Young children's cortisol levels at out-of-home child care: a meta-analysis

Vet, S.M. de; Vrijhof C.I.; Veek, S.M.C. van der; Linting, M.; Vermeer, H.J.

Early Childhood Research Quarterly 63 (2023) 204–218

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

journal homepage: www.elsevier.com/locate/ecresq

Young children's cortisol levels at out-of-home child care: A meta-analysis

Sanne M. de Vet, Claudia I. Vrijhof, Shelley M.C. van der Veek, Mariëlle Linting, Harriet J. Vermeer*

Education and Child Studies, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, the Netherlands

1. Introduction

Out-of-home child care is a familiar and necessary environment for many young children worldwide, particularly in economically advanced societies and for parents who work outside of the home (Organization for Economic Cooperation and Development, 2021). Child care has been shown to be beneficial to children's cognitive and social-emotional development (National Institute for Child Health and Human Development (NICHD), 2006), but the separation from parents and the complex interactions with peers and caregivers may also provoke physiological stress (Vermeer & Groeneveld, 2017). The potential stressfulness of the child care setting attracted attention, and many studies into this topic have emerged during the past decades (Gunnar, 2021; Gunnar et al., 1997). In the early days, research efforts to study children's stress at child care were mainly driven by more traditional gender roles and moral concerns regarding non-parental child care (e.g., Belsky, 1986). Nowadays, due to economic progress, changing gender roles, and mothers who have entered the work force more often, out-of-home child care has become an indispensable part of family life, and is also viewed as contributing to children's positive development (NICHD, 2006). Current research therefore focuses more on variations in stress at child care, and how we could improve the quality of child care for (different groups of) children in order to support their health and well-being. Therefore, it is of interest to examine how stressful child care is for young children and to understand factors related to this potential stress. The knowledge will thus add to our understanding of the meaning of child care for children's physiological well-being and related factors. This meta-analysis attempts to provide some evidence to inform child care practice and policy with regard to children's stress levels at child care.

1.1. Developmental view on stress

The most commonly used method to measure physiological stress in children is analyzing the level of the hormone cortisol at various times during the day (Vermeer & Groeneveld, 2017), collected via body material like urine, blood or saliva. Because of its non-invasive character, the collection of saliva is often preferred (Tryphonopoulos et al., 2014). Cortisol is the end product of the hypothalamic-pituitary-adrenocortical (HPA) axis, with the latter being involved in the regulation of physical and psychological stress. When stressors are absent, cortisol follows a typical pattern over the day (Tryphonopoulos et al., 2014). Such a typical pattern is characterized by a peak shortly after wakening, a sharp decrease during mid-morning, a gradual decrease from the afternoon onwards, and an evening nadir.

The typical diurnal cortisol pattern is not yet evident in newborns. Instead they display two peaks that are 12 h apart and depend on the time of birth instead of the time of day (Iwata et al., 2013; Sippell et al., 1978). As children grow older, the diurnal cortisol pattern matures in steps and gradually follows a more typical adult pattern (Gunnar & Donzella, 2002). The literature is inconclusive as to when children show a consistent diurnal cortisol patterns, with estimates ranging from the first months (De Weerth et al., 2003; Ivars et al., 2015) to the third or even fourth year of life (Gunnar & Donzella, 2002), depending on the research method and the specific definition of a consistent diurnal cortisol pattern. In general, absolute cortisol levels gradually decrease over the first year of life (Tollenaar et al., 2010), although intra- and inter-individual variability remain relatively large, especially up to eight months of age (De Weerth & Van Geert, 2002; Ivars et al., 2015; Tollenaar et al., 2010). This variability and the confounding influence of factors such as napping (Watamura et al., 2002) make the study of cortisol in young children rather complex, and therefore researchers are recommended to measure cortisol at different moments and in different settings within the same child (Gunnar & Donzella, 2002).

To a certain extent, increased cortisol secretion can be viewed as adaptive and beneficial from a developmental point of view (Tryphonopoulos et al., 2014), since it may stimulate children's learning. Persistent deviations from this pattern, however, in particular elevations in the afternoon, can be interpreted as an indicator of stress (Geoffroy et al., 2006; Vermeer & Van IJzendoorn, 2006). Although short-term carryover effects of increased cortisol levels during the day at child care into the evening at home were not found (Watamura et al., 2009), prolonged activation of the HPA-axis, especially during the first years of life, can pose a threat to or increase vulnerability of the developing brain and children's long-term health and well-being (Gunnar, 2021; Sonia et al., 2009; Tryphonopoulos et al., 2014), although the specific mechanisms remain unclear (Gunnar, 2021). Moreover, links between elevated cortisol and a suboptimal immune system have been found (Watamura et al., 2010). Finally, behaviorally inhibited children were shown to score higher on internalizing problems over time when they ex-

* Corresponding author.
E-mail address: vermeer@fsw.leidenuniv.nl (H.J. Vermeer).

https://doi.org/10.1016/j.ecresq.2022.12.015
Received 12 August 2021; Received in revised form 23 November 2022; Accepted 15 December 2022

perienced elevated cortisol levels at child care (Rosen & Schulkin, 1998). For many families, however, out-of-home child care is a necessity, and therefore the question that comes to mind is: under what conditions at child care do we see the smallest increase in children's stress levels (as measured via cortisol) compared to home, in order to optimize child care for children?

1.2. Factors related to cortisol elevations at child care

Frequently studied and theoretically relevant variables are quality of care (e.g., Watamura et al., 2009), child temperament (e.g., Gunnar et al., 2010), group size (e.g., Lisonbee et al., 2008) and hours at child care (e.g., Drugli et al., 2018). The relations between child care quality and hours at child care on the one hand and cortisol at child care on the other hand are quite well-established, with many studies finding lower child care quality and more hours at child care relating to higher cortisol levels (Vermeer & Groeneveld, 2017). Low child care quality is theoretically linked to cortisol elevations because the lack of sensitive caregiver responses hampers the co-regulation of the stress system of the child (Gunnar, 2021). The HPA-axis was indeed found to be socially regulated by caregivers, in particular by sensitive caregivers to whom the child is securely attached. More specifically, sensitive caregivers can buffer the impact of stressors on the child (Gunnar, 2021). More hours in child care are also thought to be linked to cortisol elevations, probably because of the endured and sometimes exhausting interactions at child care (Gunnar, 2021; Vermeer & Groeneveld, 2017). Outcomes may however differ depending on characteristics such as type of care (e.g., Groeneveld et al., 2010), child age (e.g., Bernard et al., 2015) and research design (e.g., Albers et al., 2016).

The associations between child temperament and cortisol and group size and cortisol secretion at child care show a more ambiguous picture (Vermeer & Groeneveld, 2017). Theoretically, a larger group size is linked indirectly to child care quality, because it is more challenging for caregivers to provide high-quality child care when more children are placed under their care, and because of increased noise levels (Vermeer & Groeneveld, 2017). Furthermore, children with a more difficult temperament are thought to be more susceptible for stressors and might therefore show higher cortisol elevations at child care (Tout et al., 1998; Watamura et al., 2003). In sum, there is a need for an updated meta-analysis.

1.3. Previous meta-analytic studies

More than a decade ago, two meta-analyses on cortisol secretion in out-of-home child care were published (Geoffroy et al., 2006; Vermeer & Van IJzendoorn, 2006). Both meta-analyses found children's cortisol levels to show the typical pattern when children were at home, whereas at child care cortisol levels increased during the day. In the meta-analysis by Vermeer & Van IJzendoorn (2006), five studies that were based on salivary cortisol were included, which showed a small to medium combined correlation between child care attendance and cortisol secretion (r =.23), only significant for children younger than 36 months of age. In the meta-analysis byGeoffroy and colleagues (2006), which included a slightly different set of studies due to divergent inclusion criteria concerning the date of publication and child care setting, the results of nine studies were combined. The authors reported a medium to large effect size ($d = 0.72$), with a larger increase over the day when children were at child care compared to when they were at home. The authors found the largest effects for studies with preschoolers, children with a more difficult temperament and children attending low-quality child care, which was however not meta-analytically tested, due to a limited number of studies at the time into stress in child care in general, let alone into specific potential correlates such as child care quality. Since the publication of the previous meta-analyses, many additional empirical studies on the topic have been published, also in countries other than the United States. Most of them were in line with the results of the meta-analyses, but others reported null results (e.g., Vermeer et al., 2010). Furthermore, in the earlier meta-analytic studies, analyses on the correlates of cortisol and the role of potential moderation by study variables were not possible, again due to the limited number of articles published at the time (Geoffroy et al., 2006; Vermeer & Van IJzendoorn, 2006). These analytical limitations and mixed results asked for a new meta-analysis into this topic.

1.4. Current study

The aims of the current study were to meta-analytically examine whether 1a) mid-morning cortisol values, 1b) mid-afternoon cortisol values, and 1c) the cortisol increase from mid-morning to mid-afternoon is higher within children when at child care compared to when at home; 2) an increase in cortisol from mid-morning to mid-afternoon at child care correlates with lower child care quality, more difficult child temperament, larger group size, and more hours at child care (all studied separately), and 3) certain study characteristics (year of publication, age group(s), research design, country, type of care, method of saliva collection, and the inclusion of covariates) moderate the outcomes of the former research questions in case of heterogeneity.

2. Method

2.1. Literature search

We conducted and reported the current study according to PRISMA guidelines for systematic reviews and meta-analyses (Moher et al., 2009). A systematic literature search was performed via the following electronic databases: EMBASE, Emcare, ERIC, ProQuest Dissertations and Theses, PsychINFO, PubMed, and Web of Science (in August 2019, but the databases were most recently checked for new studies in August 2020). We chose to use the terms *cortisol, child care* and synonyms of these terms (e.g., *hydrocortisone, day care*). Through the thesaurus (synonym dictionary) of the databases it was possible to see which specific synonyms of the terms yielded the most relevant results. In Table 1 in the Supplementary Materials, an overview of the search formulas per database can be found. The outcome of interest concerned cortisol levels, as measured in the saliva of children who are cared for in center-based or home-based child care settings. Studies on children with severe disabilities and studies on children who attended any kind of formal education (e.g., kindergarten, first grade) were excluded, because the focus was on the cortisol secretion of = children in regular preschool child care settings. Since the age when children start in formal education varies between countries (between 3 and 6 years), a strict maximum child age was not set. However, we did not include studies with children who already started primary school and attended after school child care settings. Only articles published in peer-reviewed journals, dissertations, and conference papers were included, written in English and without limit on date of publication. We came across one Korean and one Finnish article and asked the authors if these articles were available in English, which was unfortunately not the case. We are aware of possible accessibility, language, and publication bias and therefore at least the potential publication bias was taken into account in the analyses.

2.2. Selection of studies

Studies were included for the first research question on cortisol levels at child care versus home when raw (uncorrected) cortisol data (*M, SD*) were available for at least one mid-morning (between 08:30 AM and 11:30 AM) and/or one mid-afternoon (between 02:00 PM and 05:30 PM) sample of saliva per child. Samples had to be collected on at least one day at child care and at least one day at home. For the meta-analyses on correlates of cortisol at child care, we included studies if a correlation coefficient was available or, when absent, a standardized regression coeffi-

Table 1
Characteristics of the studies included in the meta-analyses.

