents. Four studies ( $k = 5$ ) did not report on ethnicity with the other articles ranging between 34.6-57.1% in terms of percentage of Non-Caucasians ( $M = 45.9\%$ ,  $SD = 9.3\%$ ). With respect to psychopathology,  $k = 3$  comparisons included healthy participants,  $k = 3$  involved participants all fulfilling diagnostic criteria for major depressive disorder (MDD),  $k = 1$  comparison included participants with mixed diagnoses and  $k = 3$  involved participants where the child maltreatment and the control sample were not matched for psychopathology. The authors of the three comparisons in which the subjects were not matched for psychopathology, however, showed that the presence of the specific mental disorder did not affect the cortisol results. Again, various instruments to assess child maltreatment were used, with the majority of studies relying on self-report data ( $k = 8$ ). The grouping of participants also differed, including the use of cut-off scores, thirds, the presence of a CPS record and the utilization of specific definitions. The data of  $n = 2$  studies were provided by corresponding authors, with the data of one study being re-grouped. See Table 9 for more details.

#### 3.2.4.2.2 Risk of bias assessment

Studies assessing the responsivity of the pituitary to CRH, on average received a total score of 69.8/100.0 ( $SD = 8.0$ , range: 58.3-83.3). All studies, or  $k = 9$  out of 10 comparisons used an established instrument to assess the experience of child maltreatment, matched their participants with respect to age and sex, and made efforts to assess psychopathology (selection of participants:  $M = 76.3/100.0$ ,  $SD = 12.4$ ). Additionally, most comparisons rescheduled the sampling when participants were sick ( $k = 8$ ), provided details on their test protocol ( $k = 10$ ), as well as on how cortisol was collected, stored and analyzed ( $k = 10$ ; appropriate measurement of cortisol in the context of the Dex-CRH:  $M = 62.3/100.0$ ,  $SD = 14.8$ ). Finally, as shown by

the relatively high scores related to the control of confounding variables ( $M = 69.0/100.0$ ,  $SD = 10.3$ ), the majority of studies assessed or controlled for a wide variety of potential influential factors. Overall, however, only a few comparisons used two different sources of information to establish the presence of child maltreatment ( $k = 2$ ), verified the ingestion of dexamethasone ( $k = 2$ ), reported on whether samples were assessed in one batch ( $k = 1$ ), or whether participants were excluded in case of working night shifts ( $k = 2$ ) and only  $k = 2$  comparisons made any effort to ensure that the exposure and control group did not differ with respect to other types of ELA or adult adversity. For details on individual scoring results of the primary studies as well as a summary of the average risk of bias scores see Appendix B Table B12 or Table B5 for individual quality items.

#### 3.2.4.2.3 Meta-analysis

The pooled effect estimates for the different indices are displayed in Table 10. For each outcome index the analysis was repeated excluding the studies focusing on the CRH test only. Results of the sensitivity analyses (where appropriate) are also presented. None of the pooled effect estimates were significant, indicating that there is no overall difference in cortisol assessed both before and after the administration of CRH and holding true for all outcome indices (all  $p > 0.05$ ). Between-study heterogeneity was however in the moderate range ( $I^2 > 50\%$ ), exceeding the level of significance ( $Q$ -statistics all  $p < 0.05$ ) for peak, delta,  $AUC_g$  and  $AUC_i$  cortisol. Due to a limited number of studies ( $k < 10$ ), however, no subgroup analyses and meta-regressions were performed. Visual inspection of traditional and counter-enhanced funnel plots as well as Egger's regression test of funnel plot asymmetry suggested the absence of small-study bias (all  $p > 0.05$ ). Forest and corresponding funnel plots for  $AUC_g$  and  $AUC_i$  cortisol are displayed in Appendix C Figs. 4.2.1 and 4.2.2.

### 3.2.5 Cumulative measures of cortisol secretion

#### 3.2.5.1 Hair cortisol concentrations

##### 3.2.5.1.1 Included studies

A total of  $n = 8$  independent studies, comprising  $k = 9$  comparisons, with an overall sample size of  $n = 978$  participants, reported on HCC. The sample size ranged from  $n = 22$  to  $n = 537$  participants and the majority of studies included mainly female subjects (with the percentage of females ranging between 50.7-100.0%,  $M = 84.4%$ ,  $SD = 19.2%$ ;  $k = 3$  contained a purely female sample). Four comparisons included children and/or adolescents and  $k = 5$  included adult participants only. The average age was 28.13 ( $SD = 14.88$ ) years with the youngest participant being about 3 years and the oldest around 79 years. With respect to ethnicity, most studies included samples composed mainly of an ethnic majority group with the percentage of Non-Caucasians ranging between 0.0-87.2% ( $M = 27.0%$ ,  $SD = 33.8%$ ) with three studies not reporting on ethnicity. Concerning psychopathology,  $k = 4$  comparisons included only healthy participants,  $k = 2$  consisted of a primarily clinical sample (with 96.0-100.0% meeting diagnostic criteria for a mental disorder) and  $k = 3$  comparisons did not report on psychopathology. Only one study used information about child maltreatment from an informant source, with the others all using self-report data. The assessment of child maltreatment and the grouping of participants into a child maltreatment and a control group varied between the studies. This refers both to the instruments used and to the grouping procedure, which included the use of cut-off scores, percentiles, clustering methods as well as the use of specific definitions. Of the six studies ( $k = 7$ ) that provided data on request, the respective authors of four studies ( $k = 5$ ) regrouped or grouped their data based on the available assessment of child maltreatment (or, in case of raw data, the (re)grouping was performed by us). For further details on the characteristics of the included studies, see Table 11.

### 3.2.5.1.2 Risk of bias assessment

Table B13 in Appendix B provides an overview of the individual scoring results of the primary studies as well as a summary of the average risk of bias scores. The detailed description of the quality items can be found in Table B6. On average a total score of 58.9/100.0 was received ( $SD = 16.6$ , range: 24.0-76.0). Regarding participant selection ( $M = 75.0/100.0$ ,  $SD = 10.8$ ), the majority of studies assessed child maltreatment with an established instrument and ensured that all the participants in the child maltreatment group did experience child maltreatment, while none of the participants in the control group did. Most studies also matched participants with respect to age, sex as well as psychopathology. However, only one study used two different sources of information to establish the presence of child maltreatment. With respect to the appropriate measure of HCC ( $M = 55.6/100.0$ ,  $SD = 15.1$ ), most studies obtained hair samples from the posterior vertex of the head and reported on a clear sampling analysis protocol as well as information about outlying or missing data. The majority of the studies however, neither assessed the experience of any ongoing life stressor ( $k = 0$ ) nor whether HCC samples were assessed in one batch ( $k = 2$ ). With respect to the appropriate control of confounding variables, the quality of the different studies varied quite strongly ( $M = 48.3/100.0$ ,  $SD = 33.3$ , range: 0.0-88.9). It should be noted, however, that several studies would have assessed the variables of interest, but since the data of four studies ( $k = 5$ ) were regrouped, the information at the group level was no longer available (this is marked accordingly in the table). No studies reported whether participants were excluded if they worked night shifts ( $k = 0$ ) and few reported whether participants in the two groups did not differ with respect to other traumatic experiences in childhood or adulthood ( $k = 3$ ). Moreover, less than half of the comparisons reported whether participants with any type of addiction ( $k = 4$ ) or pregnant women ( $k = 4$ ) were excluded as well as whether participants were comparable with respect to medication use (medications with CNS effects,  $k = 4$ ).

### 3.2.5.1.3 Meta-analysis

Pooling the results of the  $k = 9$  comparisons ( $n = 978$ ), we found no significant effect (Hedges'  $g = -0.05$ , 95% CI [-0.33; 0.24],  $p = 0.749$ ), suggesting no overall difference in HCC in the child maltreatment sample compared to the control sample (the corresponding forest plot is shown in Appendix C Fig. 5.1). There was significant, moderate, heterogeneity in the effect size estimates between studies ( $Q_8 = 17.14$ ,  $p = 0.029$ ,  $I^2 = 53.3\%$ ). The between-study heterogeneity was not caused by extreme effect sizes as there was no such outlier study. However, one study exerting a high influence on the overall effect estimate was identified (do Prado et al., 2017). Traditional and contour-enhanced funnel plots are shown in Appendix C Fig. 5.1 and visual inspection of them suggested absence of small-study bias as did the Egger's regression test of funnel plot asymmetry ( $intercept = 0.970$ ,  $p = 0.285$ ; but attention  $k < 10$ ). Excluding the study from do Prado et al. (2017) as part of the sensitivity analysis, heterogeneity decreased from  $I^2 = 53.3\%$  to  $I^2 = 10.2\%$  ( $Q_7 = 7.80$ ,  $p = 0.351$ ), yielding a small negative effect, which reached significance ( $k = 8$ ,  $n = 921$ , Hedges'  $g = -0.20$ , 95% CI [-0.34; -0.06],  $p = 0.004$ ). Despite varying effects of the primary studies, the result of the sensitivity analysis suggests an overall reduction of HCC in the child maltreatment sample compared to the control sample, with the prediction interval (-0.37; -0.03) pointing in the same direction. However, three of the five studies indicating reduced HCC in the child maltreatment compared to the control group had not used a gold-standard diagnostic tool to assess psychopathology and thus matching in this respect is not properly judgeable.

### 3.2.5.2 24-hour urinary free cortisol

#### 3.2.5.2.1 Included studies

Eleven studies assessing cortisol in urine were identified through the systematic search. Of these, only  $n = 4$  studies including a total of  $n = 110$  participants ( $n = 108$  with valid cortisol data) could finally be included. Participants were on average 22.17 ( $SD = 12.94$ ) years old and three out of the four studies comprised female participants only ( $M = 85.1\%$ ,  $SD = 29.8\%$ ). The



majority of the participants were Caucasian, with the percentage of Non-Caucasians ranging between 12.0-42.3% ( $M = 27.6\%$ ,  $SD = 15.2\%$ ). With respect to psychopathology, most of the subjects in the child maltreatment group met the criteria for a mental disorder, while the subjects in the control group were mainly healthy controls ( $n = 3$ ;  $n = 1$  did not report on psychopathology). However, in two of the studies included, the authors were able to demonstrate that the presence of the specific mental disorder did not affect the 24-hour UFC data. Three of the four studies focused exclusively on sexual abuse experiences without collecting information about other types of child maltreatment, and two studies recruited participants solely on the basis of self-identification without using any established measurement to assess child maltreatment. All data were extracted from the respective articles. See Table 12 for further details.

#### 3.2.5.2.2 Risk of bias assessment

The risk of bias assessment of the primary studies as well as the average risk of bias scores can be found in Appendix B Table B14. The detailed description of the quality items can be found in the Table B7. On average, a total score of 61.5/100.0 was received ( $SD = 13.5$ , range: 42.3-73.1; selection of participants:  $M = 65.6/100.0$ ,  $SD = 12.0$ ; appropriate measure of 24-hour UFC:  $M = 50.0/100.0$ ,  $SD = 18.4$ ; adequate controlling of confounders:  $M = 66.5/100.0$ ,  $SD = 20.7$ ). In all  $n = 4$  studies participants were matched with respect to age and gender. In addition, all four articles provide detailed information on how UFC samples were collected, stored and analyzed and all studies give a relatively good overview of participants' medication use. The  $n = 3$  studies that focused on sexual abuse did not provide information about other maltreatment experiences, reducing the quality of the grouping into a child maltreatment and a clear control group without any types of child maltreatment experiences. In only one of the four studies it was ensured that the participants did not experience any ongoing significant life stressors, assessed UFC over at least three days, provided batch analysis information, and none of the

studies ensured that the participants in the two groups did not differ with respect to other traumatic experiences in childhood or adulthood.

### 3.2.5.2.3 Meta-analysis

Pooling the results of the  $n = 4$  studies ( $n = 108$ ), the aggregate effect size was Hedges'  $g = 0.07$ , 95% CI [-0.83; 0.98],  $p = 0.874$ , suggesting no overall difference in 24-hour UFC in the child maltreatment sample compared to the control sample (the corresponding forest plot is shown in Appendix C Fig. 5.2). There was significant, high, heterogeneity in the effect size estimates ( $Q_3 = 14.79$ ,  $p = 0.002$ ,  $I^2 = 79.7\%$ ), indicating high inconsistencies between studies. No outlier study was detected, but the study from Lemieux et al., (2008) had a high influence on the overall result. Traditional and contour-enhanced funnel plots are also shown in Appendix C Fig. 5.2 and visual inspection of them suggested the absence of small-study bias as did the Egger's regression test of funnel plot asymmetry ( $intercept = -6.898$ ,  $p = 0.428$ ; but attention  $k < 10$ ). The sensitivity analysis substantially reduced heterogeneity ( $I^2 = 79.7\%$  to  $I^2 = 8.5\%$ ) and yielded a medium significant overall effect ( $n = 83$ , Hedges'  $g = 0.56$ , 95% CI [0.11; 1.00],  $p = 0.014$ ) with participants in the child maltreatment group showing higher 24-hour UFC concentrations compared to the control sample. However, considering the small sample size and the large prediction interval (-2.33; 3.45), it is unclear what the results of future studies will show. In addition, it should be noted that the study excluded in the context of the sensitivity analysis (Lemieux et al., 2008) received the highest average quality score.

## 4.0 Discussion

This series of meta-analyses, based on a systematic review of the literature, examined the existing evidence on the association between child maltreatment and cortisol metabolism including various measures of HPA axis activity. Measures of interest ranged from cortisol assessed in the context of the circadian rhythm (DC) to cortisol assessed in response to awakening (CAR), in response to the perception of a stressor (cortisol stress reactivity) and pharmacological challenges (DST, Dex-CRH test, CRH test), to cumulative measures of cortisol secretion, namely 24-hour UFC and HCC.

### 4.1 Main findings

Consistent with the findings of two previous meta-analyses (Bernard et al., 2017; Fogelman and Canli, 2018) we did not find overall differences in any of the indices related to cortisol secretion in the context of circadian activity (with the exception of evening cortisol) as well as in response to awakening (CAR) when comparing individuals with child maltreatment to those without corresponding experiences. The finding of slightly increased evening cortisol in individuals exposed to child maltreatment was mainly driven by a few studies in which the child maltreatment group also included individuals with loss experiences and thus should be interpreted with caution in the context of this meta-analysis. Individuals with a history of child maltreatment, however, appear to show a blunted cortisol stress response. Though not yet evident before being introduced to a corresponding stressor (baseline cortisol), blunting was seen in indices reflecting total cortisol production (peak, recovery cortisol) as well as in indices expressing changes in cortisol over time (delta, AUC<sub>i</sub> cortisol) following the perception of a stressor. These findings are consistent with the results of a previous meta-analysis examining the effect of ELA on cortisol response to social stress (Bunea et al., 2017), albeit with somewhat smaller effects observed in our meta-analysis. Interestingly, this blunting was not observed in studies where CRH injections (Dex-CRH test) were used to initiate the secretion of cortisol. However, the number of studies on the Dex-CRH test was much smaller compared to the

number of studies assessing cortisol in response to a stressor. In addition, no difference in the negative feedback mechanism of the HPA axis (at least at the level of the pituitary gland), measured by oral administration of dexamethasone, was found between the two groups. Finally, with respect to the few studies reporting on cumulative measures of cortisol secretion including 24-hour UFC and HCC, no differences were observed in both of these measures between those exposed to child maltreatment and those without corresponding adversity. Respective sensitivity analyses excluding influential studies, performed within the context of these two outcome indices, on the other hand, suggest increased 24-hour UFC and slightly reduced HCC in maltreated individuals. However, especially the finding of increased 24-hour UFC should be interpreted with caution, since the overall sample size was small and the large prediction interval of the pooled effect estimate suggests a high degree of uncertainty regarding the results of upcoming studies (for a summary overview about all main findings see Table 13).

#### 4.2 Between-study heterogeneity and the influence of moderators

Although no overall differences in cortisol secretion were found for the majority of the HPA axis activity measures except for cortisol assessed in response to a stressor across studies, we generally observed a significant degree of variability in the effect estimates between studies (especially for indices reflecting cortisol secretion in the context of awakening, after a stressor and following the Dex-CRH test), suggesting the likely influence of additional variables in moderating the effect of child maltreatment on cortisol regulation. Before discussing some of the moderators that systematically accounted for between-study heterogeneity, holding true for various of the HPA axis activity measures, it should be kept in mind that the majority of studies were conducted with predominantly young, female adults who belonged to an ethnic majority group and in whom child maltreatment experiences were assessed mainly through self-reports. In addition, a considerable number of studies did not report on psychopathology and studies involving clinical samples were fairly heterogeneous in terms of the predominant mental

disorder (e.g., MDD versus PTSD). Accordingly, our ability to find important moderators or relevant subgroup differences might have been limited.

#### 4.2.1 Influence of participant related characteristics

One of the moderators that explained some of the between-study heterogeneity in effect sizes (in line with findings from Bunea et al., 2017; and Zorn et al., 2017) was the proportion of females in the respective sample, with a higher proportion being associated with a stronger blunting of cortisol secretion (CAR, cortisol stress reactivity and DST). Corresponding sex differences, particularly with respect to stress reactivity, have been repeatedly reported, with men showing higher cortisol levels to psychosocial stress than women (Liu et al., 2017). Factors that influence corticosteroid binding globulin (CBG) levels and thus the level of free cortisol appear to account for some of these gender effects including the use of oral contraceptives and the production of sex steroids throughout the menstrual cycle (e.g., Foley and Kirschbaum, 2010). Both the intake of oral contraceptives and the assessment of the menstrual cycle were not adequately evaluated in many of the studies included and therefore matching of the two groups in these respects was not properly controlled. Apart from the proportion of women in the respective sample, there was little evidence that the remaining participant related characteristics such as age, ethnicity and participant diagnosis accounted for variability in the effect estimates among primary studies. Interestingly, even though psychopathology did not account for heterogeneity in the child maltreatment cortisol relationship (at least for the majority of outcome indices), a tendency for stronger blunting in clinical samples compared to healthy controls was observed for indices related to cortisol stress reactivity. Nevertheless, an attenuation of the cortisol stress response was observed in participants with child maltreatment experiences that at the time of measurement did not report a mental disorder, suggesting that alterations in HPA axis activity may be present prior to the development of mental health issues or independent of psychopathology.

#### 4.2.2 Influence of trauma related information

Another fairly consistent moderator accounting for some of the between-study heterogeneity (DC and CAR, tendency for cortisol stress reactivity) was whether or not data were (re)grouped for the purpose of this meta-analysis, with (re)grouped data showing a tendency towards smaller effects. In some of the studies that provided (re)grouped data, grouping of participants into a child maltreatment and a control group was based on relatively low severity thresholds (particularly for cortisol assessed in the context of awakening and circadian activity), which might account for this finding. Indeed, and in agreement with the observed dose-dependent relationship between child maltreatment and health impairments (e.g., Clemens et al., 2018; Norman et al., 2012), the severity of child maltreatment experiences, though difficult to assess (Jackson et al., 2019), might actually be of particular importance in explaining variability between the association of child maltreatment and HPA axis functioning. Interestingly, several studies, especially those which assessed cortisol in response to a stressor, found a stronger blunting in cortisol secretion following the perception of a stressor with an increase in the severity of child maltreatment (Lovallo et al., 2019; Ouellet-Morin et al., 2018; Trickett et al., 2014; Voellmin et al., 2015). Unfortunately, our group comparison approach did not allow us to investigate this association systematically. In line with the difficulties in defining child maltreatment and the possibility to rely on various assessment modalities (Cicchetti and Toth, 2005; Manly, 2005), studies generally differed widely in their child maltreatment assessment and grouping approaches. For instance, studies that relied on established self-reports such as the CTQ or the Childhood Experience of Care and Abuse interview (CECA) grouped their participants based on validated cut-off scores, while other studies applied specific definitions (e.g., Heim et al., 2000: “repeated abuse, once a month or more for at least 1 year”), sometimes based on self-developed assessment tools (e.g., Groër et al., 2016; Martinson et al., 2016; Smeets et al., 2007) and still others relied on the presence or absence of a specific record, such as a CPS record (e.g., Bernard et al., 2010; Cicchetti et al., 2010; Hibel et al., 2019). However,

neither the assessment (self-report, informant report, mixed) nor the grouping method (cut-offs, other, record) explained variance in the effect estimates of any of the HPA axis activity outcome indices, with the exception of larger effect sizes found in studies focusing on informant reports compared to self-reports for waking/morning cortisol (which is consistent with findings from Bernard et al., 2017). Since there were far fewer studies using informant reports as opposed to self-reports (a pattern also found among studies on the prevalence of child maltreatment; Stoltenborgh et al., 2015), comparison between these approaches might have been inappropriate and the chances of detecting differences accordingly low. Importantly, some of the included studies failed to sufficiently ensure that control participants were not subjected to any type of child maltreatment. The studies that performed poorly in terms of ensuring maltreatment did not take place in the control group were those that relied on specific records such as CPS records and studies focusing on one particular type of child maltreatment. In corresponding studies - besides the absence of a record or the corresponding type of maltreatment - no other measures were applied to ensure that control participants had not experienced any child maltreatment. This may have influenced our results, since child maltreatment experiences are rather common in the general population, as demonstrated by epidemiological studies (Witt et al., 2017). Neither the role of age at maltreatment onset nor the chronicity of the maltreatment experiences on HPA axis activity could be investigated, as very few studies applied measures that assessed these two factors in the first place - a finding in line with a recent review summarizing research on the operationalization of child maltreatment over the last 10 years (Jackson et al., 2019).

#### 4.2.3 Influence of cortisol related information

In our meta-analysis, neither sample type (blood, saliva), type of stressor (social-evaluative, other), slope type (wake-to-bed, other), whether cortisol was assessed in reference to awakening or not, nor dose of dexamethasone (0.5 mg, 1.0 mg) significantly explained variability in the various effect estimates. The only moderator related to the assessment of cortisol that explained between-study heterogeneity, was whether a cortisol response was observed in both, in only

one, or in none of the groups following the perception of a stressor, with stronger effects found in those studies that observed a cortisol response in just one of the two groups. This finding might be attributed to the fact that six out of the seven comparisons that observed a response only in one of the two groups, reported an increase in cortisol in the control group only. Overall, there was no evidence to suggest that the two components of cortisol secretion - total cortisol production and change in cortisol over time - which appear to capture different aspects of HPA axis activity (Khoury et al., 2015), are affected differently by child maltreatment. A blunting in cortisol secretion following the perception of a stressor was found in indices reflecting both total and change in cortisol over time. However, as expected, studies generally differed widely with respect to the index or indices reported (e.g., much more studies reported on peak cortisol compared to  $AUC_i$  cortisol) making comparisons between the different outcome indices difficult. Along with this, depending on the studies included, different moderators emerged, which in turn complicated the interpretation of the corresponding findings.

#### 4.2.4 Influence of several components of methodological quality

Finally, we did not observe a consistent association between study quality as assessed by our self-developed quality assessment tool (which was based on existing recommendations and guidelines) and reported effect sizes. Although, at least for some outcome indices, the quality of the individual studies seemed to explain some of the between-study heterogeneity. For instance, a tendency towards smaller negative differences in cortisol secretion following the perception of a stressor was observed in studies with a higher study quality and thus a lower risk of bias. Importantly, in studies that (re)grouped participants for the purpose of this meta-analysis - although the majority of these studies used an established instrument (an established source of information) to assess child maltreatment experiences - by (re)grouping their study participants, several other aspects of methodological quality could no longer be assessed. Therefore, our conservative approach of not awarding any points in a given case may have induced biases. In addition, two studies were able to achieve the same average total score, but



scored in completely different quality items. Since we were not able to value the importance of the various quality items, our ability to find associations with the different effect estimates might thus have been limited. Nevertheless, it should be mentioned that several quality-related aspects were generally well implemented in the majority of the studies (holding true for all HPA axis activity measures), whereas others were insufficiently addressed and controlled in most of the included publications. With respect to the selection of participants, for instance, the two comparison groups were generally matched in terms of sex and age, and most studies used an established measurement to assess child maltreatment. While exposure and the control groups were generally balanced in terms of psychopathology in those studies that reported on psychopathology, a substantial number of studies did not evaluate the presence of mental disorders at all and therefore information about matching in this respect was unavailable for several studies. This is particularly surprising as, on the one hand, child maltreatment experiences are much more common among individuals with a mental disorder and, on the other hand, psychopathology itself has been repeatedly associated with changes in various HPA axis activity measures, with sometimes opposing findings for different mental disorders (e.g., Adam et al., 2017; Chida and Steptoe, 2009; Leistner and Menke, 2018; Stalder et al., 2017; Zorn et al., 2017). Accordingly, in these studies, the presence of a possible confounding effect of psychopathology cannot be ruled out. Interestingly, a recent study examining the effects of comorbidity and adversity on HPA axis functioning in depressed patients was able to show that rather than the diagnostic groups per se, the timing of adversity appears to influence HPA axis functioning in adulthood, putting the importance of psychopathology and especially the role of diagnostic groups somewhat into perspective. In this study, an attenuated HPA axis stress response was only found in those patients with comorbid PTSD from childhood. By contrast, no alterations were seen in those with depression only, or those with depression with comorbid PTSD resulting from adult trauma (Mayer et al., 2020). Thus, these results, consistent with the findings of the present meta-analysis and those of the meta-analysis by Bunea et al. (2017),

suggest that adverse experiences during childhood indeed appear to be of particular importance in influencing the HPA axis stress response in adulthood.