Author(s) and year of publication	Country	Sample size	Design of study and covariates (yes/no)	Type of care	Child age range (M, SD) and group	Location(s) of saliva collection	Number of days and method of saliva collection	Mean sampling time(s) of saliva collection[c]
Ahnert et al. (2004)	Germany	70	Longitudinal; No	Center-based	11- 20 months (14.90, 1.70); Younger	Child care and home	6 (child care); 1 (home); cotton pads	At (expected) arrival, 30 min. later and **60 min. later** (both settings)
Albers et al. (2016)	The Netherlands	64	Longitudinal; No	Center-based	8.86 – 23.14 weeks (14.60, 2.80); Younger	Child care and home	9 (child care); 3 (home); cotton dental roll	10 AM and 4 PM (both settings)
Badanes et al. (2012)	United States	110	Cross-sectional; No	Center-based	2.03 – 5.38 years (4.03, 0.73); Older	Child care	3; cotton dental roll	9:50 AM and 3:34 PM
Bernard et al. (2015)	United States	168	Longitudinal; No	Center-based	1.20 months – 8 years (3.27 years, 2.09)	Child care and home	6 (child care); 2 (home); cotton dental roll	10:10 AM and 3:35 PM (child care); 10:53 AM and 3:35 PM (home)
De Haan et al. (1998)	United States	24	Longitudinal; No	Center-based	NR (30.49 months, 1.85); Younger	Child care and home	18/27 (child care); 2 (home); cotton dental roll; stimulants	Between 10 – 10:30 AM (both settings)
Dettling et al. (1999)	United States	70	Cross-sectional; No	Center-based	39–69 months (51.50, 7.09); Older	Child care and home	2 (per setting); cotton dental roll; stimulants	10:37 AM and 4:04 PM (child care); 10:24 AM and 4:11 PM (home)
Dettling et al. (2000)	United States	21	Cross-sectional; No	Home-based	40–69 months (52.75, 7.67); Older	Child care and home	2 (per setting); cotton dental roll; stimulants	10:30 AM and 3:30 PM (both settings)
Drugli et al. (2018)	Norway	112	Cross-sectional; Yes; Log-transformed data	Center-based and home-based	16–32 months (23.17, 3.80); Younger	Child care and home	2 (per setting); cotton dental roll	10 AM and 3 PM (both settings)
Eckstein-Madry et al. (2020)	Germany	60	Cross-sectional; No	Center-based	NR (47.40 months, 14); Older	Child care and home	2 (child care); 1 (home); NR	08:06 AM, **11:02 AM, 2:24 PM** and 6:12 PM (both settings)
Groeneveld et al. (2010)[a]	The Netherlands	45	Cross-sectional; Yes	Center-based	20–40 months (32, 4.40); Younger	Child care and home	1 (child care); 2 (home); sorbette	11:10 AM and 3:19 PM (child care); 10:59 AM and 3:10 PM (home)
Groeneveld et al. (2010)[b]	The Netherlands	71	Cross-sectional; Yes	Home-based	20–40 months (29.20, 6.30); Younger	Child care and home	1 (child care); 2 (home); sorbette	11:10 AM and 3:19 PM (child care); 10:59 AM and 3:10 PM (home)
Gunnar et al. (2010)	United States	151	Cross-sectional; Yes	Home-based	3 – 4.50 years (3.81, 0.23); Older	Child care and home	2 (per setting); cotton dental roll; stimulants	Between 10–11 AM and 3–4 PM (both settings)
Hatfield (2013)	United States	63	Cross-sectional; Yes	Center based	NR (53.92 months, 5.39); Older	Child care	2; sorbette	**8:29 AM, 10:48 AM** and **3:27 PM**
Lisonbee et al. (2008)	United States	191	Cross-sectional; No	Center-based	43–67 months (53, 4.10); Older	Child care	2; passive drool	9 AM and between 1 and 4 PM
Lumian et al. (2016)[a]	United States	118	Cross-sectional; No	Center-based	NR (4.51 years, 0.63); Older	Child care and home	3 (child care); 2 (home); cotton dental roll	10 AM and 3:25 PM (both settings)
Lumian et al. (2016)[b]	United States	78	Cross-sectional; No	Center-based	NR (3.92 years, 0.63); Older	Child care and home	3 (child care); 2 (home); cotton dental roll	9:50 AM and 4:06 PM (both settings)
Ouellet-Morin et al. (2010)	Canada	155	Longitudinal; No	Center-based and home-based	NR (2.28 years, 0.26); Younger	Child care and home	2 (per setting); sorbette	10:02 AM and 3:26 PM (child care); 8:51 AM and 5:20 PM (home)
Park and Choi (2009)	Korea	117	Cross-sectional; No	Center-based	4–5 years (NR, NR); Older	Child care	2; NR	10:30 AM and 3:30 PM
Sumner et al. (2010)	United States	42	Cross-sectional; No	Center-based	16–24 months (21, 2.48); Younger	Child care and home	2 (per setting); cotton dental roll; stimulants	10:03 AM and 3:27 PM (child care); 10:53 AM and 3:44 PM (home)
Tervahartiala et al. (2019)	Finland	106	Cross-sectional; No	Center-based	NR (2.26 years, 0.60); Younger	Child care and home	2 (1 child care day; 1 home day); swab	30 min after waking (at home), **10 AM** and **between 2–3 PM** (both settings) and before going to bed (at home)

(*continued on next page*)

Table 1 (*continued*)

Author(s) and year of publication	Country	Sample size	Design of study and covariates (yes/no)	Type of care	Child age range (M, SD) and group	Location(s) of saliva collection	Number of days and method of saliva collection	Mean sampling time(s) of saliva collection[c]
Tout et al. (1998)	United States	75	Cross-sectional; No	Center-based	2.67 – 5.83 (4.30, NR)	Child care	30; cotton dental roll; stimulants	10:30 AM and 3 PM
Vaillancourt et al. (2018)	Canada	198	Cross-sectional; No	Center-based	NR (33.62 months, 5); Older	Child care	2; sorbette	9:58 AM and 2:45 PM
Vermeer et al. (2010)[a]	Basque Country	60	Cross-sectional; Yes	Center-based	16–41 months (32.20, 4.90); Younger	Child care and home	1 (per setting); sorbette	10:40 AM and 3:26 PM (child care); 10:42 AM and 3:30 PM (home)
Vermeer et al. (2010)[b]	The Netherlands	25	Cross-sectional; Yes	Center-based	24–41 months (32.50, 4.60); Younger	Child care and home	1 (per setting); sorbette	10:40 AM and 3:26 PM (child care); 10:42 AM and 3:30 PM (home)
Watamura et al. (2003)[a]	United States	20	Cross-sectional; No	Center-based	2–16 months (10.40, NR); Younger	Child care and home	2 (per setting); cotton dental roll; stimulants	10 AM and 4 PM (child care); 10:02 AM and 4:27 PM (home)
Watamura et al. (2003)[b]	United States	35	Cross-sectional; No	Center-based	16–38 months (28.80, NR); Younger	Child care and home	2 (per setting); cotton dental roll; stimulants	10 AM and 4 PM (child care); 10:02 AM and 4:27 PM (home)
Watamura et al. (2009)	United States	65	Cross-sectional; No	Center-based	2.67–5.33 years (4,29, NR); Older	Child care and home	2 (per setting); cotton dental roll	10:48 AM and 3:48 PM (child care); 10:07 AM and 3:44 PM (home)
Watamura et al. (2002)	United States	40	Cross-sectional; No	Center-based	2.80–4.25 years (3.45, NR) for the younger preschoolers; 4.26–5.43 years (4.67, NR) for the older preschoolers;; Older	Child care and home	2 (per setting); cotton dental roll; stimulants	**10:37 AM**, 11:50 AM, 2:38 PM and **3:46 PM** (child care); 10:32 AM and 3:55 PM (home)

[a] = first independent subsample of study.
[b] = second independent subsample of study.
[c] = when more than one mid-morning and/or one mid-afternoon time slot is mentioned, the included time slot(s) is/are in bold.

cient, for the cortisol change score (mid-afternoon minus mid-morning, so that a positive score indicated an increase over the day and a negative value indicated a decrease over the day) at child care and one or more of the four correlates.

In the final stage of the selection process, studies with overlapping samples were excluded. When samples of studies (partly) overlapped, published articles in peer-reviewed journals were preferred to dissertations. When both of the studies with (partly) overlapping samples were part of published articles in peer-reviewed journals, the article with the most relevant data regarding the research questions was chosen. When the studies were equally relevant, we included the study with the biggest sample size. Finally, we excluded 12 studies because the methodology did not match our research questions. For example, studies that collected saliva samples directly or shortly after awakening or daytime naps were excluded, because these samples were more likely to reflect part of the cortisol awakening response than mid-morning or mid-afternoon cortisol values. Regarding constructs, two studies were excluded since these studies used instruments that were not comparable to other instruments measuring child care quality (Reunamo et al., 2012; Sajaniemi et al., 2011) and could not be included for the first research question either. The authors of 22 articles were asked for additional non-published raw (uncorrected) data. For 12 papers, authors responded to this request. In two cases, the requested data were no longer available. Six articles for which we did not receive responses were (partially) included on the basis of the information that was available. Four papers had to be excluded because of the lack of usable data and no response. All this resulted in the final inclusion of 24 articles, which contained 28 independent subsamples (because when samples within one article were completely independent, they were treated as separate studies, which was the case for four articles).

A flow chart of the article selection process for the meta-analyses can be found in Fig. 1. After the removal of duplicates, 3094 articles were identified. The titles of these articles were screened and 123 articles met the inclusion criteria. Then, the abstracts of the articles were screened for eligibility, followed by a discussion of (27) unclear abstracts. Of the remaining 96 articles, twenty abstracts were independently double-screened to check whether the other authors agreed with the decisions. This was the case for all abstracts. While checking the reference lists of the 56 articles that met the criteria based on the abstract and after asking some authors for additional data, we came across another eight relevant studies. The full-texts of the first randomly selected 10 articles were coded together by four researchers in order to finalize the coding form (see below) and practice coding. After that, two researchers independently coded a reliability set containing 10 randomly selected articles. The percentage of absolute agreement between coders for coding the descriptive study and sample characteristics was calculated and found to be adequate (M = 90.68%, *Min.* = 60%, *Max.* = 100%). After this, the remaining articles were coded separately. Difficulties in coding the articles were discussed during weekly meetings and if needed, a third party was consulted.

2.3. Data extraction

For the descriptive statistics, we used a self-developed standardized coding form. Information on study characteristics (e.g., year of publication, research design, country), sample characteristics (e.g., sample size, ethnicities, child age) and child care characteristics (e.g., type of care, number of groups and centers) was extracted. Data on the collection of saliva and the correlates of cortisol levels (e.g., name and type of the instrument(s), informant(s), reliability) were retrieved as

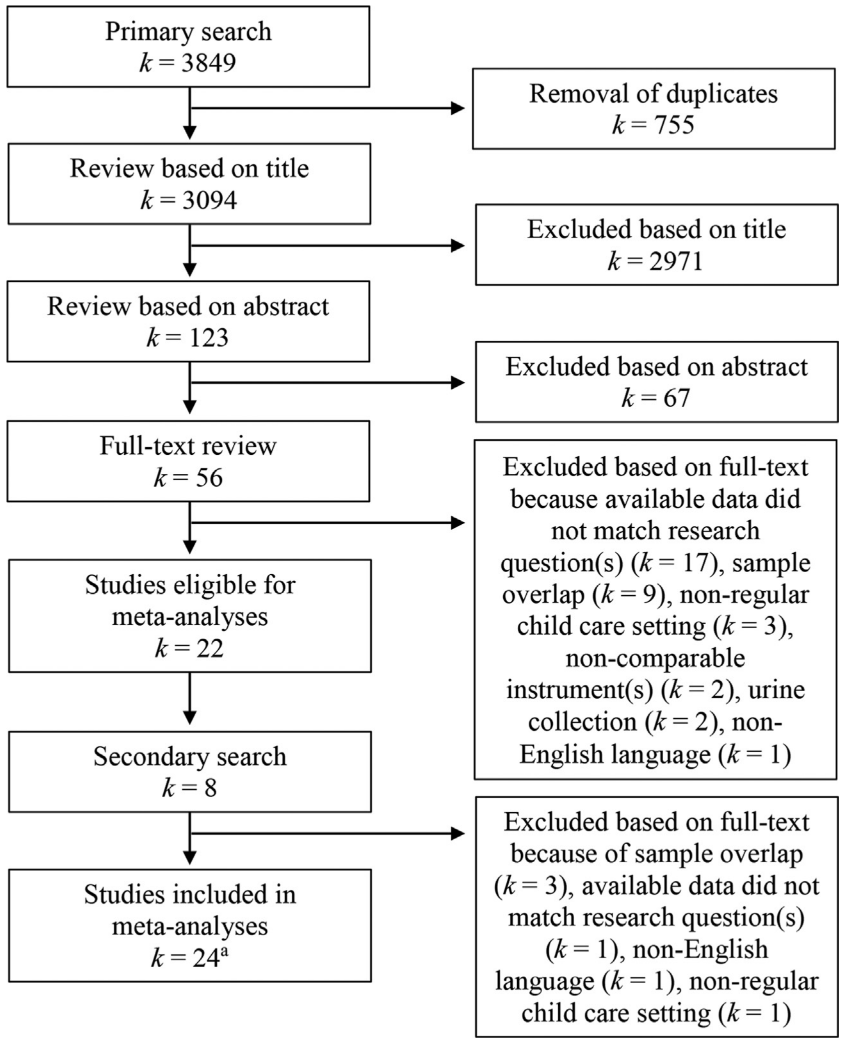

Fig. 1. Flow chart of the article selection process for the meta-analyses.

well. In total, seven meta-analyses were performed, three to answer the research question concerning cortisol levels at child care versus home and four to answer the research question concerning correlates of cortisol levels at child care. For three articles (Badanes et al., 2012; Groeneveld et al., 2010; Sumner et al., 2010), data were (partly) retrieved from the corresponding dissertation or another article about the same sample with more information on the cortisol statistics or coefficients (Badanes, 2010; Groeneveld et al., 2012; Sumner, 2009). Data for the analyses (i.e., cortisol statistics, information on times and locations of saliva collection, sample sizes, and coefficients) were thoroughly checked again for each individual study as a reliability check (and thus were not based on the coding forms alone).

When possible, uncorrected values were used. However, in some cases only correlational analyses with covariates and log-transformed cortisol values were available. In these cases, unstandardized regression coefficients were divided by the standard error to obtain t-values. See Table 1 for more information on the methodological differences between studies. When sample sizes differed between mid-morning and mid-afternoon measures within one setting (child care or home), we chose to include the smallest sample size for conservative purposes. When more time-points were reported (multiple weeks in a row) we chose to take the average of all time-points for which values for both child care and home were available, except when time-points were too far apart (such as 1 year, which was the case for Ouellet-Morin et al., 2010). In some cases, saliva was collected 2 days at child care and only 1 day at home. In these instances, we calculated a mean score for the cortisol levels at child care. For temperament, the subscales measuring negative affectivity, fearfulness, shyness, and anger and aggression were included, because these subscales were most frequently reported and are theoretically closest to the definition of a more difficult temperament. When no coefficient for a composite score of child care quality was provided, we included subscales related to both global quality (e.g., classroom quality) and process quality (e.g., caregiver-child interactions on group-, dyad- and the child-level). See Table 2 for the included measurement instruments.

2.4. Meta-analytic procedures

2.4.1. Within-child cortisol levels at child care versus home

As mentioned, a series of seven meta-analyses was conducted, performed with Comprehensive Meta-Analysis Software Version 2 (Borenstein et al., 2005). Raw (uncorrected) means, standard deviations, and sample size were used to calculate Hedge's g in order to answer the first research question on differences between mid-morning and mid-afternoon cortisol at child care compared to home. A positive value for the combined Hedge's g would indicate that the mean cortisol level was higher at child care than at home, in the morning or the afternoon. Concerning the meta-analysis on cortisol change from mid-morning to mid-afternoon at child care compared to home, the change score and the standard deviation of this change score were entered for both settings. For this analysis, a positive effect size would signify a larger cortisol increase over the day at child care compared to home (since the

Table 2
Overview of the studies that related cortisol to child care quality, child temperament, group size and/or hours at child care.

Author(s) and year of publication	Correlate(s)	Name(s) of instrument(s)	Type(s) of instrument(s)	Informant(s)	Reliability of instrument(s)	Mean number of group size and hours in care
Albers et al. (2016)	Child care quality; Child temperament	Ainsworth; IBQ-R	Dyad-level observation; Questionnaire	Researcher; Parent	Good; Fair	NA
Badanes et al. (2012)	Child care quality; Child temperament	ECERS-R; CBQ	Group-level observation; Questionnaire	Researcher; Professional caregiver	NR; Fair	NA
Dettling et al. (1999)	Child temperament	CBQ	Questionnaire	Parent and professional caregiver	NR	NA
Dettling et al. (2000)	Child care quality; Child temperament; Group size	ORCE; CBQ	Dyad-level observation; Questionnaire	Researcher; Parent and professional caregiver	Fair; NR	Group size: 3–8 versus 9–13 (Median = 8.50)
Drugli et al. (2018)	Child care quality; Child temperament; Group size; Hours at child care	CLASS-Toddler; CRITQ	Group-level observation; Questionnaire	Researcher; Parent	Good; Poor	Group size: 14.40 (M), 5.89 (SD), 4–31 (range) Hours in care: 5–7 (33%) versus 8–9 (67%) per day
Groeneveld et al. (2010)[a]	Child care quality	ECERS-R & NCKO scale for caregiver sensitivity	Group-level observation and interview; Group-level observation	Researcher; Researcher	Good; Fair	NA
Groeneveld et al. (2010)[b]	Child care quality; Child temperament	IT-CC-HOME & NCKO scale for caregiver sensitivity; TBAQ	Group-level observation and interview; Group-level observation; Questionnaire	Researcher; Parent	Good; Fair; Fair	NA
Gunnar et al. (2010)	Child care quality; Child temperament; Group size; Hours at child care	M-ORCE; M-ORCE	Dyad- and group-level observation; Child-level observation	Researcher; Professional caregiver	Good; Good	Group size and hours in care: NR
Hatfield (2013)	Child care quality	CLASS Pre-K	Group-level observation	Researcher	Good	NA
Lisonbee et al. (2008)	Child care quality; Child temperament; Group size	CIS; CBQ	Group-level observation; Questionnaire	Researcher; Parent	Good; Fair	Group size: 17 (Median), 11–24 (range)
Lumian et al. (2016)[b]	Hours at child care	NA	NA	NA	NA	Hours in care: 5 (47%) versus 2–3 (53%) days per week
Ouellet-Morin et al. (2010)	Child care quality	ITERS & ECERS	Group-level observation	Researcher	NR	NA
Park and Choi (2009)	Child care quality; Child temperament; Hours at child care	Observation Scale for Day Care Programs; Social Competence and Behavior Evaluation	Group-level observation; Questionnaire	Researcher; Professional caregiver	NR; Good	Hours in care: NR
Sumner (2009)	Child care quality	ITERS-R	Group-level observation	Researcher	Good	NA
Tervahartiala et al. (2019)	Group size; Hours at child care	NA	NA	NA	NA	Group size: 13.47 (M), 3.80 (SD) Hours in care: 20 (71.7%) versus ≤ 16 (28.3%) days per month
Tout et al. (1998)	Child temperament	SCBE	Questionnaire	Professional caregiver	Good	NA
Vaillancourt et al. (2018)	Child care quality; Hours at child care	ITERS-R & ECERS-R	Group-level observation	Researcher	NR	Hours in care: 29.35 (M), 12.55 (SD) per week
Vermeer et al. (2010)[a]	Child care quality; Child temperament; Group size; Hours at child care	ECERS-R; ICQ	Group-level observation; Questionnaire	Researcher; Parent	Good; Good	Group size: 15.40 (M), 3 (SD), 8–20 (range) Hours in care: 27.50 (M), 8.20 (SD), 10–43 (range)
Vermeer et al. (2010)[b]	Child care quality; Child temperament; Group size; Hours at child care	ECERS-R; TBAQ	Group-level observation; Questionnaire	Researcher; Parent	Good; Fair	Group size: 12.10 (M), 2.30 (SD), 5–15 (range) Hours in care: 20.30 (M), 6.80 (SD), 2–40 (range)
Watamura et al. (2003)[a]	Child care quality; Child temperament	ECERS; IBQ	Group-level observation; Questionnaire	NR; Professional caregiver	NR; Good	NA
Watamura et al. (2003)[b]	Child care quality; Child temperament	ECERS; TBAQ	Group-level observation; Questionnaire	NR; Professional caregiver	NR; Fair	NA
Watamura et al. (2009)	Child care quality	M-ORCE	Child- and group-level observation	Researcher	Fair	NA

Note. NR = not reported; NA = not applicable.
[a] = first independent subsample of study.
[b] = second independent subsample of study.c= ≤ .60 = poor, > .60 < .80 = fair, ≥ .80 = good.

home values were subtracted from the child care values). Hedge's *g* was computed by dividing the difference between the two means (or change scores) by the pooled standard deviation (Borenstein et al. 2009).