Related to participant selection, and as already indicated in the context of the assessment of child maltreatment, several studies inadequately ensured that none of the control participants were exposed to any type of child maltreatment, and only a handful of studies used two different sources to evaluate the presence of child maltreatment. Regarding appropriate assessment of cortisol in the context of the corresponding HPA axis activity measure, most studies reported on clear sampling (prohibitions) and collection instructions (i.e., how they collected, stored and analyzed samples), provided details on their test protocol and generally reported on missing and/or outlier data. By contrast, very few studies provided information on whether sampling was rescheduled if participants were sick, on batch analysis or ensured that participants were not under any current stress at the time of testing (sampling), factors known to influence cortisol results (e.g., Adam and Kumari, 2009). Furthermore, only a few studies that assessed cortisol in the context of daily activity (DC and CAR) ensured that exposure and control groups did not differ in the time of awakening as well as sampling time adherence. This is particularly surprising as the validity of the CAR measurement critically depends on the sampling schedule, with inaccurate sampling strongly biasing CAR (including morning/waking) estimates (Stalder et al., 2016). Moreover, a study investigating the variability and reliability of DC indicated that a 10-day sampling procedure would be required to obtain stable estimates of between-person differences in DSL cortisol (Segerstrom et al., 2014). Similarly, up to six assessment days might be necessary to obtain reliable CAR trait measures (Hellhammer et al., 2007). However, only a few studies included in this systematic review assessed cortisol over more than two days. Lastly, a substantial number of studies failed to adequately assess and thus control for important confounding variables. As mentioned in the context of sex differences, the matching of participants in terms of oral contraceptive use and menstrual cycle timing was insufficient in many studies. Other confounding variables, which were generally poorly assessed and

controlled for, were: smoking, medication intake with known CNS effects, and clear statements about whether pregnant women and participants working night shifts were excluded, factors also known to account for variability in cortisol results (e.g., Kudielka et al., 2012; Locatelli et al., 2009; Stalder et al., 2016; Zänkert et al., 2019). Finally, only a few of the included studies took measures to ensure that participants from the control group were not subjected to any type of ELA other than child maltreatment. According to the National Scientific Council on the Developing Child (e.g., Shonkoff et al., 2012), three types of stressors can be differentiated according to their potential to cause enduring physiological disruptions. These include: positive, tolerable and toxic stressors. “Tolerable” stress experiences include those that present a great magnitude of adversity or threat, such as the death of a family member, or a serious illness or injury. However, when buffered by a supportive adult, the risk that corresponding circumstances will cause long-term consequences for health are suggested to be greatly reduced. In contrast, toxic stress experiences include those that are experienced in the absence of a supportive, adult relationship and may cause strong, frequent, or prolonged activation of the body’s stress response system. Since child maltreatment experiences typically occur in the absence of the buffering protection of stable adult support, these experiences are suggested to be particularly toxic and thus show a great potential to induce long-long-lasting biological changes. In line with this, child maltreatment experiences do show high associations with later disease risk (e.g., Dube et al., 2001). Nevertheless, ELA and especially the experience of multiple adverse childhood experiences have been related to various health conditions later in life as well (e.g., Clark et al., 2010; Danese et al., 2009; Hughes et al., 2017). In addition, the meta-analysis conducted by Bunea et al. (2017), although slightly smaller effects were observed compared to studies focusing on child maltreatment only, showed that ELA was similarly associated with a blunted cortisol stress response. Thus, considering that experiencing adversity during childhood is rather the rule than the exception (Merrick et al., 2019), for the vast majority of studies, it cannot be ruled out that control participants have experienced other forms of ELA,

which in turn may have influenced the results of this systematic review. Considering these methodological shortcomings including the limitations associated with this series of meta-analyses, our ability to establish a consistent link between the experience of child maltreatment and HPA axis functioning may indeed be compromised. Finally, it should be noted that we decided to focus on peer-reviewed papers only to allow for a transparent and replicable search of the literature. Appropriate statistical methods (e.g., funnel plots and Egger's Regression tests) were applied to evaluate and control for publication bias. Nevertheless, the inclusion of grey literature could have counteracted the problem of including data that are not fully representative of the evidence as a whole.

4.3 Interpretation of the findings in the context of developmental programming of the HPA axis

Nevertheless, taking the above constraints into account, we found evidence of an altered cortisol stress response in individuals exposed to child maltreatment as compared to control participants. The null findings with regard to the other HPA axis activity measures (keeping in mind the various methodological shortcomings as one potential explanation) could also indicate that alterations causing aberrant cortisol secretion are less apparent at the level of the pituitary or adrenal glands, but rather, expressed in brain regions involved in stress processing (e.g., limbic brain areas including the hippocampus, the amygdala and the prefrontal cortex) and in the connectivity of these brain regions to the hypothalamus (Herman et al., 2003). In line with this idea, a review summarizing findings on the neuronal control of chronic stress adaptation, suggests that changes in HPA axis regulation following severe stress exposure might be traced back to long-term changes in the limbic input to neurons controlling stress responsiveness (Herman, 2013). Additionally, it is well known that limbic brain areas including the hippocampus, the amygdala and the prefrontal cortex widely express GRs, and therefore it is not surprising that acute and chronic stress appear to significantly affect synaptic physiology and connectivity in these regions (e.g., Myers et al., 2014). In contrast, structures involved primarily in the regulation of cortisol release in the context of circadian signals or following

awakening (i.e., the suprachiasmatic nucleus, Spiga et al., 2014) might be less affected. Thus, corresponding alterations in HPA axis activity measures that are not primarily activated by stress perception (e.g., DC, CAR, HCC, UFC) might only become apparent when cortisol is measured during periods of high life stress - when stress processing actually becomes relevant for these activity measures as well. Interestingly, findings of a longitudinal study evaluating stress exposure across the lifespan on HPA axis functioning at age 37 provide some support for this assumption (Young et al., 2020). In this study, in accordance with the theory of developmental programming of biological systems - the biological embedding model (e.g., Heim et al., 2019; Heindel et al., 2015) - individuals with adversities experienced during early or middle childhood showed a blunted cortisol response to a modified version of the TSST. This blunting of cortisol secretion following the perception of this stressor was independent of whether or not participants were experiencing current life stress. Additionally, similar cortisol stress response patterns were seen in participants with high and low cumulative stress, if these cumulative stress exposures did not involve early life stress (Young et al., 2020). These findings thus support the notion that when attempting to explain differences in the cortisol stress reactivity, it is not so much stress in general, but early childhood stress in particular that seems to be critical (supporting the biological embedding model). In contrast, flatter DSL profiles were only observed in those individuals who experienced ELA and were currently subjected to high levels of stress (Young et al., 2019). While these DSL results remain consistent with the biological embedding model, they also provide support for the assumption that alterations causing aberrant cortisol secretion likely relate to circuits of the brain involved in the processing of stress, and accordingly, meaningful differences in HPA axis activity measures that in terms of their activation do not per se require the experience of stress, are only to be expected when stress processing actually is involved (i.e., under high current life stress; see also Kuhlman et al., 2016).

#### 4.4 Conclusion and future directions

Taking into account all the findings and difficulties in the context of this series of meta-analyses, including: the unbalanced recruitment of study participants in the primary studies (e.g., predominantly young, female adults who belonged to an ethnic majority group and in whom child maltreatment experiences were assessed mainly through self-reports), the considerable number of studies that did not report on psychopathology, the limitations related to the assessment of child maltreatment (i.e., the use of various definitions and our inability to investigate the role of age at onset and the chronicity of the maltreatment experiences) and the various constraints related to the assessment of the various HPA axis activity measures (i.e., the inadequate control of state factors and confounding variables and limitations related to the reliability of the cortisol outcome measures), it becomes apparent that, on the one hand, a comprehensive conclusion about the functioning of the HPA axis in individuals who have been exposed to child maltreatment cannot be drawn at this time point, and on the other hand, our ability to find important moderators or relevant subgroup differences might have been limited. Nevertheless, child maltreatment appears to be associated with a blunted rather than an exaggerated activity when considering cortisol secretion following the perception of a stressor (while a tendency was also shown for HCC) and several moderators including the proportion of females in the sample, psychopathology and the study quality (to name a few) have been identified to account for some of the observed between-study heterogeneity. Considering that cortisol, when secreted in excess (e.g., during prolonged stress exposure like it is the case for child maltreatment), can have a variety of deleterious effects (Feelders et al., 2012), particularly in the brain (Sapolsky, 1999; Sapolsky et al., 2000), a corresponding downregulation may indeed serve an adaptive function protecting the body from these various adverse effects, an idea that has been subsumed under the so-called “attenuation hypothesis” (e.g., Kaess et al., 2018; Trickett et al., 2010). While probably adaptive in the first place, there is growing evidence linking not only an exaggerated, but increasingly also a blunted cortisol stress response to

various adverse behavioral and health outcomes (Carroll et al., 2017; de Rooij, 2013; Turner et al., 2020). Cortisol has various important anti-inflammatory and immune suppressive functions (Sapolsky et al., 2000) and several studies have shown that an attenuated cortisol stress reactivity (irrespective of cause) is associated with a stronger proinflammatory immune response (Buske-Kirschbaum et al., 2010; Janusek et al., 2017; Schwaiger et al., 2016). Interestingly, a growing number of studies suggest that inflammatory processes may precede the onset of, or be involved in the development of various types of mental disorders (Kivimäki et al., 2014; Melhem et al., 2017; Slavich et al., 2020). Accordingly, future studies should not only pay more attention to the potential moderating influence of current life stress, especially if interested in HPA axis activity measures that are not primarily regulated by stress perception alone, but, if interested in the consequences arising from an altered HPA axis activity, studies specifically should examine how an alerted cortisol secretion might be related to dysfunctions in other biological systems. In addition, by investigating the potentially moderating role of genes and epigenetic changes, knowledge of which individuals are most susceptible to the long-term consequences of child maltreatment (or ELA in general) may be further enhanced (e.g., Heim et al., 2019). However, reliable and reproducible results are only to be obtained if future studies more consistently rely on measurement tools that capture the assessment of various types of ELA, their onset, their chronicity and, in particular, these tools should permit the assessment of the perceived severity of the corresponding experiences. Related to a growing number of studies showing different neurobiological consequences of deprivation and threat experiences (e.g., Colich et al., 2020), a more fine-grained analysis of child maltreatment or adversity in general could further improve our understanding of the functioning of the HPA axis in individuals exposed to corresponding experiences. However, in order to obtain reliable and valid HPA axis activity measures, future studies must focus more consistently on cortisol assessment guidelines, which provide important information regarding various state and confounding variables, as well as information on the reliability of the corresponding outcome

activity measures (e.g., Adam and Kumari, 2009; Allen et al., 2017; Foley and Kirschbaum, 2010; Kudielka et al., 2012; Stalder et al., 2016; Stalder and Kirschbaum, 2012; Zänkert et al., 2019).

Journal Pre-proofs



## 5.0 Acknowledgements

We would like to thank all authors and associated research groups who responded to our requests, who suggested further studies for inclusion in our meta-analysis, who sent us additional information on their studies to clarify upon inclusion, and, in particular, we would like to thank all study authors who provided additional data for the purpose of this meta-analysis (alphabetically by last name): Alexander Nina, Ali Nida, Bechner Anna P., Bernard Kristin, Binder Elisabeth B., Burraston Bert O., Busso Daniel, Carpenter Linda, Carvalho Fernando Silvia, Cleare Tony, Cohoon Andrew J., Collip Dina, Cook Emily, Cordero Maria, Dahmen Brigitte, Diaz Angela, Duval Fabrice, Ehrenthal Johannes, Fischer Susanne, Fisher Philip A., Fogelman Nia, Fuchs Anna M., Giltay Erik J., Gonzalez Andrea, Gordis Elana B., Gunnar Megan R., Harkness Kate L., Heim Christine, Heinzel Sibylle, Helmink Fleur, Hillegers Manon H. J., Holz Nathalie, Howells Fleur, Hug Evelin, Ivanov Iliyan, Janusek Linda, Kaess Michael, Kaiser Roselinde H., Kathleen Brewer-Smyt, Kellner Michael, Klaassens Ellen, Kuhlman Kate R., Kumsta Robert, Kuzminskaite Erika, Lee Royce J., Linares Oriana, Lindley Steven E., Lovallo William R., Luecken Linda J., MacMillan Harreit, Malaspina Dolores, Mansueto Giovanni, Martinson Amber, Martorell Gabi, Mathews Herbert L., Mehta Divya, Melhem Nadine M., Mello Andrea F., Mesman Esther, Meston Cindy M., Monteleone Palmiero, Morris Matthew C., Negriff Sonya, Nelson Elizabeth, Noorden Martijn, O'Connor Daryl B., Otte Christian, Ouellet-Morin Isabelle, Peckins Melissa K., Pektus Andrew J., Penninx Brenda W. J. H., Phassouliotis Christina, Pizzagalli Diego A., Pollak Seth D., Pruessner Jens C., Putnam Frank, Rao Uma, Reichl Corinna, Reijnierse Esmee M., Schächinger Hartmut, Schalinski Inga, Schechter Daniel, Schindler Lena, Shenk Chad E., Shirtcliff Elizabeth, Smeets Tom, Stein Murray B., Steudte-Schmiedgen Susann, Sullivan Margaret, Sunnqvist Charlotta, Tiemeier Henning, Tyrka Audrey, Valentino Kristin, van den Brink Wim, van Winkel Ruud, van Zuiden Mirjam, Watson Stuart, White Lars O. (AMIS project). Finally, we would like to thank our

research assistant Arumaithurai Ajanigha and our biostatistician Lerch Stefan for their great help in the data extraction process.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Declarations of Interest: None**

Journal Pre-proofs

## 6.0 References

References with an asterisk ('\*') were included in this series of meta-analyses

Adam, E.K., Kumari, M., 2009. Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology* 34, 1423–1436.

<https://doi.org/10.1016/j.psyneuen.2009.06.011>

Adam, E.K., Quinn, M.E., Tavernier, R., McQuillan, M.T., Dahlke, K.A., Gilbert, K.E., 2017. Diurnal cortisol slopes and mental and physical health outcomes: A systematic review and meta-analysis. *Psychoneuroendocrinology* 83, 25–41.

<https://doi.org/10.1016/j.psyneuen.2017.05.018>

\* Alexander, N., Kirschbaum, C., Wankerl, M., Stauch, B.J., Stalder, T., Steudte-Schmiedgen, S., Muehlhan, M., Miller, R., 2018. Glucocorticoid receptor gene methylation moderates the association of childhood trauma and cortisol stress reactivity.

*Psychoneuroendocrinology* 90, 68–75. <https://doi.org/10.1016/j.psyneuen.2018.01.020>

\* Ali, N., Pruessner, J.C., 2012. The salivary alpha amylase over cortisol ratio as a marker to assess dysregulations of the stress systems. *Physiology & Behavior* 106, 65–72.

<https://doi.org/10.1016/j.physbeh.2011.10.003>

Allen, A.P., Kennedy, P.J., Dockray, S., Cryan, J.F., Dinan, T.G., Clarke, G., 2017. The Trier Social Stress Test: Principles and practice. *Neurobiology of Stress* 6, 113–126.

<https://doi.org/10.1016/j.ynstr.2016.11.001>

\* Baes, C. von W., Martins, C.M.S., Tofoli, S.M. de C., Juruena, M. F., 2014. Early Life Stress in Depressive Patients: HPA Axis Response to GR and MR Agonist. *Frontiers in Psychiatry* 5. <https://doi.org/10.3389/fpsy.2014.00002>

Berger, M., Kraeuter, A.K., Romanik, D., Malouf, P., Amminger, G.P., Sarnyai, Z., 2016.

Cortisol awakening response in patients with psychosis: Systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews* 68, 157–166.

\* Bernard, K., Butzin-Dozier, Z., Rittenhouse, J., Dozier, M., 2010. Cortisol Production

- Patterns in Young Children Living With Birth Parents vs Children Placed in Foster Care Following Involvement of Child Protective Services. *Archives of Pediatrics & Adolescent Medicine* 164. <https://doi.org/10.1001/archpediatrics.2010.54>
- Bernard, K., Frost, A., Bennett, C.B., Lindhiem, O., 2017. Maltreatment and diurnal cortisol regulation: A meta-analysis. *Psychoneuroendocrinology* 78, 57–67. <https://doi.org/10.1016/j.psyneuen.2017.01.005>
- \* Bicanic, I.A.E., Postma, R.M., Sinnema, G., De Roos, C., Olf, M., Van Wesel, F., Van de Putte, E. M., 2013. Salivary cortisol and dehydroepiandrosterone sulfate in adolescent rape victims with post traumatic stress disorder. *Psychoneuroendocrinology* 38, 408–415. <https://doi.org/10.1016/j.psyneuen.2012.06.015>
- Boileau, K., Barbeau, K., Sharma, R., Bielajew, C., 2019. Ethnic differences in diurnal cortisol profiles in healthy adults: A meta-analysis. *British Journal of Health Psychology* 24, 806–827. <https://doi.org/10.1111/bjhp.12380>
- Bonoldi, I., Simeone, E., Rocchetti, M., Codjoe, L., Rossi, G., Gambi, F., Balottin, U., Caverzasi, E., Politi, P., Fusar-Poli, P., 2013. Prevalence of self-reported childhood abuse in psychosis: A meta-analysis of retrospective studies. *Psychiatry Research* 210, 8–15. <https://doi.org/10.1016/j.psychres.2013.05.003>
- Bremner, J.D., Vermetten, E., Mazure, C.M., 2000. Development and preliminary psychometric properties of an instrument for the measurement of childhood trauma: The Early Trauma Inventory. *Depression and Anxiety* 12, 1–12. [https://doi.org/10.1002/1520-6394\(2000\)12:1<1::AID-DA1>3.0.CO;2-W](https://doi.org/10.1002/1520-6394(2000)12:1<1::AID-DA1>3.0.CO;2-W)
- \* Buchmann, A.F., Holz, N., Boecker, R., Blomeyer, D., Rietschel, M., Witt, S.H., Schmidt, M.H., Esser, G., Banaschewski, T., Brandeis, D., Zimmermann, U.S., Laucht, M., 2014. Moderating role of FKBP5 genotype in the impact of childhood adversity on cortisol stress response during adulthood. *European Neuropsychopharmacology* 24, 837–845. <https://doi.org/10.1016/j.euroneuro.2013.12.001>

- Bunea, I.M., Szentagotai-Tatar, A., Miu, A.C., 2017. Early-life adversity and cortisol response to social stress: A meta-analysis. *Translational Psychiatry* 7, 1274. <https://doi.org/10.1038/s41398-017-0032-3>
- Buske-Kirschbaum, A., Ebrect, M., Hellhammer, D.H., 2010. Blunted HPA axis responsiveness to stress in atopic patients is associated with the acuity and severeness of allergic inflammation. *Brain, Behavior, and Immunity* 24, 1347–1353. <https://doi.org/10.1016/j.bbi.2010.06.013>
- \* Carpenter, L.L., Carvalho, J.P., Tyrka, A.R., Wier, L.M., Mello, A.F., Mello, M.F., Anderson, G.M., Wilkinson, C.W., Price, L.H., 2007. Decreased Adrenocorticotrophic Hormone and Cortisol Responses to Stress in Healthy Adults Reporting Significant Childhood Maltreatment. *Biological Psychiatry* 62, 1080–1087. <https://doi.org/10.1016/j.biopsych.2007.05.002>
- \* Carpenter, L.L., Shattuck, T.T., Tyrka, A.R., Geraciotti, T.D., Price, L.H., 2011. Effect of childhood physical abuse on cortisol stress response. *Psychopharmacology* 214, 367–375. <https://doi.org/10.1007/s00213-010-2007-4>
- \* Carpenter, L.L., Tyrka, A.R., Ross, N.S., Khoury, L., Anderson, G.M., Price, L.H., 2009. Effect of Childhood Emotional Abuse and Age on Cortisol Responsivity in Adulthood. *Biological Psychiatry* 66, 69–75. <https://doi.org/10.1016/j.biopsych.2009.02.030>
- Carr, C.P., Martins, C.M.S., Stingel, A.M., Lemgruber, V.B., Juruena, M.F., 2013. The Role of Early Life Stress in Adult Psychiatric Disorders: A Systematic Review According to Childhood Trauma Subtypes. *The Journal of Nervous and Mental Disease* 201, 1007–1020. <https://doi.org/10.1097/NMD.0000000000000049>
- \* Carrion, V.G., Weems, C.F., Ray, R.D., Glaser, B., Hessel, D., Reiss, A.L., 2002. Diurnal salivary cortisol in pediatric posttraumatic stress disorder. *Biological Psychiatry* 51, 575–582. [https://doi.org/10.1016/S0006-3223\(01\)01310-5](https://doi.org/10.1016/S0006-3223(01)01310-5)
- Carroll, B.J., 1981. A Specific Laboratory Test for the Diagnosis of Melancholia:

- Standardization, Validation, and Clinical Utility. *Archives of General Psychiatry* 38, 15. <https://doi.org/10.1001/archpsyc.1981.01780260017001>
- Carroll, D., Ginty, A.T., Whittaker, A.C., Lovallo, W.R., de Rooij, S.R., 2017. The behavioural, cognitive, and neural corollaries of blunted cardiovascular and cortisol reactions to acute psychological stress. *Neuroscience & Biobehavioral Reviews* 77, 74–86. <https://doi.org/10.1016/j.neubiorev.2017.02.025>
- \* Carvalho Fernando, S., Beblo, T., Schlosser, N., Terfehr, K., Otte, C., Löwe, B., Wolf, O.T., Spitzer, C., Driessen, M., Wingenfeld, K., 2012. Associations of childhood trauma with hypothalamic-pituitary-adrenal function in borderline personality disorder and major depression. *Psychoneuroendocrinology* 37, 1659–1668. <https://doi.org/10.1016/j.psyneuen.2012.02.012>
- Chen, Y., Baram, T.Z., 2016. Toward Understanding How Early-Life Stress Reprograms Cognitive and Emotional Brain Networks. *Neuropsychopharmacology* 41, 197–206. <https://doi.org/10.1038/npp.2015.181>
- Chida, Y., Steptoe, A., 2009. Cortisol awakening response and psychosocial factors: A systematic review and meta-analysis. *Biological Psychology* 80, 265–278. <https://doi.org/10.1016/j.biopsycho.2008.10.004>
- \* Cicchetti, D., Rogosch, F.A., Gunnar, M.R., Toth, S.L., 2010. The Differential Impacts of Early Physical and Sexual Abuse and Internalizing Problems on Daytime Cortisol Rhythm in School-Aged Children. *Child Development* 81, 252–269. <https://doi.org/10.1111/j.1467-8624.2009.01393.x>
- Cicchetti, D., & Toth, S.L., 2005. Child Maltreatment. *Annual Review of Clinical Psychology* 1, 409–438. <https://doi.org/10.1146/annurev.clinpsy.1.102803.144029>
- Clark, C., Caldwell, T., Power, C., Stansfeld, S. A., 2010. Does the Influence of Childhood Adversity on Psychopathology Persist Across the Lifecourse? A 45-Year Prospective Epidemiologic Study. *Annals of Epidemiology* 20, 385–394.

<https://doi.org/10.1016/j.annepidem.2010.02.008>

Clemens, V., Huber-Lang, M., Plener, P.L., Brähler, E., Brown, R.C., Fegert, J.M., 2018.

Association of child maltreatment subtypes and long-term physical health in a German representative sample. *European Journal of Psychotraumatology* 9, 1510278.

<https://doi.org/10.1080/20008198.2018.1510278>

Clow, A., Hucklebridge, F., Stalder, T., Evans, P., Thorn, L., 2010. The cortisol awakening response: More than a measure of HPA axis function. *Neuroscience & Biobehavioral Reviews* 35, 97–103. <https://doi.org/10.1016/j.neubiorev.2009.12.011>

Colich, N.L., Rosen, M.L., Williams, E.S., McLaughlin, K.A., 2020. Biological aging in childhood and adolescence following experiences of threat and deprivation: A systematic review and meta-analysis. *Psychological Bulletin* 146, 721–764.

<https://doi.org/10.1037/bul0000270>

\* Cook, E.C., Chaplin, T.M., Sinha, R., Tebes, J.K., Mayes, L.C., 2012. The Stress Response and Adolescents' Adjustment: The Impact of Child Maltreatment. *Journal of Youth and Adolescence* 41, 1067–1077. <https://doi.org/10.1007/s10964-012-9746-y>

Cowell, R.A., Cicchetti, D., Rogosch, F.A., & Toth, S.L., 2015. Childhood maltreatment and its effect on neurocognitive functioning: Timing and chronicity matter. *Development and Psychopathology* 27, 521–533. <https://doi.org/10.1017/S0954579415000139>

Danese, A., Moffitt, T.E., Harrington, H., Milne, B.J., Polanczyk, G., Pariante, C.M., Poulton, R., Caspi, A., 2009. Adverse Childhood Experiences and Adult Risk Factors for Age-Related Disease: Depression, Inflammation, and Clustering of Metabolic Risk Markers. *Arch Pediatr Adolesc Med* 163.

<https://doi.org/10.1001/archpediatrics.2009.214>

\* De Bellis, M.D., Baum, A.S., Birmaher, B., Keshavan, M.S., Eccard, C.H., Boring, A.M., Jenkins, F.J., Ryan, N.D., 1999. A.E. Bennett Research Award. Developmental traumatology. Part I: Biological stress systems. *Biol Psychiatry* 45, 1259–1270.

[https://doi.org/10.1016/s0006-3223\(99\)00044-x](https://doi.org/10.1016/s0006-3223(99)00044-x)

\* De Bellis, M.D., Chrousos, G.P., Dorn, L.D., Burke, L., Helmers, K., Kling, M.A., Trickett, P.K., Putnam, F.W., 1994. Hypothalamic-pituitary-adrenal axis dysregulation in sexually abused girls. *J Clin Endocrinol Metab* 78, 249–255.

<https://doi.org/10.1210/jcem.78.2.8106608>

de Rooij, S.R., 2013. Blunted cardiovascular and cortisol reactivity to acute psychological stress: A summary of results from the Dutch Famine Birth Cohort Study. *International Journal of Psychophysiology* 90, 21–27.

<https://doi.org/10.1016/j.ijpsycho.2012.09.011>

Deutschbein, T., Broecker-Preuss, M., Hartmann, M., Althoff, R., Wudy, S., Mann, K., Petersenn, S., 2011. Measurement of Urinary Free Cortisol by Current Immunoassays: Need for Sex-dependent Reference Ranges to Define Hypercortisolism. *Horm Metab Res* 43, 714–719. <https://doi.org/10.1055/s-0031-1286297>

Dickerson, S.S., Kemeny, M.E., 2004. Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychological Bulletin* 130, 355–391. <https://doi.org/10.1037/0033-2909.130.3.355>

\* do Prado, C.H., Grassi-Oliveira, R., Daruy-Filho, L., Wieck, A., Bauer, M. E., 2017.

Evidence for Immune Activation and Resistance to Glucocorticoids Following Childhood Maltreatment in Adolescents Without Psychopathology.

*Neuropsychopharmacology* 42, 2272–2282. <https://doi.org/10.1038/npp.2017.137>

Drahota, A., Beller, E. (n.d.). RevMan Calculator.

<https://training.cochrane.org/resource/revman-calculator> (accessed 25 January 2021).

Drews, E., Fertuck, E.A., Koenig, J., Kaess, M., Arntz, A., 2019. Hypothalamic-pituitary-adrenal axis functioning in borderline personality disorder: A meta-analysis.

*Neuroscience & Biobehavioral Reviews* 96, 316–334.

<https://doi.org/10.1016/j.neubiorev.2018.11.008>



- Dube, S.R., Anda, R.F., Felitti, V.J., Chapman, D.P., Williamson, D.F., Giles, W.H., 2001. Childhood Abuse, Household Dysfunction, and the Risk of Attempted Suicide Throughout the Life Span: Findings From the Adverse Childhood Experiences Study. *JAMA* 286, 3089. <https://doi.org/10.1001/jama.286.24.3089>
- \* Duval, F., Crocq, M.-A., Guillon, M.-S., Mokrani, M.-C., Monreal, J., Bailey, P., Macher, J.-P., 2004. Increased adrenocorticotropin suppression following dexamethasone administration in sexually abused adolescents with posttraumatic stress disorder. *Psychoneuroendocrinology* 29, 1281–1289. <https://doi.org/10.1016/j.psyneuen.2004.03.006>
- Egger, M., Smith, G.D., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315, 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- \* England-Mason, G., Kimber, M., Khoury, J., Atkinson, L., MacMillan, H., Gonzalez, A., 2017. Difficulties with emotion regulation moderate the association between childhood history of maltreatment and cortisol reactivity to psychosocial challenge in postpartum women. *Hormones and Behavior* 95, 44–56. <https://doi.org/10.1016/j.yhbeh.2017.07.007>
- Fan, B., Ma, J., Zhang, H., Liao, Y., Wang, W., Zhang, S., Lu, C., Guo, L., 2021. Association of FKBP5 gene variants with depression susceptibility: A comprehensive meta-analysis. *Asia-Pacific Psychiatry* 13, e12464. <https://doi.org/10.1111/appy.12464>
- \* Faravelli, C., Amedei, S.G., Rotella, F., Faravelli, L., Palla, A., Consoli, G., Ricca, V., Batini, S., Sauro, C.L., Spiti, A., dell’Osso, M.C., 2010. Childhood traumata, Dexamethasone Suppression Test and psychiatric symptoms: a trans-diagnostic approach. *Psychol. Med.* 40, 2037–2048. <https://doi.org/10.1017/S0033291710000115>
- \* Faravelli, C., Mansueto, G., Palmieri, S., Lo Sauro, C., Rotella, F., Pietrini, F., Fioravanti, G., 2017. Childhood Adversity, Cortisol Levels, and Psychosis: A Retrospective Investigation. *The Journal of Nervous and Mental Disease* 205, 574–579.

<https://doi.org/10.1097/NMD.0000000000000699>

Feelders, R.A., Pulgar, S.J., Kempel, A., Pereira, A.M., 2012. MANAGEMENT OF ENDOCRINE DISEASE: The burden of Cushing's disease: clinical and health-related quality of life aspects. *European Journal of Endocrinology* 167, 311–326.

<https://doi.org/10.1530/EJE-11-1095>

\* Fischer, S., Duncko, R., Hatch, S.L., Papadopoulos, A., Goodwin, L., Frissa, S., Hotopf, M., Cleare, A.J., 2017. Sociodemographic, lifestyle, and psychosocial determinants of hair cortisol in a South London community sample. *Psychoneuroendocrinology* 76, 144–153. <https://doi.org/10.1016/j.psyneuen.2016.11.011>

\* Fisher, P.A., Stoolmiller, M., Gunnar, M.R., Burraston, B.O., 2007. Effects of a therapeutic intervention for foster preschoolers on diurnal cortisol activity.

*Psychoneuroendocrinology* 32, 892–905.

<https://doi.org/10.1016/j.psyneuen.2007.06.008>

Fogelman, N., Canli, T., 2018. Early life stress and cortisol: A meta-analysis. *Hormones and Behavior* 98, 63–76. <https://doi.org/10.1016/j.yhbeh.2017.12.014>

\* Fogelman, N., Mikhailik, A., Mueller-Alcazar, A., Bernard, K., Canli, T., 2016. Stressing over anxiety: A novel interaction of 5-HTTPLR genotype and anxiety-related phenotypes in older adults. *Psychoneuroendocrinology* 71, 36–42.

<https://doi.org/10.1016/j.psyneuen.2016.05.012>

Foley, P., Kirschbaum, C., 2010. Human hypothalamus–pituitary–adrenal axis responses to acute psychosocial stress in laboratory settings. *Neuroscience & Biobehavioral*

*Reviews* 35, 91–96. <https://doi.org/10.1016/j.neubiorev.2010.01.010>

Fries, E., Dettenborn, L., Kirschbaum, C., 2009. The cortisol awakening response (CAR): Facts and future directions. *International Journal of Psychophysiology* 72, 67–73.

<https://doi.org/10.1016/j.ijpsycho.2008.03.014>

\* Fuchs, A., Moehler, E., Resch, F., Kaess, M., 2017. The effect of a maternal history of

- childhood abuse on adrenocortical attunement in mothers and their toddlers. *Dev Psychobiol* 59, 639–652. <https://doi.org/10.1002/dev.21531>
- Gilbert, R., Widom, C.S., Browne, K., Fergusson, D., Webb, E., Janson, S., 2009. Burden and consequences of child maltreatment in high-income countries. *The Lancet* 373, 68–81. [https://doi.org/10.1016/S0140-6736\(08\)61706-7](https://doi.org/10.1016/S0140-6736(08)61706-7)
- Gold, P.W., Loriaux, D.L., Roy, A., Kling, M.A., Calabrese, J.R., Kellner, C.H., Nieman, L.K., Post, R.M., Pickar, D., Gallucci, W., Avgerinos, P., Paul, S., Oldfield, E.H., Cutler, G.B., Chrousos, G.P., 1986. Responses to Corticotropin-Releasing Hormone in the Hypercortisolism of Depression and Cushing's Disease. *N Engl J Med* 314, 1329–1335. <https://doi.org/10.1056/NEJM198605223142101>
- Granger, D.A., Hibel, L.C., Fortunato, C.K., Kapelewski, C.H., 2009. Medication effects on salivary cortisol: Tactics and strategy to minimize impact in behavioral and developmental science. *Psychoneuroendocrinology* 34, 1437–1448. <https://doi.org/10.1016/j.psyneuen.2009.06.017>
- \* Groër, M.W., Kostas-Polston, E.A., Dillahun-Aspillaga, C., Beckie, T.M., Johnson-Mallard, V., Duffy, A., Evans, M.E., 2016. Allostatic Perspectives in Women Veterans With a History of Childhood Sexual Assault. *Biological Research For Nursing* 18, 454–464. <https://doi.org/10.1177/1099800416638442>
- \* Harkness, K.L., Stewart, J.G., Wynne-Edwards, K.E., 2011. Cortisol reactivity to social stress in adolescents: Role of depression severity and child maltreatment. *Psychoneuroendocrinology* 36, 173–181. <https://doi.org/10.1016/j.psyneuen.2010.07.006>
- Harrer, M., Cuijpers, P., Ebert, D., 2019. Doing Meta-Analysis in R. <https://doi.org/10.5281/ZENODO.2551803>
- Hedges, L.V., 1982. Estimation of effect size from a series of independent experiments. *Psychological Bulletin* 92, 490–499. <https://doi.org/10.1037/0033-2909.92.2.490>

- \* Heim, C., Mletzko, T., Purselle, D., Musselman, D.L., Nemeroff, C.B., 2008. The dexamethasone/corticotropin-releasing factor test in men with major depression: role of childhood trauma. *Biol Psychiatry* 63, 398–405.  
<https://doi.org/10.1016/j.biopsych.2007.07.002>
- \* Heim, C., Newport, D.J., Heit, S., Graham, Y.P., Wilcox, M., Bonsall, R., Miller, A.H., Nemeroff, C.B., 2000. Pituitary-adrenal and autonomic responses to stress in women after sexual and physical abuse in childhood. *JAMA* 284, 592–597.  
<https://doi.org/10.1001/jama.284.5.592>
- Heim, C., Entringer, S., Buss, C., 2019. Translating basic research knowledge on the biological embedding of early-life stress into novel approaches for the developmental programming of lifelong health. *Psychoneuroendocrinology* 105, 123–137.  
<https://doi.org/10.1016/j.psyneuen.2018.12.011>
- Heindel, J.J., Balbus, J., Birnbaum, L., Brune-Drisse, M.N., Grandjean, P., Gray, K., Landrigan, P.J., Sly, P.D., Suk, W., Slechta, D.C., Thompson, C., Hanson, M., 2015. Developmental Origins of Health and Disease: Integrating Environmental Influences. *Endocrinology* 156, 3416–3421. <https://doi.org/10.1210/en.2015-1394>
- Hellhammer, J., Fries, E., Schweisthal, O.W., Schlotz, W., Stone, A.A., Hagemann, D., 2007. Several daily measurements are necessary to reliably assess the cortisol rise after awakening: State- and trait components. *Psychoneuroendocrinology* 32, 80–86.  
<https://doi.org/10.1016/j.psyneuen.2006.10.005>
- Hemmingsson, E., Johansson, K., Reynisdottir, S., 2014. Effects of childhood abuse on adult obesity: a systematic review and meta-analysis: Childhood abuse and adult obesity. *Obesity Reviews* 15, 882–893. <https://doi.org/10.1111/obr.12216>
- \* Hengesch, X., Elwenspoek, M.M.C., Schaan, V.K., Larra, M.F., Finke, J.B., Zhang, X., Bachmann, P., Turner, J.D., Vögele, C., Muller, C.P., Schächinger, H., 2018. Blunted endocrine response to a combined physical-cognitive stressor in adults with early life

- adversity. *Child Abuse & Neglect* 85, 137–144.  
<https://doi.org/10.1016/j.chiabu.2018.04.002>
- Herman, J.P., 2013. Neural control of chronic stress adaptation. *Front. Behav. Neurosci.* 7.  
<https://doi.org/10.3389/fnbeh.2013.00061>
- Herman, J.P., Figueiredo, H., Mueller, N.K., Ulrich-Lai, Y., Ostrander, M.M., Choi, D.C., Cullinan, W.E., 2003. Central mechanisms of stress integration: hierarchical circuitry controlling hypothalamo–pituitary–adrenocortical responsiveness. *Frontiers in Neuroendocrinology* 24, 151–180. <https://doi.org/10.1016/j.yfrne.2003.07.001>
- Herrenkohl, R.C., Herrenkohl, T.I., 2009. Assessing a Child’s Experience of Multiple Maltreatment Types: Some Unfinished Business. *J Fam Viol* 24, 485–496.  
<https://doi.org/10.1007/s10896-009-9247-2>
- \* Hibel, L.C., Mercado, E., Valentino, K., 2019. Child Maltreatment and Mother–Child Transmission of Stress Physiology. *Child Maltreat* 24, 340–352.  
<https://doi.org/10.1177/1077559519826295>
- Higgins, J.P.T., 2003. Measuring inconsistency in meta-analyses. *BMJ* 327, 557–560.  
<https://doi.org/10.1136/bmj.327.7414.557>
- Higgins, J.P.T., Thompson, S.G., 2002. Quantifying heterogeneity in a meta-analysis. *Statist. Med.* 21, 1539–1558. <https://doi.org/10.1002/sim.1186>
- Hoffmann, A., Spengler, D., 2012. The lasting legacy of social stress on the epigenome of the hypothalamic–pituitary–adrenal axis. *Epigenomics* 4, 431–444.  
<https://doi.org/10.2217/epi.12.34>
- Hughes, K., Bellis, M.A., Hardcastle, K.A., Sethi, D., Butchart, A., Mikton, C., Jones, L., Dunne, M.P., 2017. The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. *The Lancet Public Health* 2, e356–e366.  
[https://doi.org/10.1016/S2468-2667\(17\)30118-4](https://doi.org/10.1016/S2468-2667(17)30118-4)
- Infurna, M.R., Reichl, C., Parzer, P., Schimmenti, A., Bifulco, A., Kaess, M., 2016.

- Associations between depression and specific childhood experiences of abuse and neglect: A meta-analysis. *Journal of Affective Disorders* 190, 47–55.  
<https://doi.org/10.1016/j.jad.2015.09.006>
- IntHout, J., Ioannidis, J.P.A., Rovers, M.M., Goeman, J.J., 2016. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* 6, e010247.  
<https://doi.org/10.1136/bmjopen-2015-010247>
- Ising, M., Depping, A.-M., Siebertz, A., Lucae, S., Unschuld, P.G., Kloiber, S., Horstmann, S., Uhr, M., Müller-Myhsok, B., Holsboer, F., 2008. Polymorphisms in the FKBP5 gene region modulate recovery from psychosocial stress in healthy controls. *European Journal of Neuroscience* 28, 389–398. <https://doi.org/10.1111/j.1460-9568.2008.06332.x>
- \* Ivanov, I., Yehuda, R., Greenblatt, E., Davidow, J., Makotkine, I., Alfi, L., Newcorn, J.H., 2011. The effect of trauma on stress reactivity in aggressive youth. *Psychiatry Research* 189, 396–402. <https://doi.org/10.1016/j.psychres.2011.05.046>
- Jackson, Y., McGuire, A., Tunno, A.M., Makanui, P.K., 2019. A reasonably large review of operationalization in child maltreatment research: Assessment approaches and sources of information in youth samples. *Child Abuse & Neglect* 87, 5–17.  
<https://doi.org/10.1016/j.chiabu.2018.09.016>
- Jacoby, N., Overfeld, J., Binder, E.B., Heim, C.M., 2016. Stress Neurobiology and Developmental Psychopathology, in: Cicchetti, D. (Ed.), *Developmental Psychopathology*. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 1–45.  
<https://doi.org/10.1002/9781119125556.devpsy221>
- Jaffee, S.R., Maikovich-Fong, A.K., 2011. Effects of chronic maltreatment and maltreatment timing on children's behavior and cognitive abilities: Effects of chronic maltreatment. *Journal of Child Psychology and Psychiatry* 52, 184–194.  
<https://doi.org/10.1111/j.1469-7610.2010.02304.x>

- Janusek, L.W., Tell, D., Gaylord-Harden, N., Mathews, H.L., 2017. Relationship of childhood adversity and neighborhood violence to a proinflammatory phenotype in emerging adult African American men: An epigenetic link. *Brain, Behavior, and Immunity* 60, 126–135. <https://doi.org/10.1016/j.bbi.2016.10.006>
- Kaess, M., Whittle, S., O'Brien-Simpson, L., Allen, N.B., Simmons, J.G., 2018. Childhood maltreatment, pituitary volume and adolescent hypothalamic-pituitary-adrenal axis – Evidence for a maltreatment-related attenuation. *Psychoneuroendocrinology* 98, 39–45. <https://doi.org/10.1016/j.psyneuen.2018.08.004>
- \* Kaess, M., Whittle, S., Simmons, J.G., Jovev, M., Allen, N.B., Chanen, A.M., 2017. The Interaction of Childhood Maltreatment, Sex, and Borderline Personality Features in the Prediction of the Cortisol Awakening Response in Adolescents. *Psychopathology* 50, 188–194. <https://doi.org/10.1159/000456549>
- \* Kaiser, R.H., Clegg, R., Goer, F., Pechtel, P., Beltzer, M., Vitaliano, G., Olson, D.P., Teicher, M.H., Pizzagalli, D.A., 2018. Childhood stress, grown-up brain networks: corticolimbic correlates of threat-related early life stress and adult stress response. *Psychol. Med.* 48, 1157–1166. <https://doi.org/10.1017/S0033291717002628>
- Kaplow, J.B., Widom, C.S., 2007. Age of onset of child maltreatment predicts long-term mental health outcomes. *Journal of Abnormal Psychology* 116, 176–187. <https://doi.org/10.1037/0021-843X.116.1.176>
- \* Kaufman, J., Birmaher, B., Perel, J., Dahl, R.E., Moreci, P., Nelson, B., Wells, W., Ryan, N.D., 1997. The corticotropin-releasing hormone challenge in depressed abused, depressed nonabused, and normal control children. *Biol Psychiatry* 42, 669–679. [https://doi.org/10.1016/s0006-3223\(96\)00470-2](https://doi.org/10.1016/s0006-3223(96)00470-2)
- \* Keeshin, B.R., Strawn, J.R., Out, D., Granger, D.A., Putnam, F.W., 2014. Cortisol awakening response in adolescents with acute sexual abuse related posttraumatic stress disorder. *Depress Anxiety* 31, 107–114. <https://doi.org/10.1002/da.22154>

- Kendler, K.S., Bulik, C.M., Silberg, J., Hettema, J.M., Myers, J., Prescott, C.A., 2000. Childhood Sexual Abuse and Adult Psychiatric and Substance Use Disorders in Women: An Epidemiological and Cotwin Control Analysis. *Arch Gen Psychiatry* 57, 953. <https://doi.org/10.1001/archpsyc.57.10.953>
- Khoury, J.E., Gonzalez, A., Levitan, R.D., Pruessner, J.C., Chopra, K., Basile, V.S., Masellis, M., Goodwill, A., Atkinson, L., 2015. Summary cortisol reactivity indicators: Interrelations and meaning. *Neurobiology of Stress* 2, 34–43. <https://doi.org/10.1016/j.ynstr.2015.04.002>
- Kivimäki, M., Shipley, M.J., Batty, G.D., Hamer, M., Akbaraly, T.N., Kumari, M., Jokela, M., Virtanen, M., Lowe, G.D., Ebmeier, K.P., Brunner, E.J., Singh-Manoux, A., 2014. Long-term inflammation increases risk of common mental disorder: a cohort study. *Mol Psychiatry* 19, 149–150. <https://doi.org/10.1038/mp.2013.35>
- \* Klaassens, E.R., van Noorden, M.S., Giltay, E.J., van Pelt, J., van Veen, T., Zitman, F.G., 2009. Effects of childhood trauma on HPA-axis reactivity in women free of lifetime psychopathology. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 33, 889–894. <https://doi.org/10.1016/j.pnpbp.2009.04.011>
- \* Klaus, K., Butler, K., Gutierrez, H., Durrant, S.J., Pennington, K., 2018. Interactive effects of early life stress and CACNA1C genotype on cortisol awakening response. *Biological Psychology* 136, 22–28. <https://doi.org/10.1016/j.biopsycho.2018.05.002>
- Klengel, T., Mehta, D., Anacker, C., Rex-Haffner, M., Pruessner, J.C., Pariante, C.M., Pace, T.W.W., Mercer, K.B., Mayberg, H.S., Bradley, B., Nemeroff, C.B., Holsboer, F., Heim, C.M., Ressler, K.J., Rein, T., Binder, E.B., 2013. Allele-specific FKBP5 DNA demethylation mediates gene–childhood trauma interactions. *Nat Neurosci* 16, 33–41. <https://doi.org/10.1038/nn.3275>
- Koss, K.J., Gunnar, M.R., 2018. Annual Research Review: Early adversity, the hypothalamic-pituitary-adrenocortical axis, and child psychopathology. *Journal of Child Psychology*



- and *Psychiatry* 59, 327–346. <https://doi.org/10.1111/jcpp.12784>
- Kudielka, B.M., Gierens, A., Hellhammer, D.H., Wüst, S., Schlotz, W., 2012. Salivary Cortisol in Ambulatory Assessment—Some Dos, Some Don'ts, and Some Open Questions: *Psychosomatic Medicine* 74, 418–431.  
<https://doi.org/10.1097/PSY.0b013e31825434c7>
- Kudielka, B.M., Wüst, S., 2010. Human models in acute and chronic stress: Assessing determinants of individual hypothalamus–pituitary–adrenal axis activity and reactivity. *Stress* 13, 1–14. <https://doi.org/10.3109/10253890902874913>
- \* Kuhlman, K.R., Geiss, E.G., Vargas, I., Lopez-Duran, N.L., 2015. Differential associations between childhood trauma subtypes and adolescent HPA-axis functioning. *Psychoneuroendocrinology* 54, 103–114.  
<https://doi.org/10.1016/j.psyneuen.2015.01.020>
- Kuhlman, K.R., Repetti, R.L., Reynolds, B.M., Robles, T.F., 2016. Change in parent-child conflict and the HPA-axis: Where should we be looking and for how long? *Psychoneuroendocrinology* 68, 74–81. <https://doi.org/10.1016/j.psyneuen.2016.02.029>
- \* Kumsta, R., Schlotz, W., Golm, D., Moser, D., Kennedy, M., Knights, N., Kreppner, J., Maughan, B., Rutter, M., Sonuga-Barke, E., 2017. HPA axis dysregulation in adult adoptees twenty years after severe institutional deprivation in childhood. *Psychoneuroendocrinology* 86, 196–202.  
<https://doi.org/10.1016/j.psyneuen.2017.09.021>
- \* Kuras, Y.I., Assaf, N., Thoma, M.V., Gianferante, D., Hanlin, L., Chen, X., Fiksdal, A., Rohleder, N., 2017. Blunted Diurnal Cortisol Activity in Healthy Adults with Childhood Adversity. *Front. Hum. Neurosci.* 11, 574.  
<https://doi.org/10.3389/fnhum.2017.00574>
- \* Lee, R.J., Hempel, J., TenHarmsel, A., Liu, T., Mathé, A.A., Klock, A., 2012. The neuroendocrinology of childhood trauma in personality disorder.