Sensitivity analysis. The three meta-analyses on cortisol levels at child care versus home consisted of two paired measurements, namely multiple saliva collections for the same child in two different settings (at child care and home). To control for this dependency, one needs the specific correlation between cortisol measures at home and child care for every individual study. Since these were not available, we based the correlations on two studies of one of the authors (X). We used the weighted mean of the correlations in the Dutch samples of these studies as an educated guess for the other correlations in the analyses, which were $r = .261$ for the mid-morning cortisol levels, $r = .303$ for the mid-afternoon cortisol levels, and $r = .359$ for the cortisol change score. For the meta-analysis on cortisol change, the estimated correlation between the mid-morning and mid-afternoon values over the settings was $r = .292$ and was therefore used in the formula to calculate the standard deviations of the change scores (the square root of the outcome of the following equation: $S^2_{\bar{x}_d} = S^2_{AM} + S^2_{PM} - 2\ r\ S_{AM} S_{PM}$. In a sensitivity analysis, correlations were also set at the values of $r = .0$ (complete independence) and $r = .8$ (strong dependence) for all analyses (in addition to the estimated correlation(s)), which resulted in nine different effect sizes for the meta-analysis on cortisol change and six different effect sizes in total for the meta-analyses on mid-morning and mid-afternoon cortisol.

2.4.2. Correlates of cortisol levels at child care

For the second research question on correlates of cortisol levels at child care, correlations (or comparable parameters if uncorrected correlations were not available) between the correlates, the cortisol increase from mid-morning to mid-afternoon, and accompanying sample sizes were used to calculate the combined effect size (*r*). A positive correlation would indicate that a more difficult child temperament, larger group size and more hours at child care correlate with a larger increase in cortisol over the day at child care. For child care quality, a negative correlation would indicate lower child care quality to correlate with a larger cortisol increase at child care. Random-effects models were chosen in all cases, because we assumed that the variance between studies did not only contain measurement error, but also real variation due to differences in study design, measurement instruments, and sample characteristics (Borenstein et al., 2009). *Q*-statistics were used to examine if the sets of studies could be considered heterogeneous (Borenstein et al., 2009). If this was the case. To get an idea of the robustness of the outcomes, we calculated the fail-safe number *N*, which can be described as the number of studies of an average effect size with on average a null effect, that is necessary to reduce the combined effect size to non-significance (Rosenthal, 1979). Potential publication bias was visualized and when present, compensated by using Duval & Tweedie's trim and fill method (Duval & Tweedie, 2000a, 2000b). We transformed individual effect sizes into standardized *z*-scores and defined them as outliers if their value was smaller than -3.29 or greater than 3.29 (Tabachnick & Fidell, 2007). If outliers were present, we performed analyses both with and without the outlier(s).

2.4.3. Moderators of study outcomes

To answer the third research question, we carried out additional moderator analyses. To perform such analyses reliably, the smallest subset for each categorical moderator was set at a minimum of three studies. For all meta-analyses that showed significant heterogeneity between studies, we tested whether effect sizes differed based on year of publication, age groups (younger children versus older children), research design (cross-sectional versus longitudinal research design), country (United States versus other countries than the United States), type of care (center-based versus home-based child care settings), and method of saliva collection (cotton dental roll versus other devices). Although more specific classifications concerning age groups would have been preferred, this was not possible, since many studies included multiple age groups and did not differentiate between those age groups in their results. As an appropriate solution, we made two groups: (1) the studies with only infants, only toddlers, or with a combination of both infants and toddlers, and (2) studies with only preschoolers, or with a combination of both toddlers and preschoolers. By doing this, we created two non-overlapping groups approximately equal in size that differed in the mean child age. In addition, for the second research question, we tested whether outcomes differed based on whether covariates were included in the correlations. Child age was included as a moderator and not as a correlate, since the correlation between cortisol and child age at the individual level was only reported in some studies, while child age at study level could be tested as a moderator for all studies. Year of publication was analyzed using meta-regression, while the other variables were treated as categorical moderators. For the meta-regression analyses, the mixed effects regression model (method of moments) was chosen, and for the categorical moderator analyses it was assumed that a common among-study variance component existed across subgroups. A variable was considered a moderator if the effect sizes of categories differed significantly.

3. Results

3.1. Descriptive statistics

The characteristics of the studies included in the meta-analyses ($k = 28$) are summarized in Table 1. In Table 2, one can find the studies that included measures of child care quality, child temperament, group size, and hours at child care, which were considered correlates of cortisol secretion and were meta-analytically examined ($k = 22$). Below, the study and sample characteristics, the procedures around the saliva collection and the measurement of the correlates are described.

3.1.1. Study characteristics

Most of the studies (48.8 %) were conducted in North-America (United States: $k = 14$; Canada: $k = 2$), followed by Europe (The Netherlands: $k = 4$; Germany: $k = 2$; Scandinavia: $k = 2$; Spain: $k = 1$) and Asia ($k = 1$). The majority of studies (57.7 %) was published in or after 2010. A minority of studies ($k = 3$) focused completely on home-based child care, whereas 21 studies included only center-based child care settings. Two studies covered both types of care in their sample. Most studies ($k = 21$) were set up with a cross-sectional design, whereas five studies could be considered longitudinal, since these studies included multiple time-points across several weeks, months, or even years, with a mean of 6 time-points. For four studies, children were followed specifically during their transition into a (new) child care setting and these were all longitudinal studies.

3.1.2. Sample characteristics

The sample size of the studies contained on average 83 children (*SD* = 52.87, *range* = 20–198). Most studies ($k = 10$) included only toddlers (children between 1 and 3 years of age) or focused on both toddlers and preschoolers (the latter being children of 3 years and older; $k = 9$). Other studies included only preschoolers ($k = 3$) or infants (children under the age of 12 months; $k = 2$) or another combination of two or three age groups ($k = 2$). The mean age of the children ($k = 22$) was 2.79 years (*SD* = 1.19, *range* = 0.23–4.49). When it comes to gender distribution ($k = 25$), boys, on average, made up 51.16% of the samples (*SD* = 6.05, *range* = 38.10–64.29). The majority of studies ($k = 21$) did not report on the general health of the children. For the ones that did ($k = 5$), general health was considered to be good, with no children with medical conditions and/or disabilities. In most studies ($k = 20$), the socio-economic status (SES) of the participants was not reported. However, five studies reported that the sample mainly consisted of children belonging to families of middle-class to high SES and one study included children belonging to families of middle-class to low SES. Regarding ethnicity, on average, most participants were Caucasian (72.42%, $k = 17$), followed by African-American (12.98%, $k = 14$), mixed or Native American (10.05%, $k = 8$), Asian-American (8.60%, $k = 14$), and Hispanic ethnicity (6.93%, $k = 12$).

3.1.3. Child care characteristics

The average number of participating child care centers was 22 ($SD = 28.59$, $range = 1–120$, $k = 25$) and the average number of child care groups involved 28 ($SD = 32.92$, $range = 2–120$, $k = 21$). Information regarding child-caregiver ratio, group size, and hours and experience in care were only reported by a small subset of studies. The mean child-caregiver ratio ($k = 6$) was 7:1 ($SD = 4.45$, $range = 3:1–15:1$) and the mean group size ($k = 8$) was 11 ($SD = 4.43$, $range = 2.90–15.40$). On average, children spent 29.61 hours at child care per week ($SD = 9.12$, $range = 19.40–40.61$, $k = 8$) and were on average 0.96 years old when they entered child care ($SD = 0.50$, $range = 0.28–1.42$, $k = 4$). The average number of months children had spent at child care in total at the time of data collection was 14.11 months ($SD = 10.84$, $range = 0–27.52$, $k = 6$).

3.1.4. Saliva collection

For 21 studies, saliva was collected both at the child care center and at home. Five studies only measured cortisol at child care. The median for the number of saliva collection days at child care was 2 days ($SD = 6.80$, $range = 1–30$, $k = 26$) and for the home setting 1.85 days ($SD = .75$, $range = 1–4$, $k = 20$). The most common instrument to collect saliva was the cotton dental roll, which was used in 14 studies. Chemical stimulants were administered in nine studies, to stimulate saliva flow before saliva collection. While most studies did not report on the time it took to collect the saliva, the studies that did ($k = 11$) mostly collected saliva during 1–3 min. In less than half of the studies, authors were not precise about the procedural aspects they took into account when collecting saliva. Of the studies that reported about these aspects, 10 studies made sure children did not eat 30 min before collection, eight did the same for drinking, eight for sleeping and two for physical activity. Individual saliva samples were excluded mostly on the basis of child medicine use and insufficient saliva. In many cases ($k = 21$), the values for saliva were transformed because of a skewed distribution. In total, 14 studies reported explicitly that outliers were removed or winsorized before the analyses.

3.1.5. Correlates

Regarding the included correlates for cortisol change over the day, 14 studies measured child temperament, 17 studies were available for child care quality, 7studies for group size and 8 studies for hours at child care. Except for one study, child temperament was measured with a questionnaire filled in by the parent and/or professional caregiver. Child care quality was mostly observed at the group level by one or more of the researcher(s). The majority of constructs showed high internal reliability. Most studies reported that the participating child care centers were of moderate to high quality, although there were some exceptions (Hatfield et al., 2013; Park & Choi, 2009; Vermeer et al. 2010).

3.2. Outcomes of the meta-analyses

In Table 3, an overview of the results of the different meta-analyses is displayed. In order to explain the heterogeneity found for some groups of studies, several moderator analyses were performed. Since the sensitivity analysis yielded no significantly different effect sizes, the moderator analyses were only carried out once per analysis with the estimated correlations. The summary of the categorical moderator analyses can be found in Table 4. Since one study (Bernard et al., 2015) included children who attended primary school (34 of the 168 children varying from 5 to 8 years), we performed all analyses below (if applicable) both with and without the results of this study, but this did not yield different outcomes.

Table 3
Overview of the results of the meta-analyses.

	K	N	Effect size[a]	95% CI	Q
Mid-morning	22	1243	−.05	[−0.23, 0.12]	121.09**
Mid-afternoon	20	1093	.38**	[0.23, 0.53]	69.65**
Cortisol change	20	1093	.35**	[0.19, 0.51]	90.33**
Child care quality	17	1420	−.05	[−0.15, 0.04]	43.32**
Child temperament	13	1059	.01	[−0.05, 0.06]	11.66
Group size	7	636	.06	[−0.05, 0.17]	9.56
Hours at child care	8	784	.15*	[0.04, 0.26]	15.91*

* $p < .05$.
** $p < .001$.
[a] = for the first three rows, the effect size is the combined Hedge's g; for the last four rows, the effect size is the combined corelation.

3.2.1. Cortisol levels at child care versus home and moderators

Mid-morning cortisol. The first meta-analysis on the difference between the mean mid-morning cortisol value at child care versus home ($k = 22$, $N = 1243$) yielded a non-significant combined Hedge's g of -.05, $p = .563$ (see Fig. S2 for the forest plot). The first subsample of the study by Lumian et al. (2016) was an outlier, but when this study was left out, comparable results were obtained (Hedge's $g = .02$, $p = .607$). The combined effect size was still non-significant when a correlation of .0 or .8 for cortisol values between child care and home was used (Hedge's $g = −.05$, $p = .574$ and Hedge's $g = −.04$, $p = .604$, respectively). In sum, the results did not point towards a difference between mid-morning cortisol values at child care versus home. The Q-statistic was significant, indicating that the studies could be considered heterogeneous ($p < .001$). However, none of the potential moderators was significant.