- Psychoneuroendocrinology 37, 78–86. <https://doi.org/10.1016/j.psyneuen.2011.05.006>
- Leeb, R., Paulozzi, L., Melanson, C., Simon, T., Arias-Mayenco, I., Leeb, R.T., 2008. Child maltreatment surveillance: Uniform definitions for public health and recommended data elements, version 1.0.
- Leistner, C., Menke, A., 2018. How to measure glucocorticoid receptor's sensitivity in patients with stress-related psychiatric disorders. *Psychoneuroendocrinology* 91, 235–260. <https://doi.org/10.1016/j.psyneuen.2018.01.023>
- \* Lemieux, A., Coe, C.L., Carnes, M., 2008. Symptom severity predicts degree of T cell activation in adult women following childhood maltreatment. *Brain, Behavior, and Immunity* 22, 994–1003. <https://doi.org/10.1016/j.bbi.2008.02.005>
- \* Lemieux, A.M., Coe, C.L., 1995. Abuse-related posttraumatic stress disorder: evidence for chronic neuroendocrine activation in women. *Psychosom Med* 57, 105–115. <https://doi.org/10.1097/00006842-199503000-00002>
- \* Li, L., Chassan, R.A., Bruer, E.H., Gower, B.A., Shelton, R.C., 2015. Childhood maltreatment increases the risk for visceral obesity: Childhood Maltreatment and Visceral Obesity. *Obesity* 23, 1625–1632. <https://doi.org/10.1002/oby.21143>
- Li, S.H., Graham, B.M., 2017. Why are women so vulnerable to anxiety, trauma-related and stress-related disorders? The potential role of sex hormones. *The Lancet Psychiatry* 4, 73–82. [https://doi.org/10.1016/S2215-0366\(16\)30358-3](https://doi.org/10.1016/S2215-0366(16)30358-3)
- \* Lindley, S.E., Carlson, E.B., Benoit, M., 2004. Basal and dexamethasone suppressed salivary cortisol concentrations in a community sample of patients with posttraumatic stress disorder. *Biological Psychiatry* 55, 940–945. <https://doi.org/10.1016/j.biopsych.2003.12.021>
- Liu, J.J.W., Ein, N., Peck, K., Huang, V., Pruessner, J.C., Vickers, K., 2017. Sex differences in salivary cortisol reactivity to the Trier Social Stress Test (TSST): A meta-analysis. *Psychoneuroendocrinology* 82, 26–37. <https://doi.org/10.1016/j.psyneuen.2017.04.007>

- Liu, R.T., Scopelliti, K.M., Pittman, S.K., Zamora, A.S., 2018. Childhood maltreatment and non-suicidal self-injury: a systematic review and meta-analysis. *The Lancet Psychiatry* 5, 51–64. [https://doi.org/10.1016/S2215-0366\(17\)30469-8](https://doi.org/10.1016/S2215-0366(17)30469-8)
- Locatelli, V., Bresciani, E., Tamiazzo, L., Torsello, A., 2009. Central Nervous System-Acting Drugs Influencing Hypothalamic-Pituitary-Adrenal Axis Function, in: Loche, S., Cappa, M., Ghizzoni, L., Maghnie, M., Savage, M.O. (Eds.), *Endocrine Development*. KARGER, Basel, pp. 108–120. <https://doi.org/10.1159/000262533>
- \* Lopes, R.P., Grassi-Oliveira, R., de Almeida, L.R., Stein, L.M., Luz, C., Teixeira, A.L., Bauer, M.E., 2012. Neuroimmunoendocrine Interactions in Patients with Recurrent Major Depression, Increased Early Life Stress and Long-Standing Posttraumatic Stress Disorder Symptoms. *Neuroimmunomodulation* 19, 33–42. <https://doi.org/10.1159/000327352>
- \* Lovallo, W.R., Cohoon, A.J., Acheson, A., Sorocco, K.H., Vincent, A.S., 2019. Blunted stress reactivity reveals vulnerability to early life adversity in young adults with a family history of alcoholism. *Addiction* 114, 798–806. <https://doi.org/10.1111/add.14501>
- \* Lu, S., Gao, W., Huang, M., Li, L., Xu, Y., 2016. In search of the HPA axis activity in unipolar depression patients with childhood trauma: Combined cortisol awakening response and dexamethasone suppression test. *Journal of Psychiatric Research* 78, 24–30. <https://doi.org/10.1016/j.jpsychires.2016.03.009>
- \* Luecken, L.J., Appelhans, B.M., 2006. Early parental loss and salivary cortisol in young adulthood: the moderating role of family environment. *Dev Psychopathol* 18, 295–308. <https://doi.org/10.1017/S0954579406060160>
- \* Luecken, L.J., Kraft, A., Hagan, M.J., 2009. Negative relationships in the family-of-origin predict attenuated cortisol in emerging adults. *Hormones and Behavior* 55, 412–417. <https://doi.org/10.1016/j.yhbeh.2008.12.007>

- Lupien, S.J., McEwen, B.S., Gunnar, M.R., Heim, C., 2009. Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Reviews Neuroscience* 10, 434–445. <https://doi.org/10.1038/nrn2639>
- Mahon, P.B., Zandi, P.P., Potash, J.B., Nestadt, G., Wand, G.S., 2013. Genetic association of FKBP5 and CRHR1 with cortisol response to acute psychosocial stress in healthy adults. *Psychopharmacology* 227, 231–241. <https://doi.org/10.1007/s00213-012-2956-x>
- Manly, J.T., 2005. Advances in research definitions of child maltreatment. *Child Abuse & Neglect* 29, 425–439. <https://doi.org/10.1016/j.chiabu.2005.04.001>
- \* Martinson, A., Craner, J., Sigmon, S., 2016. Differences in HPA axis reactivity to intimacy in women with and without histories of sexual trauma. *Psychoneuroendocrinology* 65, 118–126. <https://doi.org/10.1016/j.psyneuen.2015.12.025>
- Mayer, S.E., Peckins, M., Kuhlman, K.R., Rajaram, N., Lopez-Duran, N.L., Young, E.A., Abelson, J.L., 2020. The roles of comorbidity and trauma exposure and its timing in shaping HPA axis patterns in depression. *Psychoneuroendocrinology* 120, 104776. <https://doi.org/10.1016/j.psyneuen.2020.104776>
- McCrory, E., De Brito, S.A., Viding, E., 2010. Research Review: The neurobiology and genetics of maltreatment and adversity: Research Review: The neurobiology and genetics of maltreatment and adversity. *Journal of Child Psychology and Psychiatry* 51, 1079–1095. <https://doi.org/10.1111/j.1469-7610.2010.02271.x>
- McEwen, B.S., Bowles, N.P., Gray, J.D., Hill, M.N., Hunter, R.G., Karatsoreos, I.N., Nasca, C., 2015. Mechanisms of stress in the brain. *Nature Neuroscience* 18, 1353–1363. <https://doi.org/10.1038/nn.4086>
- McEwen, B.S., Nasca, C., Gray, J.D., 2016. Stress Effects on Neuronal Structure: Hippocampus, Amygdala and Prefrontal Cortex. *Neuropsychopharmacology* 41, 3–23. <https://doi.org/10.1038/npp.2015.171>

- \* Mehta, D., Gonik, M., Klengel, T., Rex-Haffner, M., Menke, A., Rubel, J., Mercer, K.B., Pütz, B., Bradley, B., Holsboer, F., Ressler, K.J., Müller-Myhsok, B., Binder, E.B., 2011. Using Polymorphisms in FKBP5 to Define Biologically Distinct Subtypes of Posttraumatic Stress Disorder: Evidence From Endocrine and Gene Expression Studies. *Arch Gen Psychiatry* 68, 901-910.  
<https://doi.org/10.1001/archgenpsychiatry.2011.50>
- Melhem, N.M., Munroe, S., Marsland, A., Gray, K., Brent, D., Porta, G., Douaihy, A., Laudenslager, M.L., DePietro, F., Diler, R., Driscoll, H., Gopalan, P., 2017. Blunted HPA axis activity prior to suicide attempt and increased inflammation in attempters. *Psychoneuroendocrinology* 77, 284–294.  
<https://doi.org/10.1016/j.psyneuen.2017.01.001>
- \* Mello, A.F., Juruena, M.F., Maciel, M.R., Cavalcante-Nobrega, L.P., Cividanes, G.C., Fossaluzza, V., Calsavara, V., Mello, M.F., Cleare, A.J., Mari, J. de J., 2015. Factors related to the cortisol awakening response of children working on the streets and siblings, before and after 2 years of a psychosocial intervention. *Psychiatry Research* 225, 625–630. <https://doi.org/10.1016/j.psychres.2014.11.034>
- Merrick, M.T., Ford, D.C., Ports, K.A., Guinn, A.S., Chen, J., Klevens, J., Metzler, M., Jones, C.M., Simon, T.R., Daniel, V.M., Ottley, P., Mercy, J.A., 2019. Vital Signs: Estimated Proportion of Adult Health Problems Attributable to Adverse Childhood Experiences and Implications for Prevention — 25 States, 2015–2017. *MMWR Morb Mortal Wkly Rep* 68, 999–1005. <https://doi.org/10.15585/mmwr.mm6844e1>
- Meyer, J.S., Novak, M.A., 2012. Minireview: Hair Cortisol: A Novel Biomarker of Hypothalamic-Pituitary-Adrenocortical Activity. *Endocrinology* 153, 4120–4127.  
<https://doi.org/10.1210/en.2012-1226>
- \* Mielock, A.S., Morris, M.C., Rao, U., 2017. Patterns of cortisol and alpha-amylase reactivity to psychosocial stress in maltreated women. *Journal of Affective Disorders*

- 209, 46–52. <https://doi.org/10.1016/j.jad.2016.11.009>
- Miller, G.E., Chen, E., Zhou, E.S., 2007. If it goes up, must it come down? Chronic stress and the hypothalamic-pituitary-adrenocortical axis in humans. *Psychological Bulletin* 133, 25–45. <https://doi.org/10.1037/0033-2909.133.1.25>
- \* Monteleone, A.M., Monteleone, P., Volpe, U., De Riso, F., Fico, G., Giugliano, R., Nigro, M., Maj, M., 2018. Impaired cortisol awakening response in eating disorder women with childhood trauma exposure: evidence for a dose-dependent effect of the traumatic load. *Psychol. Med.* 48, 952–960. <https://doi.org/10.1017/S0033291717002409>
- Moore, A., Aitken, R., Burke, C., Gaskell, S., Groom, G., Holder, G., Selby, C., Wood, P., 1985. Cortisol Assays: Guidelines for the Provision of a Clinical Biochemistry Service. *Annals of Clinical Biochemistry: An international journal of biochemistry and laboratory medicine* 22, 435–454. <https://doi.org/10.1177/000456328502200501>
- \* Morris, M.C., Abelson, J.L., Mielock, A.S., Rao, U., 2017. Psychobiology of cumulative trauma: hair cortisol as a risk marker for stress exposure in women. *Stress* 20, 350–354. <https://doi.org/10.1080/10253890.2017.1340450>
- Myers, B., McKlveen, J.M., Herman, J.P., 2014. Glucocorticoid actions on synapses, circuits, and behavior: Implications for the energetics of stress. *Frontiers in Neuroendocrinology* 35, 180–196. <https://doi.org/10.1016/j.yfrne.2013.12.003>
- Nelson, E.C., Heath, A.C., Madden, P.A.F., Cooper, M.L., Dinwiddie, S.H., Bucholz, K.K., Glowinski, A., McLaughlin, T., Dunne, M.P., Statham, D.J., Martin, N.G., 2002. Association Between Self-reported Childhood Sexual Abuse and Adverse Psychosocial Outcomes: Results From a Twin Study. *Arch Gen Psychiatry* 59, 139. <https://doi.org/10.1001/archpsyc.59.2.139>
- Newbury, J.B., Arseneault, L., Moffitt, T.E., Caspi, A., Danese, A., Baldwin, J.R., Fisher, H.L., 2018. Measuring childhood maltreatment to predict early-adult psychopathology: Comparison of prospective informant-reports and retrospective self-

- reports. *Journal of Psychiatric Research* 96, 57–64.  
<https://doi.org/10.1016/j.jpsychires.2017.09.020>
- \* Newport, D.J., Heim, C., Bonsall, R., Miller, A.H., Nemeroff, C.B., 2004. Pituitary-adrenal responses to standard and low-dose dexamethasone suppression tests in adult survivors of child abuse. *Biol Psychiatry* 55, 10–20. [https://doi.org/10.1016/s0006-3223\(03\)00692-9](https://doi.org/10.1016/s0006-3223(03)00692-9)
- Nicolaides, N.C., Kyratzi, E., Lamprokostopoulou, A., Chrousos, G.P., Charmandari, E., 2015. Stress, the Stress System and the Role of Glucocorticoids. *Neuroimmunomodulation* 22, 6–19. <https://doi.org/10.1159/000362736>
- Norman, R.E., Byambaa, M., De, R., Butchart, A., Scott, J., Vos, T., 2012. The Long-Term Health Consequences of Child Physical Abuse, Emotional Abuse, and Neglect: A Systematic Review and Meta-Analysis. *PLoS Medicine* 9, e1001349.  
<https://doi.org/10.1371/journal.pmed.1001349>
- \* O'Connor, D.B., Green, J.A., Ferguson, E., O'Carroll, R.E., O'Connor, R.C., 2018. Effects of childhood trauma on cortisol levels in suicide attempters and ideators. *Psychoneuroendocrinology* 88, 9–16. <https://doi.org/10.1016/j.psyneuen.2017.11.004>
- O'Connor, M., Slopen, N., Becares, L., Burgner, D., Williams, D.R., Priest, N., 2020. Inequalities in the Distribution of Childhood Adversity From Birth to 11 Years. *Academic Pediatrics* 20, 609–618. <https://doi.org/10.1016/j.acap.2019.12.004>
- \* Otte, C., Neylan, T.C., Pole, N., Metzler, T., Best, S., Henn-Haase, C., Yehuda, R., Marmar, C.R., 2005. Association between childhood trauma and catecholamine response to psychological stress in police academy recruits. *Biological Psychiatry* 57, 27–32.  
<https://doi.org/10.1016/j.biopsych.2004.10.009>
- \* Ouellet-Morin, I., Odgers, C.L., Danese, A., Bowes, L., Shakoor, S., Papadopoulos, A.S., Caspi, A., Moffitt, T.E., Arseneault, L., 2011. Blunted Cortisol Responses to Stress Signal Social and Behavioral Problems Among Maltreated/Bullied 12-Year-Old

- Children. *Biological Psychiatry* 70, 1016–1023.  
<https://doi.org/10.1016/j.biopsych.2011.06.017>
- \* Ouellet-Morin, I., Robitaille, M.-P., Langevin, S., Cantave, C., Brendgen, M., Lupien, S. J., 2018. Enduring effect of childhood maltreatment on cortisol and heart rate responses to stress: The moderating role of severity of experiences. *Dev Psychopathol* 31, 497–508. <https://doi.org/10.1017/S0954579418000123>
- Palma-Gudiel, H., Córdova-Palomera, A., Leza, J.C., Fañanás, L., 2015. Glucocorticoid receptor gene (NR3C1) methylation processes as mediators of early adversity in stress-related disorders causality: A critical review. *Neuroscience & Biobehavioral Reviews* 55, 520–535. <https://doi.org/10.1016/j.neubiorev.2015.05.016>
- \* Peng, H., Long, Y., Li, J., Guo, Y., Wu, H., Yang, Y., Ding, Y., He, J., Ning, Y., 2014. Hypothalamic-pituitary-adrenal axis functioning and dysfunctional attitude in depressed patients with and without childhood neglect. *BMC Psychiatry* 14, 45. <https://doi.org/10.1186/1471-244X-14-45>
- Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R., Rushton, L., 2008. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology* 61, 991–996.  
<https://doi.org/10.1016/j.jclinepi.2007.11.010>
- \* Pierrehumbert, B., Torrisi, R., Glatz, N., Dimitrova, N., Heinrichs, M., Halfon, O., 2009. The influence of attachment on perceived stress and cortisol response to acute stress in women sexually abused in childhood or adolescence. *Psychoneuroendocrinology* 34, 924–938. <https://doi.org/10.1016/j.psyneuen.2009.01.006>
- Porter, C., Palmier-Claus, J., Branitsky, A., Mansell, W., Warwick, H., Varese, F., 2020. Childhood adversity and borderline personality disorder: a meta-analysis. *Acta Psychiatr Scand* 141, 6–20. <https://doi.org/10.1111/acps.13118>
- Pruessner, J.C., Kirschbaum, C., Meinlschmid, G., Hellhammer, D.H., 2003. Two formulas



- for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology* 28, 916–931. [https://doi.org/10.1016/S0306-4530\(02\)00108-7](https://doi.org/10.1016/S0306-4530(02)00108-7)
- \* Puetz, V.B., Zweerings, J., Dahmen, B., Ruf, C., Scharke, W., Herpertz-Dahlmann, B., Konrad, K., 2016. Multidimensional assessment of neuroendocrine and psychopathological profiles in maltreated youth. *J Neural Transm* 123, 1095–1106. <https://doi.org/10.1007/s00702-016-1509-6>
- \* Quevedo, K., Doty, J., Roos, L., Anker, J.J., 2017. The cortisol awakening response and anterior cingulate cortex function in maltreated depressed versus non-maltreated depressed youth. *Psychoneuroendocrinology* 86, 87–95. <https://doi.org/10.1016/j.psyneuen.2017.09.001>
- \* Rao, U., Morris, M.C., 2015. Cortisol Responses to Psychosocial Stress: The Role of Childhood Maltreatment and Depression. *Int J Public Ment Health Neurosci* 2.
- \* Reichl, C., Heyer, A., Brunner, R., Parzer, P., Völker, J.M., Resch, F., Kaess, M., 2016. Hypothalamic-pituitary-adrenal axis, childhood adversity and adolescent nonsuicidal self-injury. *Psychoneuroendocrinology* 74, 203–211. <https://doi.org/10.1016/j.psyneuen.2016.09.011>
- Rohatgi, A., 2012. Web Plot Digitizer. Austin, Texas, USA.
- Ryan, R., Booth, S., Spathis, A., Mollart, S., Clow, A., 2016. Use of Salivary Diurnal Cortisol as an Outcome Measure in Randomised Controlled Trials: a Systematic Review. *Annals of Behavioral Medicine* 50, 210–236. <https://doi.org/10.1007/s12160-015-9753-9>
- Sapolsky, R.M., 1999. Glucocorticoids, stress, and their adverse neurological effects: relevance to aging. *Experimental gerontology* 34, 721-732.
- Sapolsky, R.M., Romero, L.M., Munck, A.U., 2000. How Do Glucocorticoids Influence Stress Responses? Integrating Permissive, Suppressive, Stimulatory, and Preparative

- Actions<sup>1</sup>. *Endocrine Reviews* 21, 55–89. <https://doi.org/10.1210/edrv.21.1.0389>
- \* Schalinski, I., Elbert, T., Steudte-Schmiedgen, S., Kirschbaum, C., 2015. The Cortisol Paradox of Trauma-Related Disorders: Lower Phasic Responses but Higher Tonic Levels of Cortisol Are Associated with Sexual Abuse in Childhood. *PLoS ONE* 10, e0136921. <https://doi.org/10.1371/journal.pone.0136921>
- \* Schreuder, M.M., Vinkers, C.H., Mesman, E., Claes, S., Nolen, W.A., Hillegers, M.H.J., 2016. Childhood trauma and HPA axis functionality in offspring of bipolar parents. *Psychoneuroendocrinology* 74, 316–323. <https://doi.org/10.1016/j.psyneuen.2016.09.017>
- \* Schwaiger, M., Grinberg, M., Moser, D., Zang, J.C.S., Heinrichs, M., Hengstler, J.G., Rahnenführer, J., Cole, S., Kumsta, R., 2016. Altered Stress-Induced Regulation of Genes in Monocytes in Adults with a History of Childhood Adversity. *Neuropsychopharmacol* 41, 2530–2540. <https://doi.org/10.1038/npp.2016.57>
- 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. <https://doi.org/10.1016/j.psyneuen.2014.07.022>
- \* Seltzer, L.J., Ziegler, T., Connolly, M.J., Prosofski, A.R., Pollak, S.D., 2014. Stress-Induced Elevation of Oxytocin in Maltreated Children: Evolution, Neurodevelopment, and Social Behavior. *Child Dev* 85, 501–512. <https://doi.org/10.1111/cdev.12136>
- \* Shenk, C.E., Noll, J.G., Putnam, F.W., Trickett, P.K., 2010. A prospective examination of the role of childhood sexual abuse and physiological asymmetry in the development of psychopathology. *Child Abuse & Neglect* 34, 752–761. <https://doi.org/10.1016/j.chiabu.2010.02.010>
- \* Shenk, C.E., Putnam, F.W., Rausch, J.R., Peugh, J.L., Noll, J.G., 2014. A longitudinal study of several potential mediators of the relationship between child maltreatment and

- posttraumatic stress disorder symptoms. *Dev Psychopathol* 26, 81–91.  
<https://doi.org/10.1017/S0954579413000916>
- Shirazi, S.N., Friedman, A.R., Kaufer, D., Sakhai, S.A., 2015. Glucocorticoids and the Brain: Neural Mechanisms Regulating the Stress Response, in: Wang, J.-C., Harris, C. (Eds.), *Glucocorticoid Signaling: From Molecules to Mice to Man*. Springer New York, New York, NY, pp. 235–252. [https://doi.org/10.1007/978-1-4939-2895-8\\_10](https://doi.org/10.1007/978-1-4939-2895-8_10)
- Shonkoff, J.P., Garner, A.S., the Committee on psychosocial aspects of child and family health committee on early childhood, adoption, and dependent care section on developmental and behavioral pediatrics, Siegel, B.S., Dobbins, M.I., Earls, M.F., Garner, A.S., McGuinn, L., Pascoe, J., Wood, D.L., 2012. The Lifelong Effects of Early Childhood Adversity and Toxic Stress. *PEDIATRICS* 129, e232–e246.  
<https://doi.org/10.1542/peds.2011-2663>
- Slavich, G.M., Giletta, M., Helms, S.W., Hastings, P.D., Rudolph, K.D., Nock, M.K., Prinstein, M.J., 2020. Interpersonal life stress, inflammation, and depression in adolescence: Testing Social Signal Transduction Theory of Depression. *Depression and Anxiety* 37, 179–193. <https://doi.org/10.1002/da.22987>
- \* Smeets, T., Geraerts, E., Jelicic, M., Merckelbach, H., 2007. Delayed recall of childhood sexual abuse memories and the awakening rise and diurnal pattern of cortisol. *Psychiatry Res* 152, 197–204. <https://doi.org/10.1016/j.psychres.2006.07.008>
- Smith, R., Mill, J., 2011. Epigenetics and Chronic Diseases: An Overview, in: Roach, H.I., Bronner, F., Oreffo, R.O.C. (Eds.), *Epigenetic Aspects of Chronic Diseases*. Springer London, London, pp. 1–20. [https://doi.org/10.1007/978-1-84882-644-1\\_1](https://doi.org/10.1007/978-1-84882-644-1_1)
- Spencer, R.L., Chun, L.E., Hartsock, M.J., Woodruff, E.R., 2018. Glucocorticoid hormones are both a major circadian signal and major stress signal: How this shared signal contributes to a dynamic relationship between the circadian and stress systems. *Frontiers in Neuroendocrinology* 49, 52–71.

- <https://doi.org/10.1016/j.yfrne.2017.12.005>
- Spencer, R.L., Deak, T., 2017. A users guide to HPA axis research. *Physiology & Behavior* 178, 43–65. <https://doi.org/10.1016/j.physbeh.2016.11.014>
- Spiga, F., Walker, J.J., Terry, J.R., Lightman, S.L., 2014. HPA Axis-Rhythms, in: Terjung, R. (Ed.), *Comprehensive Physiology*. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 1273–1298. <https://doi.org/10.1002/cphy.c140003>
- \* Spitzer, C., Otte, C., Kuehl, L.K., May, A., Schultebrasucks, K., Hellmann-Regen, J., Wingenfeld, K., 2018. The dexamethasone corticotropin releasing hormone test in healthy and depressed women with and without childhood adversity. *Psychoneuroendocrinology* 87, 147–151. <https://doi.org/10.1016/j.psyneuen.2017.10.016>
- Stalder, T., Kirschbaum, C., 2012. Analysis of cortisol in hair – State of the art and future directions. *Brain, Behavior, and Immunity* 26, 1019–1029. <https://doi.org/10.1016/j.bbi.2012.02.002>
- Stalder, T., Kirschbaum, C., Kudielka, B.M., Adam, E.K., Pruessner, J.C., Wüst, S., Dockray, S., Smyth, N., Evans, P., Hellhammer, D.H., Miller, R., Wetherell, M.A., Lupien, S.J., Clow, A., 2016. Assessment of the cortisol awakening response: Expert consensus guidelines. *Psychoneuroendocrinology* 63, 414–432. <https://doi.org/10.1016/j.psyneuen.2015.10.010>
- Stalder, T., Steudte-Schmiedgen, S., Alexander, N., Klucken, T., Vater, A., Wichmann, S., Kirschbaum, C., Miller, R., 2017. Stress-related and basic determinants of hair cortisol in humans: A meta-analysis. *Psychoneuroendocrinology* 77, 261–274. <https://doi.org/10.1016/j.psyneuen.2016.12.017>
- \* Stein, M.B., Yehuda, R., Koverola, C., Hanna, C., 1997. Enhanced dexamethasone suppression of plasma cortisol in adult women traumatized by childhood sexual abuse. *Biol Psychiatry* 42, 680–686. [https://doi.org/10.1016/s0006-3223\(96\)00489-1](https://doi.org/10.1016/s0006-3223(96)00489-1)

- \* Steudte, S., Kirschbaum, C., Gao, W., Alexander, N., Schönfeld, S., Hoyer, J., Stalder, T., 2013. Hair Cortisol as a Biomarker of Traumatization in Healthy Individuals and Posttraumatic Stress Disorder Patients. *Biological Psychiatry* 74, 639–646.  
<https://doi.org/10.1016/j.biopsych.2013.03.011>
- Stoltenborgh, M., Bakermans-Kranenburg, M.J., Alink, L.R.A., van IJzendoorn, M.H., 2015. The Prevalence of Child Maltreatment across the Globe: Review of a Series of Meta-Analyses: Prevalence of Child Maltreatment across the Globe. *Child Abuse Review* 24, 37–50. <https://doi.org/10.1002/car.2353>
- Strüber, N., Strüber, D., Roth, G., 2014. Impact of early adversity on glucocorticoid regulation and later mental disorders. *Neuroscience & Biobehavioral Reviews* 38, 17–37. <https://doi.org/10.1016/j.neubiorev.2013.10.015>
- \* Sullivan, M.W., Bennett, D.S., Lewis, M., 2013. Individual Differences in the Cortisol Responses of Neglected and Comparison Children. *Child Maltreat* 18, 8–16.  
<https://doi.org/10.1177/1077559512449378>
- \* Sumner, J.A., McLaughlin, K.A., Walsh, K., Sheridan, M.A., Koenen, K.C., 2014. CRHR1 genotype and history of maltreatment predict cortisol reactivity to stress in adolescents. *Psychoneuroendocrinology* 43, 71–80.  
<https://doi.org/10.1016/j.psyneuen.2014.02.002>
- \* Suzuki, A., Poon, L., Papadopoulos, A.S., Kumari, V., Cleare, A.J., 2014. Long term effects of childhood trauma on cortisol stress reactivity in adulthood and relationship to the occurrence of depression. *Psychoneuroendocrinology* 50, 289–299.  
<https://doi.org/10.1016/j.psyneuen.2014.09.007>
- Teicher, M.H., Samson, J.A., 2013. Childhood Maltreatment and Psychopathology: A Case for Ecophenotypic Variants as Clinically and Neurobiologically Distinct Subtypes. *AJP* 170, 1114–1133. <https://doi.org/10.1176/appi.ajp.2013.12070957>
- Teicher, M.H., Samson, J.A., Anderson, C.M., Ohashi, K., 2016. The effects of childhood

- maltreatment on brain structure, function and connectivity. *Nat Rev Neurosci* 17, 652–666. <https://doi.org/10.1038/nrn.2016.111>
- \* Trickett, P.K., Gordis, E., Peckins, M.K., Susman, E.J., 2014. Stress Reactivity in Maltreated and Comparison Male and Female Young Adolescents. *Child Maltreat* 19, 27–37. <https://doi.org/10.1177/1077559513520466>
- Trickett, P.K., Noll, J.G., Susman, E.J., Shenk, C.E., Putnam, F.W., 2010. Attenuation of cortisol across development for victims of sexual abuse. *Dev Psychopathol* 22, 165–175. <https://doi.org/10.1017/S0954579409990332>
- Turecki, G., Meaney, M.J., 2016. Effects of the Social Environment and Stress on Glucocorticoid Receptor Gene Methylation: A Systematic Review. *Biological Psychiatry* 79, 87–96. <https://doi.org/10.1016/j.biopsych.2014.11.022>
- Turner, A.I., Smyth, N., Hall, S.J., Torres, S.J., Hussein, M., Jayasinghe, S.U., Ball, K., Clow, A.J., 2020. Psychological stress reactivity and future health and disease outcomes: A systematic review of prospective evidence. *Psychoneuroendocrinology* 114, 104599. <https://doi.org/10.1016/j.psyneuen.2020.104599>
- Vachon, D.D., Krueger, R.F., Rogosch, F.A., Cicchetti, D., 2015. Assessment of the Harmful Psychiatric and Behavioral Effects of Different Forms of Child Maltreatment. *JAMA Psychiatry* 72, 1135. <https://doi.org/10.1001/jamapsychiatry.2015.1792>
- \* van der Vegt, E.J.M., van der Ende, J., Kirschbaum, C., Verhulst, F.C., Tiemeier, H., 2009. Early neglect and abuse predict diurnal cortisol patterns in adults. A study of international adoptees. *Psychoneuroendocrinology* 34, 660–669. <https://doi.org/10.1016/j.psyneuen.2008.11.004>
- \* van Zuiden, M., Kavelaars, A., Rademaker, A.R., Vermetten, E., Heijnen, C.J., Geuze, E., 2011. A prospective study on personality and the cortisol awakening response to predict posttraumatic stress symptoms in response to military deployment. *Journal of Psychiatric Research* 45, 713–719. <https://doi.org/10.1016/j.jpsychires.2010.11.013>

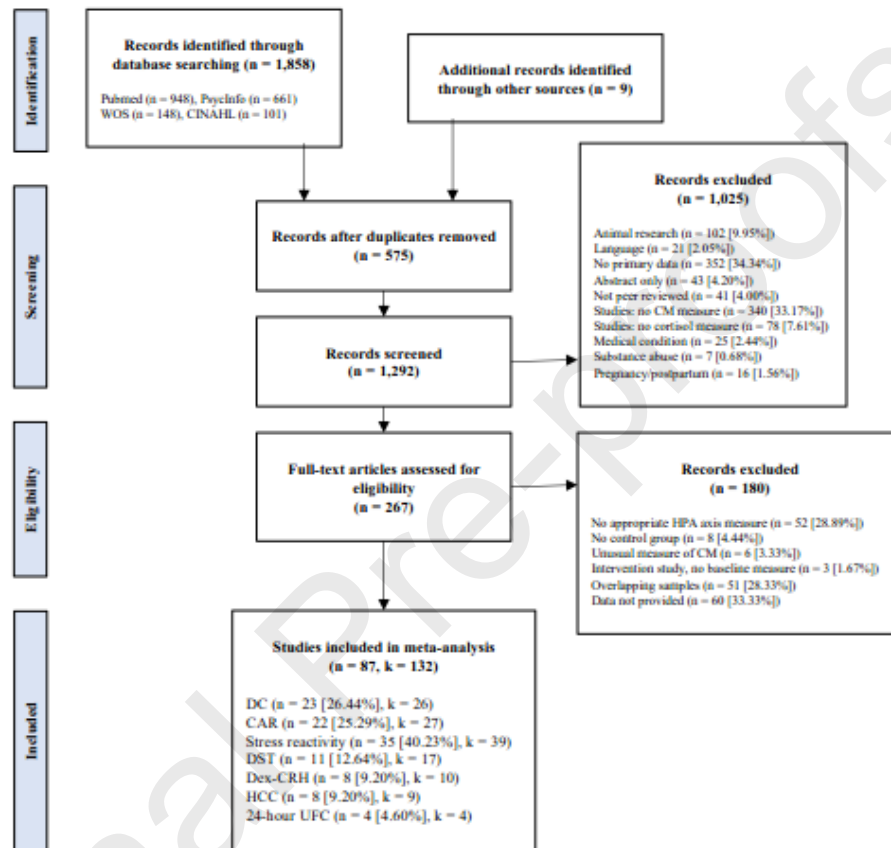
- Veroniki, A.A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., Kuss, O., Higgins, J.P., Langan, D., Salanti, G., 2016. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res. Syn. Meth.* 7, 55–79.  
<https://doi.org/10.1002/jrsm.1164>
- Viechtbauer, W., Cheung, M.W.-L., 2010. Outlier and influence diagnostics for meta-analysis. *Res. Synth. Method* 1, 112–125. <https://doi.org/10.1002/jrsm.11>
- Voellmin, A., Winzeler, K., Hug, E., Wilhelm, F.H., Schaefer, V., Gaab, J., La Marca, R., Pruessner, J.C., Bader, K., 2015. Blunted endocrine and cardiovascular reactivity in young healthy women reporting a history of childhood adversity. *Psychoneuroendocrinology* 51, 58–67. <https://doi.org/10.1016/j.psyneuen.2014.09.008>
- \* Vreeburg, S.A., Hoogendijk, W.J.G., van Pelt, J., DeRijk, R.H., Verhagen, J.C.M., van Dyck, R., Smit, J.H., Zitman, F.G., Penninx, B.W.J.H., 2009. Major Depressive Disorder and Hypothalamic-Pituitary-Adrenal Axis Activity: Results From a Large Cohort Study. *Arch Gen Psychiatry* 66, 617.  
<https://doi.org/10.1001/archgenpsychiatry.2009.50>
- Watson, S., Gallagher, P., Smith, M.S., Ferrier, I.N., Young, A.H., 2006. The dex/CRH test—Is it better than the DST? *Psychoneuroendocrinology* 31, 889–894.  
<https://doi.org/10.1016/j.psyneuen.2006.03.001>
- \* Watson, S., Owen, B.M., Gallagher, P., Hearn, A.J., Young, A.H., Ferrier, I.N., 2007. Family history, early adversity and the hypothalamic-pituitary-adrenal (HPA) axis: Mediation of the vulnerability to mood disorders. *Neuropsychiatr Dis Treat* 3, 647–653.
- \* White, L.O., Ising, M., von Klitzing, K., Sierau, S., Michel, A., Klein, A.M., Andreas, A., Keil, J., Quintero, L., Müller-Myhsok, B., Uhr, M., Gausche, R., Manly, J.T., Crowley, M.J., Kirschbaum, C., Stalder, T., 2017. Reduced hair cortisol after maltreatment mediates externalizing symptoms in middle childhood and adolescence.

- J Child Psychol Psychiatr 58, 998–1007. <https://doi.org/10.1111/jcpp.12700>
- \* Wingenfeld, K., Kuehl, L.K., Boeker, A., Schultebrasucks, K., Ritter, K., Hellmann-Regen, J., Otte, C., Spitzer, C., 2017. Stress reactivity and its effects on subsequent food intake in depressed and healthy women with and without adverse childhood experiences. *Psychoneuroendocrinology* 80, 122–130.  
<https://doi.org/10.1016/j.psyneuen.2017.03.014>
- Witt, A., Brown, R.C., Plener, P.L., Brähler, E., Fegert, J.M., 2017. Child maltreatment in Germany: prevalence rates in the general population. *Child and Adolescent Psychiatry and Mental Health* 11. <https://doi.org/10.1186/s13034-017-0185-0>
- Young, E.S., Doom, J.R., Farrell, A.K., Carlson, E.A., Englund, M.M., Miller, G.E., Gunnar, M.R., Roisman, G.I., Simpson, J.A., 2020. Life stress and cortisol reactivity: An exploratory analysis of the effects of stress exposure across life on HPA-axis functioning. *Dev Psychopathol* 1–12. <https://doi.org/10.1017/S0954579419001779>
- Young, E.S., Farrell, A.K., Carlson, E.A., Englund, M.M., Miller, G.E., Gunnar, M.R., Roisman, G.I., Simpson, J.A., 2019. The Dual Impact of Early and Concurrent Life Stress on Adults' Diurnal Cortisol Patterns: A Prospective Study. *Psychol Sci* 30, 739–747. <https://doi.org/10.1177/0956797619833664>
- Zänkert, S., Bellingrath, S., Wüst, S., Kudielka, B.M., 2019. HPA axis responses to psychological challenge linking stress and disease: What do we know on sources of intra- and interindividual variability? *Psychoneuroendocrinology* 105, 86–97.  
<https://doi.org/10.1016/j.psyneuen.2018.10.027>
- Zorn, J.V., Schür, R.R., Boks, M.P., Kahn, R.S., Joëls, M., Vinkers, C.H., 2017. Cortisol stress reactivity across psychiatric disorders: A systematic review and meta-analysis. *Psychoneuroendocrinology* 77, 25–36. <https://doi.org/10.1016/j.psyneuen.2016.11.036>



## Figure Captions

**Fig. 1.** PRISMA flow diagram describing the process of study selection. WOS = Web of Science; CINAHL = Cumulative Index to Nursing and Allied Health Literature; CM = child maltreatment; HPA axis = hypothalamic-pituitary-adrenal axis; DC = diurnal cortisol; CAR = cortisol awakening response; DST = dexamethasone suppression test; Dex-CRH = dexamethasone-corticotropin releasing hormone; HCC = hair cortisol concentrations; 24-hour UFC = 24-hour urinary free cortisol; n = number of studies; k = number of group comparisons.



**Table 1**  
Summary characteristics of included studies that reported on diurnal cortisol (DC).

Study	Total Sample					CM Sample			Control Sample			CM Assessment				Cortisol Assessment					Data		
	N	Age M(SD)	Sex	Ethn.	Psycho-path.	n	Age M(SD)	Sex	n	Age M(SD)	Sex	Method	Instrument	Grouping	Type	Type of sample	Unit	Time points	Slope	Rel.	Indices	Source	Data (re)grouped
Bernard et al., 2010	435	1.08 (0.58)	46.4	72.6	no	340 <sup>a</sup>	1.00 (0.56)	47.9	95	1.34 (0.55)	41.1	informant (record)	CPS	other (record)	NA	saliva	µg/dl	awak., bedtime	NA	2	morning, evening	provided, group	no
Carrion et al., 2002 <sup>b</sup>	82	10.78 (1.79)	41.5	36.6	yes*, PTSD	51	10.70 (1.90)	41.2	31	10.90 (1.60)	40.6	self-report (quest.)	PTSD Reaction Index	other (spec.)	EA, PA, SA, N, loss/separation <sup>o</sup>	saliva	µg/dl	pre breakfast, pre bed	NA	3	morning, evening	paper, table	no
Cicchetti et al., 2010	553	10.02 (1.87)	47.6	81.7	NA	265 <sup>c</sup>	10.00 (1.86)	38.9	288	10.08 (1.89)	55.6	informant (mixed)	MCS (CPS), other MMCI	other (record)	EA, PA, SA, N	saliva	log	9am, 4pm	NA	3	morning, evening	paper, table	no
Faravelli et al., 2010	93 <sup>d</sup>	43.60 (15.40)	50.2	NA	yes, mixed	32	NA	NA	61	NA	NA	self-report (mixed)	CECA-Q, self-dev.	other (not spec.)	PA, SA, loss <sup>o</sup>	saliva	nmol/l	8am, 8pm	other	1	all	paper, table	no
Faravelli et al., 2017 <sup>a</sup>	102	43.46 (12.24)	48.0	0.0	NA	40	43.50 (12.51)	42.0	62	43.43 (12.16)	52.0	self-report (mixed)	CECA-Q, FPI	other (spec.)	PA, SA, N, loss <sup>o</sup>	saliva	nmol/l	30min post-awak., 8pm	other	1	all	provided, group	no
Faravelli et al., 2017 <sup>b</sup>	54	43.74 (10.51)	44.0	0.0	yes, psych.	30	42.20 (11.35)	47.0	24	45.66 (9.23)	42.0	self-report (mixed)	CECA-Q, FPI	other (spec.)	PA, SA, N, loss <sup>o</sup>	saliva	nmol/l	30min post-awak., 8pm	other	1	all	provided, group	no
Fisher et al., 2007	177	4.40 (0.83)	46.3	11.0	NA	117 <sup>e</sup>	NA	46.2	60	NA	47.0	informant (record)	CWA	other (record)	EA, PA, SA, N	saliva	µg/dl	30min post awak., 30 min pre bedtime	other	2	all	paper, table	no
Fuchs et al., 2017	82	35.38 (5.07)	100.0	NA	NA	37	36.47 (4.92)	100.0	45	34.48 (5.06)	100.0	self-report (quest.)	CTQ	cut-offs	PA, SA	saliva	nmol/l	awak., 8pm	other	2	all	provided, group	no
Hibel et al., 2019 <sup>f</sup>	248	4.91 (1.14)	49.6	74.6	NA	165	4.93 (1.15)	49.7	83	4.86 (1.13)	49.4	informant (mixed)	MCS (CPS), other MMCI	other (record)	EA, PA, SA, N	saliva	µg/ml	awak., bedtime	wake-to-bed	2	all	provided, group	no
Keeshin et al., 2014	36	14.97 (1.40)	100.0	NA	yes*, mixed	24	15.04 (1.45)	100.0	12	14.84 (1.34)	100.0	informant (record)	medical record	other (record)	SA	saliva	log	awak., between 4-6pm	NA	3	morning, evening	paper, table	no
Klaassens et al., 2009	22	47.58 (11.70)	100.0	NA	no	10	47.80 (12.10)	100.0	12	47.40 (11.90)	100.0	self-report (quest.)	CTQ <sup>g</sup>	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., 11pm	NA	2	morning, evening	paper, figure	no
Kuhlman et al., 2015	79	12.73 (2.27)	40.5	22.0 <sup>h</sup>	yes, mixed	60	12.91 (2.26)	41.2	19	12.16 (2.27)	36.8	informant (quest.)	ETI	other (spec.)	EA, PA, SA,	saliva	µg/dl	awak., pre bed	wake-to-bed	2	all	provided, group	yes
Kumsta et al., 2017	57	24.02 (0.85)	61.4	NA	NA	44 <sup>i</sup>	24.15 (0.90)	68.2	13	23.55 (0.45)	38.5	informant (NA)	no instr. used	institution-alization	N	saliva	nmol/l	awak., 8pm	wake-to-bed	2	all	provided, group	no
Kuras et al., 2017	61	33.80 (2.30)	49.2	NA	no <sup>o</sup>	29	NA	NA	32	NA	NA	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., awak. + 13h	NA	2	morning, evening	paper, figure <sup>+</sup>	no
Lindley et al., 2004	16 <sup>j</sup>	40.00 (11.17)	87.5	6.2	yes, PTSD	9	34.56 (9.28)	88.9	7	47.00 (9.78)	85.7	self-report (quest.)	THS	other (spec.)	PA, SA	saliva	µg/dl	8am, 10pm	other	1	all	provided, raw	no
Lopes et al., 2012	38 <sup>k</sup>	40.29 (8.69)	100.0	NA	yes, MDD	16	41.06 (6.46)	100.0	22	39.73 (10.12)	100.0	self-report (quest.)	CTQ	other (not spec.)	EA, PA, SA, N	saliva	nmol/l	8am, 8pm	NA	1	morning, evening	paper, figure <sup>+</sup>	no
Lovallo et al., 2019 <sup>l</sup>	699	23.70 (3.14)	55.9	20.3	no	335	23.90 (3.26)	63.0	364	23.50 (3.02)	49.5	self-report (interview)	C-DIS-IV	other (spec.)	EA, PA, SA	saliva	µg/dl	awak., bedtime	wake-to-bed	1	all	provided, group	yes
Puetz et al., 2016	40	10.55 (1.71)	47.5	35.0	yes, mixed	17	10.88 (1.65)	29.0	23	10.30 (1.74)	61.0	informant (record)	SSR (medical record)	other (record)	EA, PA, N	saliva	nmol/l	30min post awak., 30min pre sleep	NA	2	morning, evening	paper, table	no
Reichl et al., 2016 <sup>a</sup>	25	16.25 (1.12)	96.0	NA	no	3	16.55 (1.34)	100.0	22	16.20 (1.12)	95.5	self-report (interview)	CECA	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., pre bed	wake-to-bed	3	all	provided, raw	yes
Reichl et al., 2016 <sup>b</sup>	25	16.30 (1.30)	92.0	NA	yes, mixed	17	16.60 (1.29)	94.1	8	15.65 (1.14)	87.5	self-report (interview)	CECA	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., pre bed	wake-to-bed	3	all	provided, raw	yes

**Table 1**

Summary characteristics of included studies that reported on diurnal cortisol (DC).

Study	Total Sample					CM Sample			Control Sample			CM Assessment				Cortisol Assessment					Data		
	N	Age M(SD)	Sex	Ethn.	Psycho-path.	n	Age M(SD)	Sex	n	Age M(SD)	Sex	Method	Instrument	Grouping	Type	Type of sample	Unit	Time points	Slope	Rel.	Indices	Source	Data (re)grouped
Schreuder et al., 2016	114	26.75 (2.85)	45.6	NA	yes, mixed	16	27.81 (2.90)	43.8	98	26.58 (2.81)	45.9	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	µg/dl	awak., 10pm	wake-to-bed	1	all	provided, group	yes
Smeets et al., 2007	22 <sup>m</sup>	40.50 (11.63)	100.0	NA	no <sup>o</sup>	13	38.77 (11.38)	100.0	9	43.00 (12.21)	100.0	self-report (interview)	self-dev.	other (spec.)	SA	saliva	nmol/l	awak., 8pm	wake-to-bed	2	all	provided, raw	no
Steutde et al., 2013	30 <sup>n</sup>	37.87 (12.04)	90.0	0.0	no	4	42.50 (11.39)	100.0	26	37.15 (12.19)	88.5	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., bedtime	wake-to-bed	1	all	provided, group	yes
van der Vegt et al., 2009	529	30.83 (1.22)	33.1	NA	NA	235 <sup>o</sup>	30.87 (1.24)	51.9	294	30.80 (1.20)	18.0	informant (interview)	self-dev.	other (spec.)	N	saliva	nmol/l	awak., bedtime	NA	1	morning, evening	paper, table	no
Vreeburg et al., 2009a <sup>p</sup>	1131	42.67 (12.27)	66.3	NA	yes, mixed <sup>d</sup>	645	43.92 (11.68)	70.9	486	41.02 (12.83)	60.3	self-report (interview)	NEMESIS CTI	other (spec.)	EA, PA, SA, N	saliva	nmol/l	awak., 11pm	wake-to-bed	1	all	provided, group	yes
Vreeburg et al., 2009b <sup>p</sup>	498	42.88 (14.46)	61.3	NA	no <sup>r</sup>	120	47.62 (12.64)	73.3	369	41.33 (14.69)	57.5	self-report (interview)	NEMESIS CTI	other (spec.)	EA, PA, SA, N	saliva	nmol/l	awak., 11pm	wake-to-bed	1	all	provided, group	yes

*Note.* N = sample size; M = mean; SD = standard deviation; The sex ratio is indicated as percentage of female participants; Ethn. = ethnicity; The ethnicity ratio is indicated as percentage of non-Caucasians; NA = not assessed; Psychopath. = psychopathology, whereby the following definitions have been used: yes = at least some of the participants met diagnostic criteria for a psychiatric disorder, no = none of the participants met diagnostic criteria for any psychiatric disorder, \* = groups are not matched with respect to psychopathology, <sup>o</sup> = no gold-standard diagnostic tool was named; PTSD = Posttraumatic Stress Disorder; psych. = psychosis; MDD = Major Depressive Disorder; CM = child maltreatment; n = sample size; quest. = questionnaire; CPS = Child Protective Services; MCS = Maltreatment Classification System; MMCI = Maternal Maltreatment Classification Interview; CECA-Q = Childhood Experiences of Care and Abuse Questionnaire; self-dev. = self-developed (the authors used a scale developed by themselves); FPI = Florence Psychiatric Interview; CWA = Child Welfare Agency; CTQ = Childhood Trauma Questionnaire; ETI = Early Trauma Inventory; no instr. used = no instrument used; THS = Trauma History Screen; C-DIS-IV = Computerized version of the Diagnostic Interview Schedule-IV; SSR = Social Services Record; NEMESIS CTI = Netherlands Mental Health Survey and Incidence Study Childhood Trauma Interview; spec. = specification (authors applied a specific definition of CM); N = neglect; EA = emotional abuse; PA = physical abuse; SA = sexual abuse; <sup>o</sup> = The grouping of participants was not just based on CM experiences but also included other traumatic experiences; awak. = awakening; Rel. = reliability, whereby the following definitions have been used: 1 = cortisol assessed over only one day, 2 = cortisol assessed over two days, 3 = cortisol assessed over more than two days; all = morning, evening, delta (evening minus morning value). <sup>a</sup> The two groups CPS-involved, stayed with birth parents and CPS-involved, placed in foster care were combined into one group. <sup>b</sup> This article did not appear in the initial search, but was suggested by the respective author as best suited for citation in this meta-analysis. <sup>c</sup> The two groups early physical/sexual abuse and maltreated without early abuse were combined into one group. <sup>d</sup> Comparison between patient groups with and without early traumata. <sup>e</sup> The two groups Multidimensional Treatment Foster Care for Preschoolers (MTFC) and Regular Foster Care (RFC) were combined into one group. <sup>f</sup> This article did not appear in the initial search, but was suggested by the respective author as best suited for citation in this meta-analysis. <sup>g</sup> Authors also administered ETI, but grouping was based on CTQ. <sup>h</sup> Percent non-Caucasians refers to total sample (N = 127). <sup>i</sup> Deprived adoptees were combined into one group and were compared to the non-deprived UK adoptees. <sup>j</sup> Comparison between PTSD patients with childhood sexual or physical abuse and those reporting no history of childhood sexual or physical abuse. <sup>k</sup> Comparison between MDD patients with early life stress and those without corresponding experiences. <sup>l</sup> This article did not appear in the initial search, but was suggested by the respective author as best suited for citation in this meta-analysis. <sup>m</sup> Exclusion of repressed memory group; CM sample: recovered and continuous memory group, control sample: control group. <sup>n</sup> The data were regrouped including only the traumatized control subjects and the non-traumatized control subjects (remark: some in the control sample may have experienced other traumatic events). <sup>o</sup> The two groups some neglect and severe neglect were combined into one group. <sup>p</sup> The data are from the Netherlands Study of Depression and Anxiety, a large cohort study; this article did not appear in the initial search, but was suggested by the respective author as best suited for citation in this meta-analysis. <sup>q</sup> Patient group: dysthymia, MDD, social phobia, panic with/without agoraphobia, generalized anxiety disorder. <sup>r</sup> No one of the participants fulfilled the diagnostic criteria for dysthymia, MDD, social phobia, panic with/without agoraphobia, or generalized anxiety disorder in the past 6 months. <sup>s</sup> It is not clearly stated in the text or in the subheading of the figure whether means and standard deviations or means and standard errors were presented; we assumed standard errors.