Mid-afternoon cortisol. The meta-analysis on the difference between the mean mid-afternoon cortisol values at child care versus home ($k = 20$, $N = 1093$) yielded a small to medium combined Hedge's g of .38 ($p < .001$), indicating that the mid-afternoon cortisol values were higher at child care than at home (within-child; see Fig. 2 for the forest plot). When the analysis was performed with a child care-home correlation of .0 or .8, combined effect sizes were comparable (Hedge's $g = .38$, $p < .001$ and Hedge's $g = .35$, $p < .001$, respectively). Duval and Tweedie's trim and fill method did not indicate publication bias. Furthermore, Rosenthal's fail-safe number demonstrated that 446 studies with an average sample size and average effect size of 0 would be needed to bring the p-value to non-significance, showing the result to be quite robust. The Q-statistic showed that the studies were heterogeneous ($p < .001$). Country was a significant moderator: the studies carried out in the United States (Hedge's $g = .54$, $p < .001$, $k = 11$, $N = 577$) as well as the studies conducted in countries other than the United States (Hedge's $g = .20$, $p = .042$, $k = 9$, $N = 516$) showed that cortisol values in the afternoon were higher at child care compared to home. However, the combined effect size for the studies from the United States was significantly higher ($p = .015$). The method of saliva collection moderated the outcome as well: the difference in afternoon cortisol at child care versus home was only significant for studies that used cotton dental rolls to measure saliva (Hedge's $g = .52$, $p < .001$, $k = 13$, $N = 724$) and not for studies that used other devices (Hedge's $g = 0.13$, $p = .350$, $k = 7$, $N = 362$), and this difference was significant ($p < .001$). There were no other significant moderators.

Cortisol change from mid-morning to mid-afternoon. The combined Hedge's g for the meta-analysis on the difference between the cortisol change from mid-morning to mid-afternoon at child care versus home ($k = 20$, $N = 1093$) was .35 (small to medium), $p < .001$. The forest plot is displayed in Fig.3, where a positive effect size indicates that the cortisol change from mid-morning to mid-afternoon was higher at child care compared to home (within-child). The first subsample of Lumian et al. (2016) again had a high standardized z-score of 4.55. Removing this study from the analysis changed the outcome slightly, but the significance of the effect did not change (Hedge's $g = .27$, $p < .001$).

Table 4
Overview of the results of the categorical moderator analyses.

	Moderator	Q	Category	k	N	Effect size[a]
Mid-morning	Age group(s)[b]	2.10	Younger children[c]	13	652	.05
			Older children[c]	8	462	−.23
	Country	2.12	United-States	12	660	.08
			Other than US	10	583	−.14
	Research design	2.06	Cross-sectional	18	964	−.09
			Longitudinal	4	279	.19
	Type of care	0.00	Center-based care	19	1092	−.04
			Home-based care	3	151	−.03
	Method of saliva collection	0.06	Cotton dental roll	15	873	−.07
			Other	7	362	−.02
Mid-afternoon	Age group(s)[b]	2.97	Younger children[c]	11	565	.27**
			Older children[c]	8	399	.55**
	Country	5.96*	United-States	11	577	.54**
			Other than US	9	516	.20*
	Type of care	0.95	Center-based care	17	942	.41**
			Home-based care	3	151	.20
	Method of saliva collection	7.77*	Cotton dental roll	13	724	0.52**
			Other	7	362	0.13
Cortisol change	Age group(s)[b]	4.76*	Younger children[c]	11	565	.20
			Older children[c]	8	399	.58**
	Country	7.97*	United-States	11	577	.55**
			Other than US	9	516	.13
	Type of care	0.62	Center-based care	11	942	.38**
			Home-based care	3	151	.19
	Moderator	***Q***	***Category***	***k***	***N***	***Effect size[a]***
	Method of saliva collection	5.30*	Cotton dental roll	13	724	0.48**
			Other	7	362	0.11
Child care quality	Age group(s)	1.64	Younger children[c]	9	478	.01
			Older children[c]	8	942	−.11
	Country	0.57	United-States	8	751	−.09
			Other than US	9	669	−.02
	Type of care	4.40*	Center-based care	14	1204	−.01
			Home-based care	3	216	−.27*
	Covariates	3.59	No	10	1038	.02
			Yes	7	838	−.16*
	Method of saliva collection	0.36	Cotton dental roll	8	666	−.02
			Other	9	754	−.08
Hours at child care	Age group(s)	0.17	Younger children[c]	4	274	.12
			Older children[c]	4	510	.17*
	Country	0.52	United States	3	391	.19*
			Other than US	5	393	.10
	Method of saliva collection	0.43	Cotton dental roll	3	336	0.10
			Other	5	448	0.18*

* $p < .05$.
** $p < .001$.
[a] = for the first three meta-analyses, the effect size is the combined Hedge's *g*; for the last two meta-analyses, the effect size is the combined corelation.
[b] = one study (Bernard et al., 2015) was excluded in this analysis, because the study includes all three age groups.
[c] = younger children: studies that included only infants, only toddlers or a combination of both infants and toddlers; older children: studies that included only preschoolers or a combination of both toddlers and preschoolers.

Duval and Tweedie's trim and fill method was implemented and did not reveal publication bias. Moreover, Rosenthal's fail-safe number showed the outcome was robust: 378 average null studies would be required to bring the total effect size to non-significance. The combined Hedge's *g*'s for the nine different correlations as described in the Method varied from 0.30 to 0.59 (see Table 2 in the Supplementary Materials for the complete overview). All effect sizes were significant, but the absolute value for Hedge's *g* varied somewhat. Fig. 5 shows the mean cortisol pattern from mid-morning to mid-afternoon in both the child care and the home setting, for illustrational purposes. The *Q*-statistic was significant as well ($p < .001$). Child age, country, and method of saliva collection were significant moderators. For child age, the studies that included younger children (infants, toddlers or a combination of both infants and toddlers) did not find a higher cortisol increase over the day at child care (Hedge's $g = .20$, $p = .067$, $k = 11$, $N = 565$), while the studies that included older children (preschoolers or a combination of both toddlers and preschoolers) did (Hedge's $g = .58$, $p < .001$, $k = 8$, $N = 399$). This difference was significant ($p = .029$). Furthermore, studies that were carried out in the United States showed significantly higher cortisol increases over the day at child care compared to home (Hedge's $g = .55$, $p < .001$, $k = 11$, $N = 577$), while studies that were carried out outside of the United States did not find different cortisol patterns over the day (Hedge's $g = .13$, $p = .236$, $k = 9$, $N = 516$). The difference in effect size between studies from the United States and other countries was significant ($p < .005$). The method of saliva collection moderated the outcome as well: the difference in cortisol change over the day at child care versus home was only significant for studies that used cotton dental rolls to measure saliva (Hedge's $g = .48$, $p < .001$, $k = 13$, $N = 724$) and not for studies that used other devices (Hedge's $g = 0.11$, $p = .387$, $k = 7$, $N = 362$), and the difference was significant ($p < .001$). Year of publication and type of care did not explain the variance between studies.

Study name	Time point	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Albers et al. (2016)	Combined	0.261	0.149	0.022	-0.031	0.553	1.752	0.080
Bernard et al. (2015)	Combined	0.338	0.110	0.012	0.123	0.554	3.081	0.002
Dettling et al. (1999)	Blank	0.317	0.167	0.028	-0.010	0.645	1.901	0.057
Dettling et al. (2000)	Blank	0.076	0.260	0.068	-0.433	0.585	0.293	0.769
Drugli et al. (2018)	Blank	0.556	0.138	0.019	0.285	0.827	4.025	0.000
Eckstein-Madry et al. (2020)	Blank	0.358	0.228	0.052	-0.089	0.804	1.569	0.117
Groeneveld et al. (2010a)	Blank	0.158	0.255	0.065	-0.342	0.658	0.619	0.536
Groeneveld et al. (2010b)	Blank	0.124	0.172	0.030	-0.213	0.461	0.721	0.471
Gunnar et al. (2010)	Blank	0.299	0.113	0.013	0.077	0.521	2.637	0.008
Lumian et al. (2016a)	Blank	1.270	0.212	0.045	0.854	1.686	5.981	0.000
Lumian et al. (2016b)	Blank	0.940	0.159	0.025	0.628	1.253	5.897	0.000
Ouellet-Morin et al. (2010)	Combined	0.351	0.112	0.013	0.131	0.571	3.131	0.002
Sumner et al. (2010)	Blank	0.717	0.201	0.041	0.322	1.111	3.562	0.000
Tervahartiala et al. (2019)	Blank	-0.062	0.123	0.015	-0.303	0.179	-0.503	0.615
Vermeer et al. (2010a)	Blank	-0.295	0.185	0.034	-0.658	0.067	-1.596	0.111
Vermeer et al. (2010b)	Blank	0.381	0.237	0.056	-0.084	0.846	1.605	0.109
Watamura et al. (2002)	Blank	0.306	0.382	0.146	-0.442	1.055	0.801	0.423
Watamura et al. (2003a)	Blank	0.251	0.312	0.097	-0.361	0.862	0.804	0.421
Watamura et al. (2003b)	Blank	0.682	0.277	0.077	0.139	1.226	2.461	0.014
Watamura et al. (2009)	Blank	0.650	0.183	0.034	0.291	1.009	3.546	0.000
		0.381	0.075	0.006	0.233	0.529	5.049	0.000

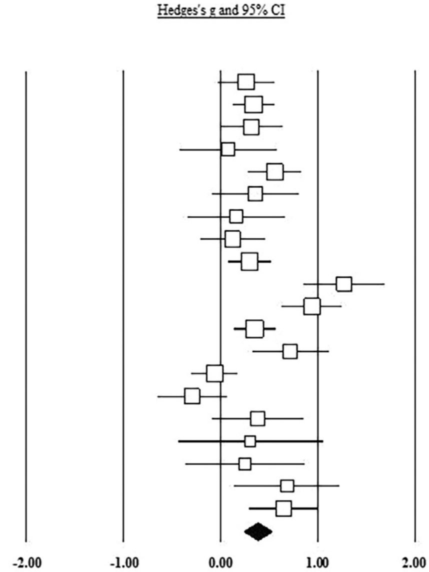

Fig. 2. Forest plot for mid-afternoon cortisol at child care versus home.