**Table 2**Sum of  
different cortisol outcome indices.

Outcome indices	Random-effects model					Heterogeneity measures					
	<i>N</i>	<i>k</i>	Hedges' <i>g</i>	95% CI	<i>p</i>	tau <sup>2</sup>	<i>I</i> <sup>2</sup>	<i>Q</i>	<i>p</i>	Pred. int.	Egger's Test <sup>a</sup>
<b>Morning cortisol</b>	5212	26	-0.02	-0.11; 0.06	0.586	0.01	29.1%	35.27	0.083	-0.25; 0.21	-0.268 (0.514)
Morning cortisol sensitivity analysis *	3646	24	-0.00	-0.09; 0.08	0.939	0.01	21.5%	29.29	0.171	-0.21; 0.20	-
<b>Evening cortisol</b>	5188	26	0.10	0.03; 0.18	0.008	0.01	2.0%	25.51	0.434	-0.08; 0.29	0.621 (0.067)
Evening cortisol sensitivity analysis *	2949	23	0.11	0.02; 0.20	0.013	0.01	0.0%	18.92	0.650	-0.07; 0.29	-
<b>DSL cortisol</b>	3390	17	0.00	-0.11; 0.11	0.987	0.01	23.5%	20.92	0.181	-0.28; 0.28	0.246 (0.609)
DSL cortisol sensitivity analysis *	2259	16	0.01	-0.13; 0.14	0.933	0.02	27.9%	20.80	0.143	-0.34; 0.36	-

Note. DSL = diurnal slope; 95% CI = 95% confidence interval; Pred. int. = prediction interval. <sup>a</sup> Intercept and (*p* values) displayed. Morning cortisol: \* exclusion of influential studies Bernard et al., 2010 and Vreeburg et al., 2009a. Evening cortisol: \* exclusion of influential studies Bernard et al., 2010, Lovallo et al., 2019 and Vreeburg et al., 2009a. DSL cortisol: \* exclusion of influential study Vreeburg et al., 2009a. See Appendix C Figs. 1.1 to 1.3 for corresponding forest and funnel plots.

**Table 3**

Summary characteristics of included studies that reported on cortisol assessed in the context of the cortisol awakening response (CAR).

Study	Total Sample				CM Sample			Control Sample			CM Assessment				Cortisol Assessment					Data			
	N	Age M(SD)	Sex	Ethn.	Psycho-path.	n	Age M(SD)	Sex	n	Age M(SD)	Sex	Method	Instrument	Grouping	Type	Type of sample	Unit	Time points <sup>a</sup>	Time point peak	Rel.	Indices	Source	Data (re)grouped
Bicanic et al., 2013	89	15.89 (1.86)	100.0	NA	yes*, PTSD	52	16.10 (1.98)	100.0	37	15.60 (1.65)	100.0	self-report (interview)	self-dev.	other (spec.)	SA	saliva	nmol/l	awak., 15, 30, 45, 60min	NA	1	AUC <sub>g</sub>	paper, table	no
Fuchs et al., 2017	82	35.38 (5.07)	100.0	NA	NA	37	36.47 (4.92)	100.0	45	34.48 (5.06)	100.0	self-report (quest.)	CTQ	cut-offs	PA, SA	saliva	nmol/l	awak., 30min	30min	2	awak., peak	provided, group	no
Kaess et al., 2017	66	14.91 (0.43)	42.4	NA	NA	15	14.79 (0.44)	53.3	51	14.95 (0.42)	39.2	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., 30, 60min	NA	2	awak., AUC <sub>g</sub>	provided, raw	yes
Keeshin et al., 2014	36	14.97 (1.40)	100.0	NA	yes*, mixed	24	15.04 (1.45)	100.0	12	14.84 (1.34)	100.0	informant (record)	medical record	other (record)	SA	saliva	log	awak., 30min	30min	3	awak., peak, delta	paper, table	no
Klaassens et al., 2009	22	47.58 (11.70)	100.0	NA	no	10	47.80 (12.10)	100.0	12	47.40 (11.90)	100.0	self-report (quest.)	CTQ <sup>b</sup>	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., 30, 45, 60min	30min	2	awak., peak, 60min, AUC <sub>g</sub>	paper, figure & table	no
Klaus et al., 2018	103	34.54 (10.80)	0.0	0.0	no <sup>o</sup>	39 <sup>c</sup>	35.16 (10.85)	0.0	64 <sup>c</sup>	34.17 (10.83)	0.0	self-report (quest.)	CTES	cut-offs	PA, SA, other <sup>o</sup>	saliva	nmol/l	awak., 30min	30min	2	awak., peak, delta	paper, figure & table	no
Kuhlman et al., 2015	79	12.73 (2.27)	40.5	22.0 <sup>d</sup>	yes, mixed	60	12.91 (2.26)	41.2	19	12.16 (2.27)	36.8	informant (quest.)	ETI	other (spec.)	EA, PA, SA	saliva	µg/dl	awak., 45min	45min	2	awak., peak, delta	provided, group	yes
Kumsta et al., 2017	57	24.02 (0.85)	61.4	NA	NA	44 <sup>e</sup>	24.15 (0.90)	68.2	13	23.55 (0.45)	38.5	informant (NA)	no instr. used	institutionalization	N	saliva	nmol/l	awak., 30, 45min	45min	2	awak., peak, delta	provided, group	no
Kuras et al., 2017	61	33.80 (2.30)	49.2	NA	no <sup>o</sup>	29	NA	NA	32	NA	NA	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., 30, 60min	30, 60min	2	awak., peak, 60min	paper, figure <sup>+</sup>	no
Li et al., 2015	75	37.92 (11.67)	72.0	46.7	yes*, MDD	38	39.70 (10.90)	68.4	37	36.10 (12.30)	75.7	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	µg/dl	awak., 15, 30, 60min	30min	1	awak., peak, 60min, AUC <sub>g</sub>	paper, figure	no
Lu et al., 2016a	35	23.80 (3.89)	51.4	100.0	yes, MDD	18	23.70 (4.13)	44.4	17	23.90 (3.75)	58.8	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	ng/ml	awak., 30, 45, 60min	30min	1	awak., peak, 60min, AUC <sub>g</sub> , AUC <sub>i</sub>	paper, figure	no
Lu et al., 2016b	45	21.65 (3.46)	62.2	100.0	no	23	21.50 (3.91)	60.9	22	21.80 (3.01)	63.6	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	ng/ml	awak., 30, 45, 60min	30min	1	awak., peak, 60min, AUC <sub>g</sub> , AUC <sub>i</sub>	paper, figure	no
Mello et al., 2015 <sup>f</sup>	63	10.63 (2.15)	47.6	NA	NA	25	10.40 (2.06)	56.0	38	10.79 (2.22)	42.1	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	µg/dl	awak., 30min	0, 30min	1	awak., peak, delta	provided, raw	yes
Monteleone et al., 2018a	44	26.50 (8.10)	100.0	NA	yes, AN	24	27.90 (8.70)	100.0	20	21.80 (7.10)	100.0	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., 15, 30, 60min	15, 60min	1	all	provided, group	no
Monteleone et al., 2018b	36	28.90 (8.70)	100.0	NA	yes, BN	22	31.00 (9.40)	100.0	14	25.70 (6.80)	100.0	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., 15, 30, 60min	30min	1	all	provided, group	no
Peng et al., 2014a	58	28.61 (7.32)	46.6	NA	yes, MDD	28	28.87 (6.28)	46.4	30	28.37 (8.27)	46.7	self-report (quest.)	CTQ	cut-offs	N	saliva	nmol/l	awak., 30min	30min	2	delta	paper, table <sup>+</sup>	no
Peng et al., 2014b	51	28.09 (4.69)	47.1	NA	no	22	28.37 (5.28)	45.5	29	27.87 (4.28)	48.3	self-report (quest.)	CTQ	cut-offs	N	saliva	nmol/l	awak., 30min	30min	2	delta	paper, table <sup>+</sup>	no
Quevedo et al., 2017	55	14.91 (1.64)	63.6	50.9	yes*, MDD <sup>g</sup>	35	14.84 (1.55)	77.1	20	15.04 (1.83)	40.0	self-report (interview)	K-SADS-P	other (spec.)	PA, SA, N	saliva	nmol/l	awak., 30, 60min	30min	3	awak., peak, 60min, delta	paper, table	no
Reichl et al., 2016a	25	16.25 (1.12)	96.0	NA	no	3	16.55 (1.34)	100.0	22	16.20 (1.12)	95.5	self-report (interview)	CECA	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., 30, 60min	30, 60min	3	all	provided, raw	yes
Reichl et al., 2016b	24	16.32 (1.33)	91.7	NA	yes, mixed	16	16.65 (1.32)	93.8	8	15.65 (1.14)	87.5	self-report (interview)	CECA	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., 30, 60min	30min	3	all	provided, raw	yes

**Table 3**

Summary characteristics of included studies that reported on cortisol assessed in the context of the cortisol awakening response (CAR).

Study	Total Sample					CM Sample			Control Sample			CM Assessment				Cortisol Assessment					Data		
	N	Age M(SD)	Sex	Ethn.	Psychopath.	n	Age M(SD)	Sex	n	Age M(SD)	Sex	Method	Instrument	Grouping	Type	Type of sample	Unit	Time points <sup>a</sup>	Time point peak	Rel.	Indices	Source	Data (re)grouped
Schreuder et al., 2016	114	26.75 (2.85)	45.6	NA	yes, mixed	16	27.81 (2.90)	43.8	98	26.58 (2.81)	45.9	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	µg/dl	awak., 30, 60min	30min	1	awak., peak, 60min, delta, AUC <sub>g</sub>	provided, group	yes
Smeets et al., 2007	22 <sup>b</sup>	40.50 (11.63)	100.0	NA	no <sup>o</sup>	13	38.77 (11.38)	100.0	9	43.00 (12.21)	100.0	self-report (interview)	self-dev.	other (spec.)	SA	saliva	nmol/l	awak., 15, 30, 45min	30min	2	awak., peak, delta, AUC <sub>g</sub> , AUC <sub>i</sub>	provided, raw	no
Stuedte et al., 2013	30 <sup>i</sup>	37.87 (12.04)	90.0	0.0	no	4	42.50 (11.39)	100.0	26	37.15 (12.19)	88.5	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	nmol/l	awak., 30min	30min	1	awak., peak, delta	provided, group	yes
van der Vegt et al., 2009	529	30.83 (1.22)	33.1	NA	NA	235	30.87 (1.24)	51.9	294	30.80 (1.20)	18.0	informant (interview)	self-dev.	other (spec.)	N	saliva	nmol/l	awak., 30min	30min	1	awak., peak	paper, table	no
van Zuiden et al., 2011	317	30.19 (9.69)	0.0	NA	NA	152	30.20 (9.64)	0.0	165	30.17 (9.76)	0.0	self-report (quest.)	ETI-SF	other (spec.)	EA, PA, SA	saliva	nmol/l	awak., 15, 30, 60min	30min	1	awak., peak, 60min, delta, AUC <sub>g</sub>	provided, group	yes
Vreeburg et al., 2009a <sup>i</sup>	996	42.35 (12.34)	65.9	NA	yes, mixed <sup>k</sup>	565	43.68 (11.82)	69.6	431	40.60 (12.79)	61.0	self-report (interview)	NEMESIS CTI	other (spec.)	EA, PA, SA, N	saliva	nmol/l	awak., 30, 45, 60min	30min	1	all	provided, group	yes
Vreeburg et al., 2009b <sup>i</sup>	431	42.71 (14.62)	59.4	NA	no <sup>l</sup>	106	47.97 (12.54)	69.8	325	41.00 (14.86)	56.0	self-report (interview)	NEMESIS CTI	other (spec.)	EA, PA, SA, N	saliva	nmol/l	awak., 30, 45, 60min	30min	1	all	provided, group	yes

*Note.* N = sample size; M = mean; SD = standard deviation; The sex ratio is indicated as percentage of female participants; Ethn. = ethnicity; The ethnicity ratio is indicated as percentage of non-Caucasians; NA = not assessed; Psychopath. = psychopathology, whereby the following definitions have been used: yes = at least some of the participants met diagnostic criteria for a psychiatric disorder, no = none of the participants met diagnostic criteria for any psychiatric disorder, \* = groups are not matched with respect to psychopathology, ° = no gold-standard diagnostic tool was named; PTSD = Posttraumatic Stress Disorder; MDD = Major Depressive Disorder; AN = Anorexia Nervosa; BN = Bulimia Nervosa; CM = child maltreatment; n = sample size; quest. = questionnaire; self-dev. = self-developed (the authors used a scale developed by themselves); CTQ = Childhood Trauma Questionnaire; CTES = Childhood Traumatic Events Scale; ETI = Early Trauma Inventory; no instr. used = no instrument used; K-SADS-P = Kiddie Schedule for Affective Disorders and Schizophrenia for School-Aged Children Present Version; CECA = Childhood Experiences of Care and Abuse; SF = short form; NEMESIS CTI = Netherlands Mental Health Survey and Incidence Study Childhood Trauma Interview; spec. = specification (authors applied a specific definition of CM); SA = sexual abuse; PA = physical abuse; EA = emotional abuse; N = neglect; ° = The grouping of participants was not just based on CM experiences but also included other traumatic experiences; awak. = awakening; Rel. = reliability, whereby the following definitions have been used: 1 = cortisol assessed over only one day, 2 = cortisol assessed over two days, 3 = cortisol assessed over more than two days; AUC<sub>g</sub> = area under the curve with respect to ground; delta = peak minus awakening levels; AUC<sub>i</sub> = area under the curve with respect to increase. <sup>a</sup> Time points cortisol was sampled. <sup>b</sup> Authors also administered ETI, but grouping was based on CTQ. <sup>c</sup> ELS AA/AG and ELS GG were combed into one group and compared to no ELS AA/AG and no ELS GG group. <sup>d</sup> Percent non-Caucasians refers to total sample (N = 127). <sup>e</sup> Deprived adoptees were combined into one group and were compared to the non-deprived UK adoptees. <sup>f</sup> Raw data were provided; participants with sampling adherence of +/- 5 min were included. <sup>g</sup> Patients were matched with respect to MDD but not with respect to PTSD. <sup>h</sup> Exclusion of repressed memory group; CM sample: recovered and continuous memory group, control sample: control group. <sup>i</sup> The data were regrouped including only the traumatized control subjects and the non-traumatized control subjects (remark: some in the control sample may have experienced other traumatic events). <sup>j</sup> The data are from the Netherlands Study of Depression and Anxiety, a large cohort study. The article did not appear in the initial search, but was suggested by the respective author as best suited for citation in this meta-analysis. <sup>k</sup> Patient group: dysthymia, MDD, social phobia, panic with/without agoraphobia, generalized anxiety disorder. <sup>l</sup> No one of the participants fulfilled the diagnostic criteria for dysthymia, MDD, social phobia, panic with/without agoraphobia, or generalized anxiety disorder in the past 6 months. <sup>m</sup> It is not clearly stated in the text or in the subheading of the figure whether means and standard deviations or means and standard errors were presented; we assumed standard errors.

**Table 4**

Summary statistics for random-effects models of included studies that reported on cortisol assessed in the context of the cortisol awakening response (CAR), displayed separately for the different cortisol outcome indices.

Outcome indices	Random-effects model					Heterogeneity measures					
	<i>N</i>	<i>k</i>	Hedges' <i>g</i>	95% CI	<i>p</i>	tau <sup>2</sup>	<i>I</i> <sup>2</sup>	<i>Q</i>	<i>p</i>	Pred. int.	Egger's Test <sup>a</sup>
<b>Awakening cortisol</b>	3342	24	-0.04	-0.13; 0.05	0.334	0.01	17.3%	27.82	0.223	-0.24; 0.15	-0.732 (0.054)
Awakening cortisol sensitivity analysis *	2271	22	-0.05	-0.14; 0.04	0.277	0.00	0.0%	19.15	0.576	-0.14; 0.05	-
<b>Peak cortisol</b>	3275	23	-0.02	-0.19; 0.15	0.812	0.10	58.3%	52.72	0.000	-0.71; 0.67	-0.636 (0.254)
Peak cortisol sensitivity analysis °	3204	21	0.02	-0.09; 0.13	0.701	0.02	36.5%	31.48	0.049	-0.28; 0.32	-
Peak cortisol sensitivity analysis *	3239	22	0.04	-0.10; 0.17	0.596	0.04	43.5%	37.16	0.016	-0.38; 0.45	-
<b>60 min post awakening cortisol</b> <sup>°</sup>	2276	14	0.02	-0.13; 0.17	0.798	0.03	41.6%	22.26	0.052	-0.36; 0.40	-0.095 (0.876)
60min post awakening cortisol sensitivity analysis °	2240	13	0.05	-0.07; 0.16	0.429	0.01	18.2%	14.66	0.261	-0.18; 0.28	-
<b>Delta cortisol</b> <sup>°</sup>	2536	18	-0.03	-0.17; 0.11	0.719	0.03	41.8%	29.23	0.033	-0.42; 0.36	-0.562 (0.287)
Delta cortisol sensitivity analysis °*	2500	17	0.02	-0.06; 0.10	0.622	0.00	0.0%	15.34	0.500	-0.07; 0.11	-
<b>AUC<sub>g</sub> cortisol</b>	2327	15	-0.13	-0.41; 0.15	0.370	0.22	73.1%	51.99	0.000	-1.18; 0.92	-0.963 (0.243)
AUC <sub>g</sub> cortisol sensitivity analysis °	2211	12	-0.18	-0.40; 0.05	0.118	0.08	61.9%	28.87	0.002	-0.88; 0.52	-
<b>AUC<sub>i</sub> cortisol</b>	1654	9	-0.07	-0.48; 0.33	0.723	0.28	70.3%	26.96	0.001	-1.41; 1.26	-0.359 (0.719)

Note. 95% CI = 95% confidence interval; Pred. int. = prediction interval. <sup>a</sup> Intercept and (*p* values) displayed. Awakening cortisol: \* exclusion of influential studies Li et al., 2015 and Vreeburg et al., 2009a. Peak cortisol: ° exclusion of outlier studies Lu et al., 2016a and Monteleone et al., 2018b; \* exclusion of influential study Monteleone et al., 2018b. ° Due to convergence problems associated with the Fisher scoring algorithm, the DerSimonian-Laird estimator was used to estimate  $\tau^2$ . 60 min post awakening cortisol: ° exclusion of outlier study Monteleone et al., 2018b. Delta cortisol: °\* exclusion of outlier and influential study Monteleone et al., 2018b. AUC<sub>g</sub> cortisol: ° exclusion of outlier studies Lu et al., 2016a, Lu et al., 2016b and Monteleone et al., 2018b. See Appendix C Figs. 2.1 to 2.6 for corresponding forest and funnel plots.

**Table 5**  
Summary characteristics of included studies that reported on cortisol assessed in the context of a stressor (cortisol stress reactivity).

Study	Total Sample					CM Sample			Control Sample			CM Assessment			Stressor				Cortisol Assessment			Data		
	N	Age M(SD)	Sex	Ethn.	Psycho-path.	n	Age M(SD)	Sex	n	Age M(SD)	Sex	Method	Instru-ment	Grouping	Type	Task	Duration stressor	Time point peak	Cortisol response	Type of sample	Unit	Indices	Source	Data (re)grouped
Alexander et al., 2018	200	23.72 (2.85)	50.0	0.0	no	32	23.81 (3.01)	56.3	168	23.70 (2.83)	48.8	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	TSST	15min	25min peak ident.	yes, both	saliva	nmol/l	all	provided, group	yes
Ali and Pruessner, 2012	37	25.77 (5.37)	51.4	NA	no	20	25.95 (5.61)	50.0	17	25.56 (5.22)	52.9	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	TSST	20min	30min peak ident.	yes, both	saliva	nmol/l	all	provided, group	no
Buchmann et al., 2014	195	19.00 (0.00)	53.9	NA	yes*, mixed	15	19.00 (0.00)	53.3	180	19.00 (0.00)	53.9	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	TSST	20min	20min peak ident.	yes, both	blood	ng/ml	bl, peak, end, delta	provided, group	yes
Carpenter et al., 2007	50	29.11 (11.31)	66.0	24.0	no	23	35.00 (12.90)	74.0	27	24.10 (6.60)	59.0	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	TSST	25min	30min peak ident.	yes, both	blood	nmol/l (adj. for age)	bl, peak, end, delta	paper, table	no
Carpenter et al., 2011	110	30.45 (11.13)	100.0	NA	yes*, mixed	20	36.50 (12.50)	100.0	90	29.10 (10.40)	100.0	self-report (quest.)	CTQ	cut-offs	PA	TSST	25min	45min peak ident.	yes, both	saliva	nmol/l	bl, peak, end	paper, figure+	no
Cook et al., 2012	175	15.36 (1.01)	51.8	NA	NA	86	15.39 (1.03)	55.4	89	15.30 (0.97)	48.2	self-report (quest.)	CTQ	cut-offs	EA, PA, N	TSST-C	15min	33min NA	yes, both	saliva	NA	delta	paper, table	no
England-Mason et al., 2017	120	32.20 (4.26)	100.0	NA	NA	25	30.47 (4.36)	100.0	95	32.65 (4.14)	100.0	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	stroop task	17min	37min peak ident.	yes, only in CG sample	saliva	nmol/l	all	provided, group	yes
Fogelman et al., 2016a	73 <sup>a</sup>	63.78 (7.39)	58.9	8.5 <sup>c</sup>	no	24	65.00 (9.21)	66.7	49	63.18 (6.33)	55.1	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	TSST	15min	25min peak ident.	yes, both	saliva	nmol/l	all	provided, group	yes
Fogelman et al., 2016b	85 <sup>b</sup>	23.86 (7.32)	0.0	4.5 <sup>d</sup>	no	20	24.65 (7.51)	0.0	65	23.62 (7.30)	0.0	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	TSST	15min	25min peak ident.	yes, both	saliva	nmol/l	all	provided, group	yes
Harkness et al., 2011	71	15.39 (2.11)	67.6	11.0	yes*, mixed	26	15.88 (1.75)	69.2	45	15.11 (2.26)	66.7	self-report (interview)	CECA	cut-offs	EA, PA, SA, N	TSST	30min	30min peak ident.	yes, both	saliva	µg/dl	bl, peak, end, delta, AUC <sub>i</sub>	provided, group	yes
Heim et al., 2000a	26	29.77 (6.37)	100.0	38.5	yes*, mixed	14	30.21 (5.46)	100.0	12	29.25 (7.52)	100.0	self-report (interview)	ETI	other (spec.)	PA, SA	TSST	20min	30min peak ident.	yes, both	blood	nmol/l	peak	paper, table	no
Heim et al., 2000b	23	33.35 (6.00)	100.0	13.0	yes*, MDD <sup>e</sup>	13	32.38 (7.64)	100.0	10	34.60 (8.38)	100.0	self-report (interview)	ETI	other (spec.)	PA, SA	TSST	20min	30, 45min peak not ident.	yes, both	blood	nmol/l	peak	paper, table	no
Hengesch et al., 2018	44	22.15 (2.91)	56.8	NA	NA	22 <sup>f</sup>	22.50 (3.04)	63.6	22 <sup>f</sup>	21.80 (2.80)	50.0	mixed	no instr. used	other (self-ident., agency)	insti-tute. <sup>g</sup>	CPT + PASAT	3min	25min peak ident.	yes, both	saliva	nmol/l	bl, delta, AUC <sub>i</sub>	paper, table	no
Ivanov et al., 2011	25	10.55 (1.02)	36.0	88.0	yes, ADHD	10	10.78 (1.12)	50.0	15	10.39 (0.95)	26.7	mixed (quest.)	CLES, PTSRI	other (spec.)	PA, SA, loss <sup>h</sup>	watching emot. video	11min	end of viewing• no peak	no	saliva	µg/dl	bl, peak, end, AUC <sub>g</sub> , AUC <sub>i</sub>	provided, raw	no
Kaiser et al., 2018	55	27.62 (6.43)	100.0	43.0 <sup>e</sup>	yes, mixed	42	28.43 (6.65)	100.0	13	25.00 (5.03)	100.0	self-report (interview)	TAQ	cut-offs	PA, SA, WV, peers <sup>g</sup>	MAST	10.75min	28min peak ident.	yes, both	saliva	ng/ml	bl, peak, end, delta, AUC <sub>g</sub>	provided, group	yes
Kuhlman et al., 2015	91	12.75 (2.27)	40.3	22.0 <sup>h</sup>	yes, mixed	72	12.91 (2.26)	41.2	19	12.16 (2.27)	36.8	informant (quest.)	ETI	other (spec.)	EA, PA, SA	SE-CPT	5min	25min peak ident.	yes, both	saliva	µg/dl	all	provided, group	yes
Lovallo et al., 2019 <sup>i</sup>	699	23.70 (3.14)	55.9	20.3	no	335	23.90 (3.26)	63.0	364	23.50 (3.02)	49.5	self-report (interview)	C-DIS-IV	other (spec.)	EA, PA, SA	public speaking, arithmet. task	45min	30min peak ident.	yes, both	saliva	µg/dl	bl, peak, end, delta	provided, group	yes



**Table 5**  
Summary characteristics of included studies that reported on cortisol assessed in the context of a stressor (cortisol stress reactivity).