Study name	Time point	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Albers et al. (2016)	Combined	-0.029	0.140	0.020	-0.304	0.245	-0.210	0.834
Bernard et al. (2015)	Combined	0.292	0.103	0.011	0.090	0.495	2.830	0.005
Dettling et al. (1999)	Blank	0.400	0.162	0.026	0.081	0.718	2.460	0.014
Dettling et al. (2000)	Blank	0.186	0.251	0.063	-0.306	0.678	0.740	0.459
Drugli et al. (2018)	Blank	0.377	0.128	0.016	0.127	0.627	2.957	0.003
Eckstein-Madry et al. (2020)	Blank	0.215	0.214	0.046	-0.205	0.634	1.003	0.316
Groeneveld et al. (2010a)	Blank	0.145	0.244	0.060	-0.334	0.625	0.595	0.552
Groeneveld et al. (2010b)	Blank	0.209	0.166	0.028	-0.116	0.535	1.260	0.208
Gunnar et al. (2010)	Blank	0.224	0.108	0.012	0.013	0.435	2.083	0.037
Lumian et al. (2016a)	Blank	2.233	0.284	0.081	1.676	2.790	7.857	0.000
Lumian et al. (2016b)	Blank	0.812	0.147	0.022	0.524	1.099	5.531	0.000
Ouellet-Morin et al. (2010)	Combined	0.145	0.104	0.011	-0.058	0.349	1.397	0.162
Sumner et al. (2010)	Blank	0.926	0.206	0.042	0.522	1.330	4.491	0.000
Tervahartiala et al. (2019)	Blank	0.093	0.118	0.014	-0.138	0.325	0.791	0.429
Vermeer et al. (2010a)	Blank	-0.150	0.175	0.030	-0.492	0.192	-0.860	0.390
Vermeer et al. (2010b)	Blank	0.138	0.220	0.049	-0.294	0.570	0.626	0.531
Watamura et al. (2002)	Blank	0.474	0.380	0.145	-0.271	1.220	1.247	0.212
Watamura et al. (2003a)	Blank	-0.029	0.294	0.086	-0.605	0.548	-0.097	0.923
Watamura et al. (2003b)	Blank	0.436	0.250	0.062	-0.054	0.925	1.745	0.081
Watamura et al. (2009)	Blank	0.438	0.167	0.028	0.111	0.765	2.623	0.009
		0.350	0.082	0.007	0.190	0.510	4.279	0.000

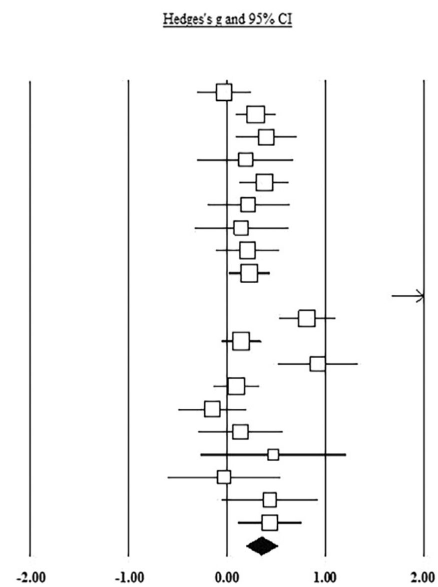

Fig. 3. Forest plot for cortisol change from mid-morning to mid-afternoon at child care versus home.

3.2.2. Correlates of cortisol levels at child care and moderators

Child care quality. The 17 studies on the correlation between child care quality and the increase in cortisol from mid-morning to mid-afternoon at child care ($N = 1420$) yielded a non-significant combined effect size ($r = -.05$, $p = .254$; see Fig. S2 for the forest plot). Separate analyses on the subscales process quality (e.g., caregiver-child interactions, $k = 8$) and global quality ($k = 13$) yielded non-significant results as well ($r = -.10$, $p = .119$ and $r = -.06$, $p = .252$, respectively). Overall, child care quality did not seem to be related to the cortisol change over the day at child care. The *Q*-value showed studies were heterogeneous ($p < .001$). Type of care moderated the outcome, with the correlation being significant for the studies that included only home-based child care ($r = -.27$, $p = .017$, $k = 3$, $N = 216$) and not significant for the studies that included (mainly) center-based child care centers ($r = -.01$, $p = .813$, $k = 14$, $N = 1204$), which differed significantly ($p = .036$). This indicates that the correlation between lower child care quality and cortisol increase over the child care day was present only for studies that included home-based child care settings. Country, year of publication, child age, the inclusion of covariates, and the method of saliva collection did not moderate the outcome.

Child temperament. The combined effect size for the correlation between child temperament and the cortisol change at child care ($k = 13$, $N = 1059$) was not significant ($r = .01$, $p = .657$; see Figure S3 for the forest plot). Analyses on different subscales yielded non-significant results as well: negative affect ($k = 7$, $r = .05$, $p = .418$), angry-aggressive temperament ($k = 6$, $r = .04$, $p = .547$) and anxious-withdrawn temperament ($k = 6$, $r = .05$, $p = .338$). The results indicated an absence of an association between child temperament and cortisol increase over the day at child care. The Q-value showed studies were homogeneous ($p = .473$).

Group size. The seven studies on group size in association with cortisol change ($N = 636$) showed a combined non-significant effect size of $r = .06$ ($p = .271$), see Figure S4 for the forest plot. The results indicated that group size and cortisol increase over the day at child care did not

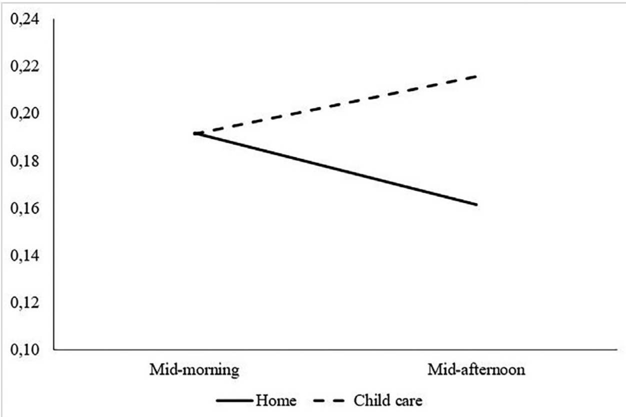

Fig. 4. Forest plot for the correlation between cortisol increase and hours at child care.

correlate. The Q-value was not significant either, implying that studies were homogeneous (p = .145).

Hours at child care. The combined effect size for the eight studies on the correlation between hours at child care and cortisol change ($N = 784$) yielded a significant but small effect size of $r = .15$ ($p = 0.010$). The more hours the child spent at child care (per day, week or month), the higher the cortisol increase from mid-morning to mid-afternoon (see Fig. 4 for the forest plot). The trim and fill method did not demonstrate publication bias and Rosenthal's fail-safe number indicated that 24 additional studies with an average sample size and effect size of zero would be needed to bring the p-value to non-significance, which indicates the outcome was rather robust. The Q-value was significant (p = .026), which indicates that the effect size varied significantly between studies. However, no significant moderators were found.

4. Discussion

Out-of-home child care is more used and valued than ever, partly because of the recent pandemic, which showed us clearly that accessible and high-quality out-of-home child care is vital to keep our societies functioning. The current study was an extension of two previous meta-analyses (Geoffroy et al., 2006; Vermeer & Van IJzendoorn, 2006). We not only included more studies but also added meta-analytical evidence with regard to potential correlates and moderators, since it is relevant to study whether and under what conditions young children's cortisol levels are elevated at child care. In the first series of meta-analyses, the within-child differences in cortisol levels between the child care setting and the home environment were investigated. As expected, no differences were found for the mid-morning cortisol values. Furthermore, results showed that both the mid-afternoon cortisol values and the cortisol increase from mid-morning to mid-afternoon were higher on child care days than on days that children were at home, with small to medium effect sizes. On average, an increase in cortisol over the day was found for the child care setting, as opposed to a decrease for the home setting. Furthermore, the within-child differences between settings were moderated by child age, country, and method of saliva collection. Regarding the correlates, we found a positive association between hours at child care and cortisol change over the day. Furthermore, a negative relation between child care quality and cortisol change was found, although only for studies that included home-based child care settings. Child temperament and group size were not associated with cortisol increases.

4.1. Cortisol levels at child care versus home

The higher cortisol levels at child care versus home that we found correspond with the outcomes of the former meta-analyses, and the effect sizes of the current study lie in between the effect sizes found in these earlier studies (Geoffroy et al., 2006; Vermeer & Van IJzendoorn, 2006). Some general explanations for this higher cortisol secretion in child care compared to the home setting are the potential stressfulness of the separation from parents and the complex interactions with peers and professional caregivers, as being separated from their parents and navigating in a room full of varying adults and peers are major developmental tasks for all young children (Vermeer & Groeneveld, 2017). Another related explanation might be that higher cortisol levels in the afternoon and the accompanying increase over the child care day may be attributed to child allostatic overload, which is the physiological accumulation of exposure to stress, when the demands exceed what the body can cope with. However, cortisol was measured mid-morning and therefore we could not study the immediate reactions to arriving at the child care center. The difference between the mid-morning and mid-afternoon values between settings is nonetheless striking. Since all mentioned factors are inherent to out-of-home child care and therefore difficult to disentangle and avoid, these explanations remain tentative, and do not truly add to our understanding. Therefore, it is interesting to look at correlates and moderators, now that we have confirmed the difference in cortisol secretion between the child care and home setting.

4.2. Moderators of cortisol levels at child care versus home

With regard to the moderators, the analyses showed that child age moderated the outcomes for the diurnal cortisol change score. Studies that included younger children (infants, toddlers or a combination of both infants and toddlers) did not find a higher cortisol increase over the day, while the studies that included older children (preschoolers or a combination of both toddlers and preschoolers) did. This finding resembles one of the earlier meta-analyses (Geoffroy et al., 2006). The moderating role of children's age on cortisol secretion in child care has been suggested to be related to social interaction. For infants, interactions with peers are rare, and while school-aged children have gained some social skills to handle social interactions, these are still limited in toddlers and preschoolers. This might lead to a curvilinear relation between child age and cortisol (Geoffroy et al., 2006; Vermeer & Groeneveld, 2017; Vermeer & Van IJzendoorn, 2006), with cortisol levels peaking for toddlers and preschoolers (Vermeer & Groeneveld, 2017). In the current study, we could only make a general distinction between studies that included younger (infants, toddlers or a combination of both infants and toddlers) and studies that included older children (preschoolers or a combination of both toddlers and preschoolers). Studies including children with a more defined age range are needed to examine this curvilinear hypothesis. The higher cortisol increases for older children could also be attributed to other variables, such as hours in care and the amount of experience in care (Geoffroy et al., 2006; Watamura et al., 2003), and less frequent but more scheduled napping time. Indeed, it was found that cortisol decreases when children are sleeping or only resting (Watamura et al., 2002), so this might also have driven the different results for both age groups. Moreover, it should be noted that the study of cortisol secretion in infants has some particular challenges due to infants' early-staged biological maturation (Tollenaar et al., 2010; Tryphonopoulos et al., 2014), which makes their cortisol levels somewhat difficult to compare directly to cortisol secretion in older children (toddlers and preschoolers). Therefore, it seems too soon to conclude that younger children experience less stress in child care.

The outcome for diurnal cortisol change was also moderated by country: the studies that were carried out in the United States showed that children displayed significantly higher cortisol increases over the day at child care compared to home, while the studies that were conducted outside of the United States did not find this difference. Since many studies conducted in the United States also included relatively older children, the relation with country might be an artefact driven by the above-mentioned moderator effect of child age. However, country

Study name	Outcome	Statistics for each study				
		Correlation	Lower limit	Upper limit	Z-Value	p-Value
Park & Choi (2009)	day	0.360	0.191	0.508	4.024	0.000
Drugli et al. (2018)	day	0.083	-0.104	0.264	0.869	0.385
Vermeer et al. (2010a)	week	0.140	-0.226	0.471	0.746	0.456
Vermeer et al. (2010b)	week	0.171	-0.240	0.530	0.810	0.418
Gunnar et al. (2010)	week	-0.059	-0.219	0.104	-0.706	0.480
Vaillancourt et al. (2018)	week	0.084	-0.068	0.232	1.085	0.278
Lumian et al. (2016b)	week	0.325	0.111	0.511	2.924	0.003
Tervahartiala et al. (2019)	month	0.122	-0.071	0.305	1.241	0.215
		0.148	0.036	0.256	2.589	0.010

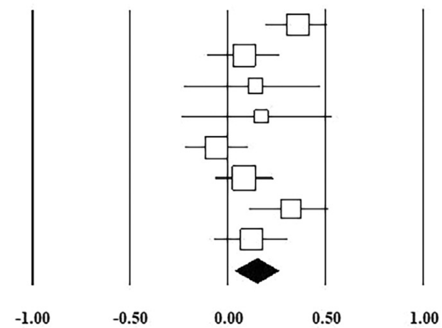

Fig. 5. Mean cortisol change (in μg/dl) from mid-morning to mid-afternoon for the child care and home setting.

(and not child age) was also found to be a moderator of mid-afternoon cortisol: although both studies carried out in the United States and studies that were conducted elsewhere showed higher cortisol values in the afternoon at child care, this effect was significantly higher for the studies from the United States. In the United States, it is common for children to attend full-time child care (Alexander & Network, 2005), whereas this is less common for children in some other included countries (such as the Netherlands). The long hours at child care in the United States could partly explain the difference, considering the significant correlation between hours at child care and cortisol increase that we found. Another explanation might be that in the United States, the emphasis in child care settings seems to lie more on educational goals than in Europe. The accompanying educational activities that are implemented to reach these educational standards could be more demanding for children and subsequently result in higher cortisol levels. Indeed, Bassok et al. (2016) found that academic expectations in child care centers have increased between 1998 and 2010 in the United States.

The moderation of outcomes by method of saliva collection was surprising. Cotton devices are nowadays less used than synthetic devices (Puhakka & Peltola, 2020), mostly because it is thought that the molecular structure of cotton might interfere with the immunoassay (Tryphonopoulos et al., 2014). However, this does not explain why we found a difference between the child care and home setting for studies that used cotton devices in both settings and studies that used other devices. However, cotton dental rolls are in general less reliable, and therefore might have resulted in a difference between settings by coincidence. Furthermore, the significant difference between settings for cotton dental rolls could be a result of confounding factors, such as child age and country (the other moderators).

4.3. Correlates of diurnal cortisol change at child care

We also studied specific correlates that could explain the variance in cortisol levels at child care. The meta-analysis on the association between hours at child care and cortisol increase over the day showed a small effect. The combined studies reported a positive association of cortisol increase with hours at child care, but the individual studies differed in the way they measured the amount of hours. Therefore, it is not exactly clear yet what matters most regarding elevated cortisol levels: how many hours per day the child attends the child care setting or how many hours per week or even per month. All quantities could explain the correlation, since allostatic overload, as mentioned before, might play a role both on a daily and on a weekly basis. Lumian et al. (2016a) found that children attending full-day child care showed higher cortisol increases over the day compared to children attending half-day child care and the same result was found for children attending full-time child care versus children attending part-time child care (Lumian et al., 2016b). However, this was only tested between and not within children. It seems that children attending half-day child care have the opportunity to recover in the afternoon and children attending part-time child care have the opportunity to recover on home days.

The meta-analysis on child care quality did not show a significant relation with cortisol increase. However, when type of care was included as a moderator, the relation between quality of care and cortisol increase was significant for studies with home-based child care settings and not for studies that included center-based child care. This result needs to be interpreted cautiously though, since only three studies into home-based child care were included. A speculative explanation for this finding could be that in home-based child care settings, most of the time only one or two caregivers are present, who therefore determine the quality of the setting for a larger part. In center-based child care, multiple caregivers are responsible, which might lead to less overall variation in child care quality for center-based compared to home-based child care. However, quality measures are mostly designed for center-based child care settings, and results are therefore hard to compare. Moreover, both child care settings differ in more respects, and therefore the reasons behind these findings are probably more complex, with other factors (e.g., other regulations, smaller group sizes, lower mean child age in home-based care compared to center-based care) playing a confounding role. The fact that a main effect for child care quality was absent, is however a remarkable result, since earlier reviews concluded that child care quality seems to matter, based on the studies conducted so far. It might be that limited variance in child care quality was responsible for the absence of significant results in the current meta-analysis, as most studies included child care centers of moderate to high quality. Furthermore, a large portion of the used instruments measuring child care quality is focused on structural quality (e.g., space and furnishings) and not so much on process quality (e.g., child-caregiver interactions), with the latter being more directly linked to child outcomes. This makes caution in interpreting the results related to child care quality even more important.

According to the current meta-analysis, an association between child temperament and an increasing cortisol pattern in child care is absent. It is possible that having a more difficult temperament only acts as a moderator, for example between child care quality and cortisol (Pluess & Belsky, 2009), or is not related to cortisol increase over the day, but to higher absolute levels, which could not be investigated in the current study.

The correlation between group size and cortisol increase was not significant either. Speculatively, it could be that the total number of adults at the child care group is most influential, which translates into the child-caregiver ratio. Although one can argue that more caregivers (a lower child-caregiver ratio) allow for more personal attention per child (e.g., De Schipper et al., 2006), too high a number of adults can have a negative effect on received caregiving due to the "displacement effect", which has been described as the risk for professional caregivers to focus more on interactions with each other than on interactions with the children (Vermeer et al., 2010). Therefore, the lack of a significant result could also be the consequence of a curvilinear relation between group size and cortisol secretion.

Finally, as the mean child age for the study samples together was rather low ($M = 2.79$ years), this could have influenced all findings substantively, as child care settings for younger children (infants) differ in multiple aspects (e.g., group size, schedule, and napping times) from child care settings for toddlers and preschoolers. This intertwining of child age and other substantial factors should therefore be kept in mind when interpreting the results.

4.4. Limitations and future directions

Although the current study contributes to the ongoing scientific discourse, some limitations are worth mentioning. First, not all relevant studies could be included because not all authors reported the information we needed and were unable to or did not deliver the information on request. We therefore tried to find a balance between being too strict and too inclusive when selecting studies. Second and related to this, dissimilarities across studies made it difficult to compare effect sizes. However, we made conservative choices when in doubt, although the exclusion of incompatible articles might have led to inadequate power to detect certain associations and differences for the correlational and moderator analyses. Therefore, larger meta-analyses with future studies will be necessary in order to be able to draw firmer conclusions.

Furthermore, although the within-child design of this study has many advantages, it does not allow for comparisons with children that are full-time cared for at home. By making this comparison, we could gain more insight into whether the cortisol levels of children attending out-of-home child care on home days are comparable to the cortisol levels of children that do not attend out-of-home child care at all. This would subsequently give us more information on whether children's cortisol levels return to baseline or whether levels remain slightly elevated (or less decreased) when children are at home, after the child care day ends or on another day.

Regarding future research, we require more studies in more different countries to test the cross-cultural validity of the outcomes. Additionally, we should try to include child care settings with a larger variety in child care quality. It might be challenging to find settings of low quality willing to participate in research, but we should persist in this endeavor when the experiences of children in settings of lower quality are now largely overlooked in most studies.

In order to unravel potential long-term consequences of elevated cortisol levels in out-of-home child care, longitudinal studies should be given priority as well, such as the study by Roisman et al. (2009) that found more experience in center-based child care in the first three years of life to uniquely predict a lower cortisol awakening response at age 15. The same holds for experimental studies, in which the causes of elevated cortisol levels could be closely examined by changing certain aspects of the child care setting while keeping all other factors constant, which was done for example in the study by De Schipper and colleagues (2006), by altering the child-caregiver ratio. Of course, researchers should make careful ethical considerations when designing experimental studies.

Additionally, we want to stress the importance for future studies into this topic to include as much information on sample characteristics (e.g., group size and stability) and methods as possible and to control for confounding variables (e.g., eating and napping times), since this benefits and stimulates replication and overview studies. Moreover, additional information on the home circumstances of participating children in cortisol studies would be important to include (e.g., SES, attachment), since this allows for comparing stress at the home versus child care setting more meaningfully.

Studying interactions between child care characteristics, such as hours at child care and child care quality, is also vital for understanding the complex interplay of factors. Broekhuizen et al. (2017) for example found children spending 3.5 days or more in high quality child care centers to show lower levels of externalizing behavior problems one year later. Other potentially important variables to study in relation to children's physiological stress at child care are peer sociometric status, the qualifications of professional caregivers, and the structure of activities throughout the day. As these variables are now not very often included in studies but are theoretically plausible, this is an important direction for future research as well.

More studies into different settings in which the child is cared for by others than the parent, such as home-based child care or a one-on-one babysitter, could also be a valuable addition. This was for example done in a recent study by Tervahartiala and colleagues (2019), who found that overall cortisol levels were higher for the in-house child care group than for the out-of-home child care group (Tervahartiala et al., 2019). Finally, since the method of saliva collection was a significant moderator, researchers should carefully select their method of collection.

When considering the practical implications of the outcomes of the current study, we should first and foremost ask ourselves whether elevated cortisol levels are necessarily problematic, since not all elevations are clinical and therefore biologically significant (Gunnar et al., 2010). However, the consistently reported higher cortisol levels and cortisol increases in child care compared to home are robust and deserve our attention, although effect sizes are small to moderate. On the basis of the current paper, we would suggest parents and practitioners to consider half-day and/or part-time child care, when possible, to alleviate the burden on the child's stress system. If part-time care is not possible, child care centers could opt for more resting or quiet time during the day or create time-out spaces where children could play on their own for a while to regain homeostasis. Several studies indicating that children's cortisol levels return to baseline levels once they are home underscore the potential of such an implication (e.g., Watamura et al., 2009). One should however bear in mind that the effect size for the association between hours at child care and children's cortisol increases was small and more research into the ideal quantity for out-of-home child care (for different groups of children) is highly necessary.

The quest for additional variables co-explaining higher cortisol levels in child care compared to home therefore continues, since we only found that hours in care and, in some cases, quality of care were associated with cortisol increases. Alternatively, a cautious conclusion might also be that child care can be stressful in itself, and only part of the variance can be explained by characteristics of the child or the child care setting.