Luecken et al., 2009	76 <sup>j</sup>	18.90 (0.97)	48.7	25.0	NA	19	NA	NA	48	NA	NA	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	role-play task	10min	10min peak ident.	yes, both	saliva	µg/dl	all	provided, raw	yes
Luecken and Appelhans, 2006	88	19.30 (1.63)	62.5	22.7	NA	16	19.81 (1.83)	56.3	72	19.18 (1.57)	63.9	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	speech task	8min	23min peak ident.	yes, only in CG sample	saliva	µg/dl	all	provided, raw	yes
Martinson et al., 2016	50	19.70 (4.19)	100.0	11.8	yes*, mixed	26	21.38 (7.59)	100.0	24	18.92 (1.14)	100.0	self-report (mixed)	self-dev., SCID-I/P	other (spec.)	SA	self-discl. and rel. - build. task	45min	15min peak ident.	yes, only in CG sample	saliva	nmol/l	bl, peak, end	paper, table	no
Mielock et al., 2017	52	27.44 (9.47)	100.0	61.5	yes, MDD	26	26.20 (9.50)	100.0	26	28.70 (9.40)	100.0	self-report (interview)	CAI	cut-offs	PA, SA, WV	TSST	15min	25, 35min peak not ident.	yes, both	saliva	µg/dl	all	provided, group	no
O'Connor et al., 2018	145	26.93 (9.39)	60.7	26.8	yes, mixed	65	29.95 (10.15)	57.8	80	24.24 (7.59)	63.7	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	MAST	10min	20, 30min peak not ident.	yes, both	saliva	nmol/l	all	provided, group	yes
Otte et al., 2005	76	28.00 (5.00)	13.2	NA	no	16	27.00 (4.00)	6.0	60	28.00 (5.00)	15.0	self-report (interview)	LSC-R	other (spec.)	PA, N, non-int. trauma <sup>o</sup>	watching emot. video	20min	40min peak ident.	yes, both	saliva	ng/dl	bl, peak	paper, figure	no
Ouellet-Morin et al., 2011	190	12.00 (0.00)	49.5	8.4	NA	64	12.00 (0.00)	54.7	126	12.00 (0.00)	46.8	informant (interview)	self-develop.	other (spec.)	PA, SA, bully. <sup>o</sup>	PST	13min	25, 35min peak not ident.	yes, both	saliva	nmol/l	bl, peak	paper, figure	no
Ouellet-Morin et al., 2018	155	24.10 (3.70)	0.0	NA	NA	56	24.00 (3.60)	0.0	99	24.10 (3.70)	0.0	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	TSST	15min	25min peak ident.	yes, both	saliva	µg/dl	all	provided, group	no
Pierrehumbert et al., 2009	44 <sup>k</sup>	33.02 (7.12)	100.0	NA	yes, mixed	27	33.38 (6.01)	100.0	17	32.46 (8.87)	100.0	self-report (interview)	ETI	other (spec.)	SA	TSST	15min	NA peak not ident.	yes, both	saliva	nmol/l	bl, delta, AUC <sub>i</sub>	paper, table	no
Rao and Morris, 2015	17 <sup>l</sup>	18.93 (4.72)	52.9	35.3	yes, MDD	7	18.40 (3.90)	42.9	10	19.30 (5.40)	60.0	mixed (interview)	CAI	cut-offs	PA, SA, WV	TSST	15min	25min peak ident.	yes, only in CG sample	saliva	µg/dl	bl, peak, end, AUC <sub>g</sub>	paper, table & figure	no
Schalinski et al., 2015	33 <sup>m</sup>	34.60 (10.40)	100.0	88.0	yes, mixed	16	34.30 (11.65)	100.0	17	34.90 (9.40)	100.0	self-report (interview)	ETI	other (clustering)	EA, PA, SA	trauma interview	99min	99min peak ident.	yes, only in CG sample	saliva	nmol/l	all	provided, group	yes
Schwaiger et al., 2016	60	52.02 (5.09)	66.7	NA	no	30	52.57 (5.52)	66.7	30	51.47 (4.64)	66.7	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	TSST	15min	25min peak ident.	yes, both	blood	ng/ml	all	provided, raw	no
Seltzer et al., 2014	73	9.54 (1.15)	52.1	47.0	NA	37 <sup>n</sup>	9.40 (1.16)	54.1	36 <sup>n</sup>	9.68 (1.14)	50.0	informant (mixed)	CPS, CTS	other (record)	PA	TSST-C	16min	31min peak ident.	yes, only in CG sample	saliva	µg/dl	all	provided, raw	no
Shenk et al., 2014	104	16.98 (1.18)	100.0	59.0	NA	47	16.75 (1.10)	100.0	57	17.13 (1.21)	100.0	informant (record)	CPS	other (record)	PA, SA, N	perform. and interp. task	15.45min	25.45min• no peak	no	saliva	µg/dl	bl, peak, end, AUC <sub>g</sub> , AUC <sub>i</sub>	provided, group	no
Shenk et al., 2010	144	18.09 (3.47)	100.0	45.8	NA	60	18.54 (3.62)	100.0	84	17.76 (3.34)	100.0	informant (record)	CPS	other (record, spec.)	SA	cognitive task	10min	15min peak ident.	yes, both	saliva	µg/dl	AUC <sub>i</sub>	paper, table	no
Sullivan et al., 2013	64	6.89 (0.95)	52.0	88.0	NA	30	6.98 (0.88)	48.0	34	6.72 (1.08)	56.0	informant (mixed)	CPS, CTS-PC	other (record)	N	self-eval. task	4min	20min after last failed task• no peak	no	saliva	µg/dl	bl, peak, end	paper, table	no
Sumner et al., 2014	158	14.93 (1.39)	53.8	47.0	NA	61	15.30 (1.30)	60.7	97	14.70 (1.40)	49.5	self-report (mixed)	CTQ, CECA	cut-offs,	EA, PA, SA	TSST	15min	20min peak ident.	yes, both	saliva	nmol/l	bl, peak, end, AUC <sub>i</sub>	paper, figure	no

**Table 5**  
Summary characteristics of included studies that reported on cortisol assessed in the context of a stressor (cortisol stress reactivity).

Author	Year	M	SD	Sex	Psychopath.	n	M	SD	Sex	Ethn.	M	SD	Instrument	Group	Other	EA, PA, SA, N	Stressor	Time	Peak	Sample	Unit	Analysis	Match		
Suzuki et al., 2014a	41	44.89 (13.06)	58.5	39.0	no	17	44.30 (12.50)	52.9	24	45.30 (13.70)	62.5	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	watching emot. pictures	NA	NA	NA	yes, only in CM sample	saliva	nmol/l	bl, peak, end, delta	paper, figure <sup>a</sup>	no
Suzuki et al., 2014b	38	51.91 (11.30)	71.1	0.0	yes, MDD	20	52.10 (12.00)	80.0	18	51.70 (10.80)	61.1	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	watching emot. pictures	NA	NA	NA	yes, both	saliva	nmol/l	bl, peak, end, delta	paper, figure <sup>a</sup>	no
Trickett et al., 2014	454	10.93 (1.16)	46.7	88.7	NA	303	10.84 (1.16)	50.0	151	11.11 (1.15)	40.0	informant (record)	DCFS	other (record)	EA, PA, SA, N	TSST-C	14min	24min	24min	yes, both	saliva	µg/dl	bl, peak, end	paper, table	no
Wingenfeld et al., 2017a	59	34.43 (12.02)	100.0	NA	no	22	36.50 (12.20)	100.0	37	33.20 (11.90)	100.0	self-report (interview)	ETI <sup>o</sup>	other (spec.)	PA, SA	TSST	15min	40min	40min	yes, both	saliva	nmol/l	bl, peak, AUC <sub>g</sub> , AUC <sub>i</sub>	paper, table & figure	no
Wingenfeld et al., 2017b	84	35.00 (11.21)	100.0	NA	yes, MDD	32	34.20 (10.50)	100.0	52	35.50 (11.70)	100.0	self-report (interview)	ETI <sup>o</sup>	other (spec.)	PA, SA	TSST	15min	40min	40min	yes, both	saliva	nmol/l	bl, peak, AUC <sub>g</sub> , AUC <sub>i</sub>	paper, table & figure	no

*Note.* N = sample size; M = mean; SD = standard deviation; The sex ratio is indicated as percentage of female participants; Ethn. = ethnicity; The ethnicity ratio is indicated as percentage of non-Caucasians; NA = not assessed; Psychopath. = psychopathology, whereby the following definitions have been used: yes = at least some of the participants met diagnostic criteria for a psychiatric disorder, no = none of the participants met diagnostic criteria for any psychiatric disorder, \* = groups are not matched with respect to psychopathology; MDD = Major Depressive Disorder; ADHD = Attention Deficit Hyperactivity Disorder; CM = child maltreatment; n = sample size; quest. = questionnaire; CTQ = Childhood Trauma Questionnaire; CECA = Childhood Experiences of Care and Abuse; ETI = Early Trauma Inventory; no instr. used = no instrument used; CLES = Codington Life Events Scale; PTSRI = Posttraumatic Stress Reaction Index; TAQ = Traumatic Antecedents Questionnaire; C-DIS-V = Computerized version of the Diagnostic Interview Schedule-IV; self-dev. = self-developed (the authors used a scale developed by themselves); SCID-I/P = Structured Clinical Interview for DSM-IV-TR axis I disorders, research version, patient edition; CAI = Childhood Adversity Interview; LSC-R = Life Stressor Checklist-Revised; CPS = Child Protective Services; CTS-PC = Parent-Child Conflict Tactics scale; DCFS = Department of Children and Family Services; spec. = specification (authors applied a specific definition of CM); self-ident. = self-identification (participants self-identified as victims of CM); EA = emotional abuse; PA = physical abuse; SA = sexual abuse; N = neglect; institute. = institutionalization; non-intent. trauma = non-intentional trauma; bully = bullying; <sup>o</sup> = The grouping of participants was not just based on CM experiences but also included other traumatic experiences; TSST = Trier Social Stress Test; TSST-C = Trier Social Stress Test for Children; CPT = Cold Pressor Test; PASAT = Paced Auditory Serial Addition Task; emot. = emotional; MAST = Maastricht Acute Stress Test; SE-CPT = Socially Evaluated Cold Pressor Task; arithmet. = arithmetic's; self-discl. = self-disclosure; rel.-build. = relationship-building; PST = Psychosocial Stress Test; perform. = performance; interper. = interpersonal; self-eval. = self-evaluative; ident. = identical; adj. = adjusted; bl = baseline (if possible, value just prior stressor onset was extracted); delta = peak values minus baseline value; AUC<sub>g</sub> = Area under the curve with respect to ground; AUC<sub>i</sub> = Area under the curve with respect to increase; • no cortisol peak was observed in both groups, therefore the time point closest to 25min post stressor onset was extracted. <sup>a</sup> Community-dwelling older adult sample. <sup>b</sup> Community-dwelling younger adult sample. <sup>c</sup> Percent non-Caucasians refers to total sample (N = 82). <sup>d</sup> Percent non-Caucasians refers to total sample (N = 88). <sup>e</sup> CM sample and control sample were matched with respect to MDD but not with respect to PTSD. <sup>f</sup> Female and male participants with early life adversity were combined into one group and were compared to female and male control participants. <sup>g</sup> Percent non-Caucasians refers to total sample (N = 42). <sup>h</sup> Percent non-Caucasians refers to total sample (N = 127). <sup>i</sup> This article did not appear in the initial search, but was suggested by the respective author as best suited for citation in this meta-analysis. <sup>j</sup> Demographic data for total sample (age, percent females, percent Non-Caucasians) refers to N = 76, cortisol data available for N = 67. <sup>k</sup> Demographic data for total sample and CM and control sample refers to N = 44, cortisol data available for N = 35. <sup>l</sup> Comparison between participants with MDD and CM and those with MDD without CM. <sup>m</sup> The data were regrouped, taking into account only the patient group with stress-related disorders. <sup>n</sup> Female and male maltreated participants were combined into one group and were compared to female and male control participants. <sup>o</sup> Authors also administered CTQ, but grouping was based on ETI. <sup>p</sup> It is not clearly stated in the text or in the subheading of the figure whether means and standard deviations or means and standard errors were presented; we assumed standard errors.

**Table 6**

Summary statistics for random-effects models of included studies that reported on cortisol assessed in the context of a stressor (cortisol stress reactivity), displayed separately for the different cortisol outcome indices.

Outcome indices	Random-effects model					Heterogeneity measures					
	<i>N</i>	<i>k</i>	Hedges' <i>g</i>	95% CI	<i>p</i>	tau <sup>2</sup>	<i>I</i> <sup>2</sup>	<i>Q</i>	<i>p</i>	Pred. int.	Egger's Test <sup>a</sup>
<b>baseline cortisol</b>	3895	35	-0.12	-0.28; 0.04	0.152	0.18	69.7%	112.38	0.000	-0.99; 0.75	-0.587 (0.411)
baseline cortisol sensitivity analysis °	3780	32	-0.07	-0.19; 0.05	0.278	0.06	52.9%	65.87	0.000	-0.59; 0.46	-
baseline cortisol sensitivity analysis *	3845	34	-0.07	-0.21; 0.06	0.290	0.09	59.9%	82.23	0.000	-0.71; 0.56	-
<b>peak cortisol</b>	3867	35	-0.27	-0.51; -0.02	0.033	0.47	78.3%	156.87	0.000	-1.68; 1.14	-1.088 (0.175)
peak cortisol sensitivity analysis °	3731	31	-0.18	-0.29; -0.08	0.001	0.03	42.6%	52.30	0.007	-0.57; 0.20	-
peak cortisol sensitivity analysis *	3817	34	-0.18	-0.33; -0.03	0.019	0.13	65.6%	95.81	0.000	-0.94; 0.58	-
<b>delta cortisol</b>	2678	24	-0.19	-0.32; -0.06	0.004	0.04	46.1%	42.66	0.008	-0.63; 0.25	-1.475 (0.016)
delta cortisol sensitivity analysis °*	2640	23	-0.14	-0.25; -0.04	0.010	0.02	30.7%	31.76	0.082	-0.44; 0.15	-
<b>recovery cortisol</b>	3407	29	-0.28	-0.48; -0.07	0.008	0.24	73.4%	105.22	0.000	-1.30; 0.74	-1.603 (0.038)
recovery cortisol sensitivity analysis °	3254	25	-0.14	-0.24; -0.04	0.004	0.01	19.1%	29.65	0.197	-0.39; 0.11	-
recovery cortisol sensitivity analysis *	3357	28	-0.19	-0.32; -0.06	0.004	0.06	55.5%	60.65	0.000	-0.69; 0.31	-
<b>AUC<sub>g</sub> cortisol</b>	1614	20	-0.17	-0.35; 0.02	0.081	0.11	62.5%	50.72	0.000	-0.89; 0.56	-1.596 (0.257)
AUC <sub>g</sub> cortisol sensitivity analysis °	1591	19	-0.20	-0.37; -0.03	0.021	0.08	55.7%	40.62	0.002	-0.81; 0.41	-
<b>AUC<sub>i</sub> cortisol</b>	1992	23	-0.22	-0.34; -0.10	0.000	0.03	31.4%	32.06	0.076	-0.58; 0.14	-1.458 (0.156)

Note. 95% CI = 95% confidence interval; Pred. int. = prediction interval. <sup>a</sup> Intercept and (*p* values) displayed. Baseline cortisol: ° exclusion of outlier studies Carpenter et al., 2007, Ivanov et al., 2011, Suzuki et al., 2014a; \* exclusion of influential study Carpenter et al., 2007. Peak cortisol: ° exclusion of outlier studies Carpenter et al., 2007, Heim et al., 2000b, Ivanov et al., 2011, Suzuki et al., 2014b; \* exclusion of influential study Carpenter et al., 2007. Delta cortisol: °\* exclusion of outlier and influential study Suzuki et al., 2014b. Recovery cortisol: ° exclusion of outlier studies Ali & Pruessner, 2012, Carpenter et al., 2007, Ivanov et al., 2011, Suzuki et al., 2014a; \* exclusion of influential study Carpenter et al., 2007. AUC<sub>g</sub> cortisol: ° exclusion of outlier study Ivanov et al., 2011. See Appendix C Figs. 3.1 to 3.6 for corresponding forest and funnel plots.

**Table 7**

Summary characteristics of included studies that reported on cortisol assessed in the context of the dexamethasone suppression test (DST).

Study	Total Sample				CM Sample			Control Sample			CM Assessment				Cortisol Assessment					Data		
	N	Age M(SD)	Sex	Ethn.	Psychopath.	n	Age M(SD)	Sex	n	Age M(SD)	Sex	Method	Instrument	Grouping	Type	Type of sample	Unit	Time points <sup>a</sup>	Indices <sup>b</sup>	Dex-dose <sup>c</sup>	Source	Data (re)grouped
Baes et al., 2014	20 <sup>d</sup>	38.80 (9.84)	75.0	40.0	yes, MDD	13	39.50 (9.73)	76.9	7	37.40 (11.38)	71.4	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	blood	NA	9am	post	0.5mg 10pm	paper, text	no
Carvalho et al., 2012a	19	27.05 (6.21)	94.7	NA	yes, BPD	14	27.86 (6.59)	92.9	5	24.80 (4.92)	100.0	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	nmol/l	7:30am	pre, post, delta	0.5mg 11pm	provided, raw	yes
Carvalho et al., 2012b	25	32.72 (8.78)	52.0	NA	yes, MDD	12	34.42 (9.55)	66.7	13	31.15 (8.06)	38.5	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	nmol/l	7:30am	pre, post, delta	0.5mg 11pm	provided, raw	yes
Carvalho et al., 2012c	40	32.70 (10.39)	67.5	NA	no	8	42.63 (6.41)	75.0	32	30.22 (9.73)	65.6	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	nmol/l	7:30am	pre, post, delta	0.5mg 11pm	provided, raw	yes
Duval et al., 2004	28	15.95 (1.93)	82.1	0.0	yes*, PTSD	14	16.20 (1.90)	85.7	14	15.70 (2.00)	78.6	self-report (interview)	SCID-I, checklist	other (self-ident.)	SA	blood	nmol/l	8am	pre, post	1.0mg 11pm	paper, table & figure	no
Faravelli et al., 2010	93 <sup>e</sup>	43.60 (15.40)	50.2	NA	yes, mixed	32	NA	NA	61	NA	NA	self-report (mixed)	CECA, self-dev.	other (not spec.)	PA, SA, loss <sup>o</sup>	saliva	nmol/l	8am	pre, post	0.5mg 11pm	paper, table	no
Lindley et al., 2004	16 <sup>f</sup>	40.00 (11.17)	87.5	6.2	yes, PTSD	9	34.56 (9.28)	88.9	7	47.00 (9.78)	85.7	self-report (quest.)	THS	other (spec.)	PA, SA	saliva	µg/dl	8am	pre, post, delta	0.5mg 10pm	provided, raw	no
Lu et al., 2016a	35	23.80 (3.89)	51.4	100.0	yes, MDD	18	23.70 (4.13)	44.4	17	23.90 (3.75)	58.8	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	blood	ng/ml	8am	pre, post	1.0mg 11pm	paper, table	no
Lu et al., 2016b	45	21.65 (3.46)	62.2	100.0	no	23	21.50 (3.91)	60.9	22	21.80 (3.01)	63.6	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	blood	ng/ml	8am	pre, post	1.0mg 11pm	paper, table	no
Mehta et al., 2011a	63 <sup>g</sup>	41.76 (13.12)	68.3	87.5	yes, mixed	22	42.23 (11.53)	72.7	41	41.51 (14.03)	65.9	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	blood	µg/dl	8-9am	pre, post, delta	0.5mg 11pm	provided, group	yes
Mehta et al., 2011b	32 <sup>g</sup>	40.63 (12.85)	71.9	90.5	yes, PTSD	18	41.06 (13.48)	77.8	14	40.07 (12.47)	64.3	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	blood	µg/dl	8-9am	pre, post, delta	0.5mg 11pm	provided, group	yes
Newport et al., 2004a	38	30.90 (7.00)	100.0	36.8	yes*, mixed	19	33.30 (6.40)	100.0	19	28.50 (6.90)	100.0	self-report (interview)	ETI	other (spec.)	PA, SA	blood	µg/dl	8am	pre, post	0.5mg 11pm	paper, table	no
Newport et al., 2004b	26	32.98 (7.87)	100.0	19.2	yes, MDD	16	32.40 (7.80)	100.0	10	33.90 (8.30)	100.0	self-report (interview)	ETI	other (spec.)	PA, SA	blood	µg/dl	8am	pre, post	0.5mg 11pm	paper, table	no
Schreuder et al., 2016	114	26.75 (2.85)	45.6	NA	yes, mixed	16	27.81 (2.90)	43.8	98	26.58 (2.81)	45.9	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	saliva	µg/dl	awak.	pre, post, delta	1.0mg 10pm	provided, group	yes
Stein et al., 1997	40	31.47 (6.70)	100.0	NA	yes*, mixed	19	32.20 (6.70)	100.0	21	30.80 (6.80)	100.0	self-report (interview)	self-dev.	other (spec.)	SA	blood	nmol/l	8am	pre, post	0.5mg 11pm	paper, table	no
Vreeburg et al., 2009a <sup>h</sup>	1112	42.71 (12.24)	65.7	NA	yes, mixed <sup>i</sup>	635	43.92 (11.69)	70.7	477	41.10 (12.78)	59.1	self-report (interview)	NEMESIS CTI	other (spec.)	EA, PA, SA, N	saliva	nmol/l	awak.	pre, post, delta	0.5mg 11pm	provided, group	yes
Vreeburg et al., 2009b <sup>h</sup>	476	42.93 (14.49)	60.7	NA	no <sup>j</sup>	116	47.88 (12.36)	72.4	360	41.34 (14.78)	56.9	self-report (interview)	NEMESIS CTI	other (spec.)	EA, PA, SA, N	saliva	nmol/l	awak.	pre, post, delta	0.5mg 11pm	provided, group	yes

Note. N = sample size; M = mean; SD = standard deviation; The sex ratio is indicated as percentage of female participants; Ethn. = ethnicity; The ethnicity ratio is indicated as percentage of non-Caucasians; NA = not assessed; Psychopath. = psychopathology, whereby the following definitions have been used: yes = at least some of the participants met diagnostic criteria for a psychiatric disorder, no = none of the participants met diagnostic criteria for any psychiatric disorder, \* = groups are not matched with respect to psychopathology; MDD = Major Depressive Disorder; BPD = Borderline Personality Disorder; PTSD = Posttraumatic Stress Disorder; CM = child maltreatment; n = sample size; quest. = questionnaire; CTQ = Childhood Trauma Questionnaire; SCID-I = Structured Clinical Interview for DSM-IV Axis I Disorders; CECA = Childhood Experiences of Care and Abuse; self-dev. = self-developed (the authors used a scale developed by themselves); THS = Trauma History Screen; ETI = Early Trauma Inventory; NEMESIS CTI = Netherlands Mental Health Survey and Incidence Study Childhood Trauma Interview; self-ident. = self-identification (participants self-identified as victims of CM); spec. = specification (authors applied a specific definition of CM); EA = emotional abuse; PA = physical abuse; SA = sexual abuse; N = neglect; <sup>o</sup> The grouping of participants was not just based on the experience of CM but also included loss experiences; awak. = awakening; Dex-dose = dose of dexamethasone administered to participants. <sup>a</sup> Timepoints of cortisol sampling. <sup>b</sup> Indices: pre = cortisol assessed prior dexamethasone administration; post = cortisol assessed after dexamethasone administration, delta = post values minus pre values. <sup>c</sup> Dose and timepoint of dexamethasone administration. <sup>d</sup> Comparison between depressed patients with and without early life stress. <sup>e</sup> Comparison between patient groups with and without early trauma. <sup>f</sup> Comparison between PTSD patients with childhood sexual or physical abuse and those reporting no history of childhood sexual or physical abuse. <sup>g</sup> The sample represents a highly traumatized, low-income cohort; data provided separately for those patients all fulfilling diagnostic criteria for PTSD and those patients without PTSD. <sup>h</sup> The data are from the Netherlands Study of Depression and Anxiety, a large cohort study; this article did not appear in the initial search, but was suggested by the respective author as best suited for citation in this meta-analysis. <sup>i</sup> Patient group: dysthymia, MDD, social phobia, panic with/without

Journal Pre-proofs

**Table 8**

Summary statistics for random-effects models of included studies that reported on cortisol assessed in the context of the dexamethasone suppression test (DST), displayed separately for the different cortisol outcome indices.

Outcome indices	Random-effects model					Heterogeneity measures					
	<i>N</i>	<i>k</i>	Hedges' <i>g</i>	95% CI	<i>p</i>	tau <sup>2</sup>	<i>I</i> <sup>2</sup>	<i>Q</i>	<i>p</i>	Pred. int.	Egger's Test <sup>a</sup>
<b>pre-DST cortisol</b>	2201	16	0.07	-0.09; 0.23	0.402	0.03	40.5%	25.21	0.047	-0.35; 0.49	-0.096 (0.854)
pre-DST cortisol sensitivity analysis <sup>°</sup>	2163	15	0.12	-0.01; 0.25	0.074	0.01	0.0%	13.50	0.487	-0.14; 0.38	-
pre-DST cortisol sensitivity analysis*	613	14	0.03	-0.19; 0.24	0.816	0.05	32.0%	19.10	0.120	-0.53; 0.58	-
<b>post-DST cortisol</b>	2214	17	0.01	-0.18; 0.20	0.936	0.06	44.1%	28.64	0.027	-0.56; 0.57	-0.483 (0.343)
<b>delta DST cortisol</b>	1888	9	-0.13	-0.31; 0.06	0.178	0.02	34.0%	12.12	0.146	-0.54; 0.29	-0.642 (0.317)

Note. 95% CI = 95% confidence interval; Pred. int. = prediction interval. <sup>a</sup> Intercept and (*p* values) displayed. <sup>°</sup> Exclusion of outlier study Newport et al., 2004a. \* Exclusion of influential studies Vreeburg et al., 2009a and Vreeburg et al., 2009b. See Appendix C Figs. 4.1.1 to 4.1.3 for corresponding forest and funnel plots.

**Table 9**

Summary characteristics of included studies that reported on cortisol assessed in the context of the combined dexamethasone-corticotropin releasing hormone (Dex-CRH) test.

Study	Total Sample					CM Sample			Control Sample			CM Assessment				Cortisol Assessment					Data		
	N	Age M(SD)	Sex	Ethn.	Psycho-path.	n	Age M(SD)	Sex	n	Age M(SD)	Sex	Method	Instru-ment	Grouping	Type	Type of sample	Unit	Time points	Indices	Dex-dose	CRH-dose	Source	Data (re)grouped
Carpenter et al., 2009	230	29.23 (10.37)	56.5	NA	no	41	35.90 (11.60)	66.0	189	27.78 (9.52)	54.0	self-report (quest.)	CTQ	cut-offs	EA	blood	nmol/l <sup>a</sup>	2:59pm - 5pm (6 samples)	bl, peak	1.5mg 11pm	100µg 3pm	paper, figure <sup>+</sup>	no
Heim et al., 2008a	21	31.64 (8.76)	0.0	57.1	yes, MDD	15	32.30 (8.70)	0.0	6	30.00 (9.50)	0.0	self-report (interview)	ETI	other (spec.)	PA, SA	blood	µg/dl	2pm - 5pm (9 samples)	peak, delta, AUC <sub>g</sub> , AUC <sub>i</sub>	1.5mg 11pm	1µg/kg 3pm	paper, table	no
Heim et al., 2008b	28	30.25 (8.44)	0.0	53.6	yes*, mixed	14	31.40 (8.00)	0.0	14	29.10 (9.00)	0.0	self-report (interview)	ETI	other (spec.)	PA, SA	blood	µg/dl	2pm - 5pm (9 samples)	peak, delta, AUC <sub>g</sub> , AUC <sub>i</sub>	1.5mg 11pm	1µg/kg 3pm	paper, table	no
Klaassens et al., 2009	22	47.58 (11.70)	100.0	NA	no	10	47.80 (12.10)	100.0	12	47.40 (11.90)	100.0	self-report (quest.)	CTQ <sup>b</sup>	cut-offs	EA, PA, SA, N	blood	nmol/l	3pm - 4:45pm (7 samples)	AUC <sub>g</sub>	1.5mg 11pm	100µg 3:02pm	paper, table	no
Lee et al., 2012	24	33.42 (9.72)	41.7	41.7	yes*, PD	10	39.10 (11.37)	40.0	14 <sup>c</sup>	29.36 (5.93)	42.9	self-report (quest.)	CTQ	other (thirds)	EA, PA, SA, N	blood	µg/dl	4pm - 6pm (7 samples) <sup>d</sup>	all	1.5mg 11pm	1µg/kg 4pm	provided, raw	no
Spitzer et al., 2018a	86	35.50 (11.26)	100.0	NA	yes, MDD	35	35.20 (10.90)	100.0	51	35.70 (11.60)	100.0	self-report (mixed)	CTQ, ETI	other (spec.)	PA, SA	saliva	nmol/l	1:30pm - 4:30pm (7 samples)	delta <sup>e</sup> , AUC <sub>g</sub>	1.5mg 11pm	100µg 2:30pm	paper, table	no
Spitzer et al., 2018b	58	34.29 (11.85)	100.0	NA	no	21	34.80 (10.90)	100.0	37	34.00 (12.50)	100.0	self-report (mixed)	CTQ, ETI	other (spec.)	PA, SA	saliva	nmol/l	1:30pm - 4:30pm (7 samples)	delta <sup>e</sup> , AUC <sub>g</sub>	1.5mg 11pm	100µg 2:30pm	paper, table	no
Watson et al., 2007	40 <sup>f</sup>	48.63 (7.34)	47.5	NA	yes, mixed	25	49.12 (6.83)	52.0	15	47.80 (8.32)	40.0	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	blood	nmol/l	3pm - 6pm (9 samples)	all	1.5mg 11pm	100µg 3pm	provided, raw	yes
De Bellis et al., 1994 <sup>g</sup>	26	11.65 (2.74)	100.0	42.3	yes*, Dysthymia	13	11.20 (2.60)	100.0	13	12.10 (2.90)	100.0	informant (record)	CPS	other (record)	SA	blood	nmol/l	6pm - 10pm (13 samples)	AUC <sub>g</sub> , AUC <sub>i</sub>	no dex adm.	1µg/kg 8pm	paper, text	no
Kaufman et al., 1997 <sup>g</sup>	26 <sup>h</sup>	9.75 (1.16)	53.8	34.6	yes, MDD	13	9.60 (1.40)	53.9	13	9.90 (0.90)	53.9	mixed	PSS, medical records	other (not spec.)	EA, PA, SA <sup>i</sup>	blood	µg/dl	5pm - 8pm (9 samples)	bl <sup>j</sup> , peak, AUC <sub>g</sub> , AUC <sub>i</sub>	no dex adm.	1µg/kg 5:30pm	paper, table	no

Note. N = sample size; M = mean; SD = standard deviation; The sex ratio is indicated as percentage of female participants; Ethn. = ethnicity; The ethnicity ratio is indicated as percentage of non-Caucasians; NA = not assessed; Psychopath. = psychopathology, whereby the following definitions have been used: yes = at least some of the participants met diagnostic criteria for a psychiatric disorder, no = none of the participants met diagnostic criteria for any psychiatric disorder, \* = groups are not matched with respect to psychopathology; MDD = Major Depressive Disorder; PD = Personality Disorder; CM = child maltreatment; n = sample size; quest. = questionnaire; CTQ = Childhood Trauma Questionnaire; ETI = Early Trauma Inventory; CPS = Child Protective Services; PSS = Psychosocial Schedule for School Aged Children; spec. = specification (authors applied a specific definition of CM); EA = emotional abuse; PA = physical abuse; SA = sexual abuse; N = neglect; bl = baseline (if possible, sample just prior CRH injection was extracted); AUC<sub>g</sub> = area under the curve with respect to ground; AUC<sub>i</sub> = area under the curve with respect to increase. Dex-dose = dose of dexamethasone administered to participants; CRH-dose = dose of corticotropin-releasing hormone administered to participants. <sup>a</sup> All extracted values are adjusted for age, gender, and effects of four other maltreatment subtypes. <sup>b</sup> Authors also administered ETI, but grouping was based on CTQ. <sup>c</sup> The low CTQ PD group and the normal control group were combined and compared to the high CTQ PD group. <sup>d</sup> Sample assessed at -150min prior CRH injection was not included in analyses (therefore 7 instead of 8 samples; baseline sample at 4pm). <sup>e</sup> Delta defined as difference between maximum value after CRH injection and the mean of the three baseline measures. <sup>f</sup> The data were regrouped, taking into account only the patient groups. <sup>g</sup> No dexamethasone was administered (CRH test only). <sup>h</sup> Comparison between the patient groups with and without abuse. <sup>i</sup> 54% were subjected to ongoing EA. <sup>j</sup> Baseline defined as the mean of the three pre CRH infusion samples. <sup>+</sup> It is not clearly stated in the text or in the subheading of the figure whether means and standard deviations or means and standard errors were presented; we assumed standard errors.

**Table 10**

Summary statistics for random-effects models of included studies that reported on cortisol assessed in the context of the combined dexamethasone-corticotropin releasing hormone (Dex-CRH) test, displayed separately for the different cortisol outcome indices.

Outcome indices	Random-effects model					Heterogeneity measures					
	<i>N</i>	<i>k</i>	Hedges' <i>g</i>	95% CI	<i>p</i>	tau <sup>2</sup>	<i>I</i> <sup>2</sup>	<i>Q</i>	<i>p</i>	Pred. int.	Egger's Test <sup>a</sup>
<b>pre-CRH cortisol</b>	320	4	-0.08	-0.35; 0.18	0.540	0	0.0%	1.32	0.725	-0.66; 0.50	-1.434 (0.070)
pre CRH cortisol (Dex-CRH studies only)	294	3	-0.04	-0.32; 0.24	0.782	0	0.0%	0.54	0.763	-1.86; 1.78	-
pre CRH cortisol sensitivity analysis <sup>*</sup>	90	3	-0.26	-0.68; 0.16	0.228	0	0.0%	0.21	0.902	-3.00; 2.48	-
<b>peak cortisol</b>	369	6	0.11	-0.39; 0.61	0.668	0.25	68.3%	15.77	0.008	-1.45; 1.66	3.018 (0.106)
peak cortisol (Dex-CRH studies only)	343	5	0.03	-0.53; 0.60	0.906	0.28	69.0%	12.89	0.012	-1.89; 1.95	-
<b>delta cortisol</b>	257	6	-0.05	-0.42; 0.33	0.804	0.10	50.3%	10.05	0.074	-1.08; 0.98	2.038 (0.393)
<b>AUC<sub>g</sub> cortisol</b>	329	9	-0.09	-0.45; 0.26	0.602	0.16	54.8%	17.72	0.023	-1.12; 0.93	1.775 (0.379)
AUC <sub>g</sub> cortisol (Dex-CRH studies only)	277	7	-0.12	-0.56; 0.33	0.606	0.22	61.0%	15.39	0.018	-1.46; 1.23	-
AUC <sub>g</sub> cortisol sensitivity analysis <sup>o</sup>	308	8	-0.21	-0.47; 0.05	0.106	0.02	25.3%	9.37	0.227	-0.69; 0.27	-
<b>AUC<sub>i</sub> cortisol</b>	163	6	0.18	-0.38; 0.74	0.522	0.31	63.3%	13.64	0.018	-1.56; 1.93	4.839 (0.410)
AUC <sub>i</sub> cortisol (Dex-CRH studies only)	111	4	0.21	-0.56; 0.99	0.590	0.45	68.8%	9.60	0.022	-3.14; 3.56	-

Note. 95% CI = 95% confidence interval; Pred. int. = prediction interval. <sup>a</sup> Intercept and (*p* values) displayed. <sup>\*</sup> Exclusion of influential study Carpenter et al., 2009. <sup>o</sup> Exclusion of outlier study Heim et al., 2008a. See Appendix C Figs. 4.2.1 to 4.2.2 for corresponding forest and funnel plots.



**Table 11**

Summary characteristics of included studies that reported on hair cortisol concentrations (HCC).

Study	Total Sample					CM Sample			Control Sample			CM Assessment				Cortisol Assessment		Data	
	N	Age M(SD)	Sex	Ethn.	Psycho-path.	n	Age M(SD)	Sex	n	Age M(SD)	Sex	Method	Instrument	Grouping	Type	Unit	Source	Data (re)grouped	
do Prado et al., 2017	57	15.39 (1.81)	57.9	NA	no	30	16.47 (1.25)	50.0	27	14.19 (1.57)	66.7	self-report (quest.)	CTQ	other (percentile)	EA, PA, SA, N	pg/mg	paper, figure	no	
Fischer et al., 2017	135	50.31 (14.73)	71.9	9.0 <sup>a</sup>	NA	43	53.10 (13.90)	81.0	92	49.00 (15.00)	67.0	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	pg/mg	provided, group	yes	
Groër et al., 2016	81 <sup>b</sup>	46.17 (10.53)	100.0	45.7	NA	27	47.30 (10.70)	100.0	54	45.60 (10.50)	100.0	self-report (quest.)	self-dev.	other (spec.)	SA	ng/mg	paper, table	no	
Morris et al., 2017	22 <sup>c</sup>	24.71 (3.38)	100.0	18.2	no	7	22.60 (2.30)	100.0	15	25.70 (3.40)	100.0	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	pg/mg	provided, group	no	
Reichl et al., 2016a	24	16.27 (1.10)	91.7	NA	no	3	16.55 (1.34)	100.0	21	16.23 (1.09)	90.5	self-report (interview)	CECA	cut-offs	EA, PA, SA, N	pg/mg	provided, raw	yes	
Reichl et al., 2016b	25	16.27 (1.30)	96.0	NA	yes, mixed	17	16.56 (1.30)	100.0	8	15.65 (1.14)	87.5	self-report (interview)	CECA	cut-offs	EA, PA, SA, N	pg/mg	provided, raw	yes	
Schalinski et al., 2015	39 <sup>d</sup>	34.95 (10.74)	100.0	87.2	yes, mixed	17	33.70 (11.70)	100.0	22	35.91 (10.10)	100.0	self-report (interview)	ETI	other (clustering) <sup>e</sup>	EA, PA, SA	pg/mg	provided, group	yes	
Stedte et al., 2013	58 <sup>f</sup>	39.16 (13.40)	91.4	0	no	10	42.20 (11.22)	100.0	48	38.52 (13.83)	89.6	self-report (quest.)	CTQ	cut-offs	EA, PA, SA, N	pg/mg	provided, group	yes	
White et al., 2017	537	9.98 (3.13)	50.7	2.2	NA	245	9.86 (3.24)	46.1	292	10.08 (3.03)	54.5	informant (mixed)	MCS (CPS), MMCI	other (record & spec.)	EA, PA, SA, N	pg/mg	provided, group	no	

*Note.* N = sample size; M = mean; SD = standard deviation; The sex ratio is indicated as percentage of female participants; Ethn. = ethnicity; The ethnicity ratio is indicated as percentage of non-Caucasians; NA = not assessed; Psychopath. = psychopathology, whereby the following definitions have been used: yes = at least some of the participants met diagnostic criteria for a psychiatric disorder, no = none of the participants met diagnostic criteria for any psychiatric disorder; CM = child maltreatment; n = sample size; quest. = questionnaire; CTQ = Childhood Trauma Questionnaire; self-dev. = self-developed (the authors used a scale developed by themselves); CECA = Childhood Experiences of Care and Abuse; ETI = Early Trauma Inventory; MCS = Maltreatment Classification System; CPS = Child Protective Services; MMCI = Maternal Maltreatment Classification Interview; spec. = specification (authors applied a specific definition of CM); EA = emotional abuse; PA = physical abuse; SA = sexual abuse; N = neglect. <sup>a</sup>Percent non-Caucasians refers to total sample (N = 139). <sup>b</sup>Three outliers in hair cortisol were removed, however unclear to which group this applied. <sup>c</sup>Comparison between women with abuse/neglect without recent interpersonal violence exposure and non-trauma controls. <sup>d</sup>The data were regrouped, taking into account only the patient group with stress-related disorders. <sup>e</sup>Grouping was based on k-means clustering method based on ETI sum score excluding the general trauma subscale (remark: the sample represents a refugee's sample with all having experienced some type of trauma during their life). <sup>f</sup>The data were regrouped, taking into account only the traumatized control subjects and the non-traumatized control subjects (remark: some in the control sample may have experienced other traumatic events); data provided for a total sample of N = 58 instead of N = 53 as presented in the paper.

**Table 12**

Summary characteristics of included studies that reported on 24-hour urinary free cortisol (24-hour UFC).

Study	Total Sample					CM Sample			Control Sample			CM Assessment				Cortisol Assessment		Data	
	N	Age M(SD)	Sex	Ethn.	Psycho-path.	n	Age M(SD)	Sex	n	Age M(SD)	Sex	Method	Instrument	Grouping	Type	Unit	Source	Data (re)grouped	
De Bellis et al., 1994	26	11.65 (2.74)	100.0	42.3	yes*, Dysthymia	13	11.20 (2.60)	100.0	13 <sup>a</sup>	12.10 (2.90)	100.0	informant (record)	CPS	other (record)	SA	nmol/m <sup>2</sup> x day	paper, table	no	
De Bellis et al., 1999	42 <sup>b</sup>	10.46 (1.22)	40.5	28.6	yes*, PTSD	18	10.40 (1.40)	44.4	24	10.50 (1.10)	37.5	mixed	CPS, trauma interview	other (spec.)	PA, SA, WV	µg/day (adj. for SES)	paper, table	no	
Lemieux and Coe, 1995	17 <sup>c</sup>	35.30 (6.30) <sup>d</sup>	100.0	NA	NA	8	NA	100.0	9	NA	100.0	self-report	no instrument used	other (self-ident.)	SA	µg/day (adj. for conc. creatinine & body weight)	paper, table	no	
Lemieux et al., 2008	25 <sup>e</sup>	31.25 (6.13)	100.0	12.0	yes*, MDD	13	31.30 (6.50)	100.0	12	31.20 (6.00)	100.0	self-report	no instrument used	other (self-ident.)	SA	µg/day	paper, table	no	

*Note.* N = sample size; M = mean; SD = standard deviation; The sex ratio is indicated as percentage of female participants; Ethn. = ethnicity; The ethnicity ratio is indicated as percentage of non-Caucasians; NA = not assessed; Psychopath. = psychopathology, whereby the following definitions have been used: yes = at least some of the participants met diagnostic criteria for a psychiatric disorder, no = none of the participants met diagnostic criteria for any psychiatric disorder, \* = groups are not matched with respect to psychopathology; PTSD = Posttraumatic Stress Disorder; MDD = Major Depressive Disorder; CM = child maltreatment; n = sample size; CPS = Child Protective Services; spec. = specification (authors applied a specific definition of CM); self-ident. = self-identification (participants self-identified as victims of CM); SA = sexual abuse; PA = physical abuse; WV = witnessing domestic violence; adj. = adjusted; SES = socioeconomic status; conc. = concentration. <sup>a</sup> 24-hour UFC available for n = 11. <sup>b</sup> Comparison between PTSD and control children. <sup>c</sup> Comparison between women who experienced childhood sexual abuse without PTSD and controls. <sup>d</sup> Age refers to total sample (N = 28) including PTSD group. <sup>e</sup> Comparison between women with a history of childhood sexual abuse without PTSD and controls; one cortisol specimen was lost due to technical error, however unclear to which group this applied.

**Table 13**

Summary table of main findings from meta-analyses, meta-regression, and subgroup analyses (mixed/fixed-effects models) for the various HPA axis activity measures and related outcome indices.

	<i>k</i>	<i>N</i>	Meta-analyses		Moderator analyses	
			Random-effects models	Between-study heterogeneity ( <i>I</i> <sup>2</sup> )	Significant moderators (meta-regression)	Significant comparisons (subgroup analyses)
<b>DC</b>	26	5248				
Morning cortisol	26	5212	no overall significant effect	low-moderate (29.1%)	tendency for higher morning cortisol in older-aged samples; tendency for lower morning cortisol with higher study quality (in terms of appropriate measure of DC)	reduced morning cortisol in studies (1) that used informant reports, (2) that applied CPS records or cut-offs, (3) where original data were extracted
Evening cortisol	26	5188	small overall positive effect <sup>o</sup>	low (2.0%)	-	larger positive effect size estimates in studies (1) focusing on other types of ELA; (2) including original data
DSL cortisol	17	3390	no overall significant effect	low (23.5%)	-	-
<b>CAR</b>	27	3545				
Awakening cortisol	24	3342	no overall significant effect	low (17.3%)	-	reduced awakening cortisol in studies where original data were extracted
Peak cortisol	22	3215	no overall significant effect	moderate-high (59.8%)	-	-
60 min post awakening cortisol	14	2276	no overall significant effect	low-moderate (41.6%)	-	-
Delta cortisol	18	2536	no overall significant effect	low-moderate (41.8%)	tendency for lower delta cortisol in studies with higher proportion of females	-
AUC <sub>g</sub> cortisol	15	2327	no overall significant effect	moderate-high (73.1%)	tendency for lower AUC <sub>g</sub> cortisol with higher study quality (in terms of appropriate measure of confounders)	reduced AUC <sub>g</sub> cortisol where original data were extracted
AUC <sub>i</sub> cortisol	9	1654	no overall significant effect	moderate-high (70.3%)	<i>k</i> < 10 (no meta-regression performed)	<i>k</i> < 10 (no subgroup analyses performed)
<b>Cortisol stress reactivity</b>	39	4284				
Baseline cortisol	35	3895	no overall significant effect	moderate-high (69.7%)	no meta-regression performed	no subgroup analyses performed
Peak cortisol	35	3867	small overall negative effect*	high (78.3%)	-	-
Recovery cortisol	29	3407	small overall negative effect*	moderate-high (73.4%)	tendency for lower recovery cortisol in studies with higher proportion of females	-
Delta cortisol	24	2678	small overall negative effect*	low-moderate (46.1%)	tendency for lower delta cortisol in studies (1) with higher proportion of females; (2) with a higher proportion of participants fulfilling diagnostic criteria for a mental disorder	stronger negative effect estimates in studies (1) that included clinical samples, (2) involved participants with other types of ELA, (3) that observed a cortisol response in only one of the two comparison groups
AUC <sub>g</sub> cortisol	20	1614	no overall significant effect	moderate-high (62.5%)	-	-
AUC <sub>i</sub> cortisol	23	1992	small overall negative effect*	low-moderate (31.4%)	tendency for higher AUC <sub>i</sub> cortisol in studies with higher study quality (in terms of selection of participants)	stronger negative effect estimates in studies that observed a cortisol response in only one of the two comparison groups
<b>DST</b>	17	2222				
Pre-DST cortisol	16	2201	no overall significant effect	low-moderate (40.5%)	no meta-regression performed	no subgroup analyses performed

Post-DST cortisol	17	2214	no overall significant effect	low-moderate (44.1%)	tendency for lower post-DST cortisol in studies with higher proportion of females	increased post-DST cortisol in studies focusing on other types of ELA
Delta	9	1888	no overall significant effect	low-moderate (34.0%)	$k < 10$ (no meta-regression performed)	$k < 10$ (no subgroup analyses performed)
<b>Dex-CRH test</b>	10	561				
Pre-CRH cortisol	4	320	no overall significant effect	low (0.0%)	$k < 10$ (no meta-regression performed)	$k < 10$ (no subgroup analyses performed)
Peak cortisol	6	369	no overall significant effect	moderate-high (68.3%)	$k < 10$ (no meta-regression performed)	$k < 10$ (no subgroup analyses performed)
Delta cortisol	6	257	no overall significant effect	moderate-high (50.3%)	$k < 10$ (no meta-regression performed)	$k < 10$ (no subgroup analyses performed)
AUC <sub>g</sub> cortisol	9	329	no overall significant effect	moderate-high (54.8%)	$k < 10$ (no meta-regression performed)	$k < 10$ (no subgroup analyses performed)
AUC <sub>i</sub> cortisol	6	163	no overall significant effect	moderate-high (63.3%)	$k < 10$ (no meta-regression performed)	$k < 10$ (no subgroup analyses performed)
<b>HCC</b>	9	978	no overall significant effect	moderate-high (53.3%)	$k < 10$ (no meta-regression performed)	$k < 10$ (no subgroup analyses performed)
<b>24-hour UFC</b>	4	108	no overall significant effect	high (79.7%)	$k < 10$ (no meta-regression performed)	$k < 10$ (no subgroup analyses performed)

*Note.* DC = diurnal cortisol; DSL = diurnal slope cortisol; CAR = cortisol awakening response; DST = dexamethasone suppression test; Dex-CRH test = combined dexamethasone-corticotropin releasing hormone test; HCC = hair cortisol concentrations; 24-hour UFC = 24-hour urinary free cortisol; CM = child maltreatment; AUC<sub>g</sub> = area under the curve with respect to ground; AUC<sub>i</sub> = area under the curve with respect to increase; ° Positive effect = overall increased cortisol levels in child maltreatment group compared to their respective control group. \* Negative effect = overall reduced cortisol levels in child maltreatment group compared to their respective control group. - No significant moderators or/and subgroup comparisons were identified.

## Highlights

- Changes in HPA axis functioning assumed as a result of child maltreatment (CM).
- This meta-analysis revealed blunted cortisol stress reactivity in CM exposed group.
- No overall effects were found for any other HPA axis activity measure.
- Sex and study quality accounted for some of the between-study heterogeneity.
- Methodological flaws of primary studies hamper comprehensive conclusion.

Journal Pre-proofs