5. Conclusion

The current study meta-analytically reviewed available research on young children's physiological reactions to out-of-home child care. The finding that children show both higher cortisol levels in the afternoon and larger increases over the day at child care compared to home was replicated and therefore proved to be robust. Furthermore, the amount of hours at child care was found to be associated with cortisol increases, with more hours in care related to larger increases in cortisol at child care. Type of care, child age, country, and method of saliva collection moderated part of the outcomes. Although some questions remain unanswered, with the results of the current study we have shed light on potential correlates of elevated cortisol levels and implications for future research, which will increasingly expand our understanding of young children's experiences in out-of-home child care, and how we can reduce their stress levels in order for them to thrive.

Funding

This work was supported by a grant from the Netherlands Organization for Health Research and Development (grant number 544001004).

The funding source was not involved in the preparation of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Sanne M. de Vet: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft. **Claudia I. Vrijhof:** Conceptualization, Writing – review & editing. **Shelley M.C. van der Veek:** Writing – review & editing. **Mariëlle Linting:** Formal analysis, Methodology, Writing – review & editing. **Harriet J. Vermeer:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to express their gratitude to research assistant Tessa Verlaan for her help with the coding of the articles and to the many authors who responded to the request for additional data.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecresq.2022.12.015.

References

Ahnert, L., Gunnar, M. R., Lamb, M. E., & Barthel, M. (2004). Transition to child care: Associations with infant-mother attachment, infant negative emotion, and cortisol elevations. *Child Development, 75*(3), 639–650 References marked with an * were included in the meta-analyses. https://doi.org/10.1111/j.1467-8624.2004.00698.x.

Albers, E. M., Beijers, R., Riksen-Walraven, J., Sweep, F. C. G. J., & de Weerth, C. (2016). Cortisol levels of infants in center care across the first year of life: Links with quality of care and infant temperament. *Stress, 19*(1), 8–17. https://doi.org/10.3109/10253890.2015.1089230.

Alexander, D., & Network, N. E. C. C. R. (2005). *Child care and child development: Results from the NICHD study of early child care and youth development*. Guilford Press.

Badanes, L. S. (2010). *Cortisol reactivity across the day at child care: Examining the contributions of child temperament and attachment to mother and lead teacher* (Doctorial dissertation). *ProQuest*. https://search-proquest-com.ezproxy.leidenuniv.nl:2443/doYiew/304880122?pq-origsite=primo.

Badanes, L. S., Dmitrieva, J., & Watamura, S. E. (2012). Understanding cortisol reactivity across the day at child care: The potential buffering role of secure attachments to caregivers. *Early Childhood Research Quarterly, 27*(1), 156–165. https://doi.org/10.1016/j.ecresq.2011.05.005.

Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade? *AERA Open, 1*(4), 1–31. https://doi.org/10.1177/2332858415616358.

Belsky, J. (1986). Infant day care: A cause for concern? *Zero to Three, 7*(1), 1–7.

Bernard, K., Peloso, E., Laurenceau, J. P., Zhang, Z., & Dozier, M. (2015). Examining change in cortisol changes during the 10-week transition to a new child-care setting. *Child Development, 86*(2), 456–471. https://doi.org/10.1111/cdev.12304.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley.

Borenstein, M., Rothstein, D., & Cohen, J. (2005). *Comprehensive meta-analysis: A computer program for research synthesis*. Biostat: [Computer software].

Broekhuizen, M. L., van Aken, M. A. G., Dubas, J. S., & Leseman, P. P. M. (2017). Child care quality and Dutch 2- and 3-year-olds' socio-emotional outcomes: Does the amount of care matter? *Infant and Child Development, 27*(1), e2043 -n/a. https://doi.org/10.1002/icd.2043.

De Haan, M., Gunnar, M. R., Tout, K., Hart, J., & Stansbury, K. (1998). Familiar and novel contexts yield different associations between cortisol and behavior among 2-year-old children. *Developmental Psychobiology, 33*(1), 93–101. https://doi.org/10.1002/(SICI)1098-2302(199807)33:1<93::AID-DEV8>3.0.CO;2-N.

De Schipper, E. J., Riksen-Walraven, J. M., & Geurts, S. A. E. (2006). Effects of child-caregiver ratio on the interactions between caregivers and children in child-care centers: An experimental study. *Child Development, 77*(4), 861–874. https://doi.org/10.1111/j.1467-8624.2006.00907.x.

Dettling, A. C., Gunnar, M. R., & Donzella, B. (1999). Cortisol levels of young children in full-day childcare centers: Relations with age and temperament. *Psychoneuroendocrinology, 24*(5), 519–536. https://doi.org/10.1016/S0306-4530(99)00009-8.

Dettling, A. C., Parker, S. W., Lane, S., Sebanc, A., & Gunnar, M. R. (2000). Quality of care and temperament determine changes in cortisol concentrations over the day for young children in childcare. *Psychoneuroendocrinology, 25*(8), 819–836. https://doi.org/10.1016/S0306-4530(00)00028-7.

De Weerth, C., & Van Geert, P. (2002). A longitudinal study of basal cortisol in infants: Intra-individual variability, circadian rhythm and developmental trends. *Infant Behavior and Development, 25*(4), 375–398. https://doi.org/10.1016/S0163-6383(02)00141-8.

De Weerth, C., Zijl, R. H., & Buitelaar, J. K. (2003). Development of cortisol circadian rhythm in infancy. *Early Human Development, 73*(1–2), 39–52. https://doi.org/10.1016/S0378-3782(03)00074-4.

Drugli, M. B., Solheim, E., Lydersen, S., Moe, V., Smith, L., & Berg-Nielsen, T. S. (2018). Elevated cortisol levels in Norwegian toddlers in childcare. *Early Child Development and Care, 188*(12), 1682–1693. https://doi.org/10.1080/03004430.2016.1278368.

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*, 89–98. https://doi.org/10.2307/2669529.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*, 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x.

Eckstein-Madry, T., Piskernik, B., & Ahnert, L (2021). Attachment and stress regulation in socioeconomically disadvantaged children: Can public childcare compensate? *Infant Mental Health Journal, 42*(6), 839–850. https://doi.org/10.1002/imhj.21878.

Geoffroy, M.-C., Côté, S. M., Parent, S., & Séguin, J. R. (2006). Daycare attendance, stress, and mental health. *The Canadian Journal of Psychiatry, 51*(9), 607–615. https://doi.org/10.1177/070674370605100909.

Groeneveld, M. G., Vermeer, H. J., Van IJzendoorn, M. H., & Linting, M. (2010). Children's wellbeing and cortisol levels in home-based and center-based childcare. *Early Childhood Research Quarterly, 25*(4), 502–514. https://doi.org/10.1016/j.ecresq.2009.12.004.

Groeneveld, M. G., Vermeer, H. J, Van IJzendoorn, M. H., & Linting, M. (2012). Stress, cortisol and well-being of caregivers and children in home-based child care: A case for differential susceptibility. *Child Care Health and Development, 38*(2), 251–260. https://doi.org/10.1111/j.1365-2214.2010.01194.x.

Gunnar, M. R. (2021). Forty years of research on stress and development: What have we learned and future directions. *American Psychologist, 76*(9), 1372–1384. https://doi.org/10.1037/amp0000893.

Gunnar, M. R., & Donzella, B. (2002). Social regulation of the cortisol levels in early human development. *Psychoneuroendocrinology, 27*(1-2), 199–220. https://doi.org/10.1016/S0306-4530(01)00045-2.

Gunnar, M. R., Kryzer, E., Van Ryzin, M. J., & Phillips, A. (2010). The rise in cortisol in family day care: Associations with aspects of care quality, child behavior, and child sex. *Child Development, 81*(3), 851–869. https://doi.org/10.1111/j.1467-8624.2010.01438.x.

Gunnar, M. R., Tout, K., de Haan, M., Pierce, S., & Stanbury, K. (1997). Temperament, social competence, and adrenocortical activity in preschoolers. *Developmental Psychobiology, 31*(1), 65–85. https://doi.org/10.1002/(SICI)1098-2302(199707)31:1<65::AID-DEV6>3.0.CO;2-S.

Hatfield, B. E., Hestenes, L. L., Kintner-Duffy, V. L., & O'Brien, M (2013). Classroom emotional support predicts differences in preschool children's cortisol and alpha-amylase levels. *Early Childhood Research Quarterly, 28*(2), 347–356. https://doi.org/10.1016/j.ecresq.2012.08.001.

Ivars, K., Nelson, N., Theodorsson, A., Theodorsson, E., Ström, J. O., & Mörelius, E. (2015). Development of salivary cortisol circadian rhythm and reference intervals in full-term infants. *PLoS ONE, 10*(6), 1–13. https://doi.org/10.1371/journal.pone.0129502.

Iwata, O., Okamura, H., Saitsu, H., Saikusa, M., Kanda, H., Eshima, N., Iwata, S., Maeno, Y., & Matsuishi, T. (2013). Diurnal cortisol changes in newborn infants suggesting entrainment of peripheral circadian clock in utero and at birth. *Journal of Clinical Endocrinology and Metabolics, 98*, 25–32.

Lisonbee, J. A., Mize, J., Payne, A. L., & Granger, D. A. (2008). Children's cortisol and the quality of teacher-child relationships at child care. *Child Development, 79*(6), 1818–1832. https://doi.org/10.1111/j.1467-8624.2008.01228.x.

Lumian, D. S., Dmitrieva, J., Mendoza, M. M., Badanes, L. S., & Watamura, S. E. (2016). The impact of program structure on cortisol changeing in children attending out-of-home child care. *Early Childhood Research Quarterly, 34*, 92–103. https://doi.org/10.1016/j.ecresq.2015.09.004.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, The PRISMA (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *PLOS Medicine, 6*(7), Article e1000097. https://doi.org/10.1371/journal.pmed1000097.

National Institute of Child Health and Human Development (2006). *The NICHD study of early child care and youth development. Findings for children up to age 4,5 years.* https://www.nichd.nih.gov/sites/default/files/publications/pubs/Documents/SECCYD_06.pdf

Organization for Economic Co-operation and Development (2021). *Family Database, Public policies for families and children, Formal care and education for very young children, PF3.2: Enrolment in childcare and pre-school.* https://www.oecd.org/els/soc/PF3_2_Enrolment_childcare_preschool.pdf

Ouellet-Morin, I., Tremblay, R. E., Boivin, M., Meaney, M., Kramer, M., & Côté, S. M (2010). Diurnal cortisol secretion at home and at child care: A prospective study of 2-year-old toddlers. *Journal of Child Psychology and Psychiatry, 51*(3), 295–303. https://doi.org/10.1111/j.1469-7610.2009.02167.x.

*Park, K. J., & Choi, J. Y. (2009). Children's cortisol levels at child care in South Korea, presented at the SRCD Biennial Meeting, Denver, 2009.

Pluess, M., & Belsky, J. (2009). Differential susceptibility to rearing experience: The case of childcare. *Journal of Child Psychology and Psychiatry, 50*(4), 396–404. https://doi.org/10.1111/j.1469-7610.2008.01992.x.

Puhakka, I. J. A., & Peltola, M. J. (2020). Salivary cortisol reactivity to psychological stressors in infancy: A meta-analysis. *Psychoneuroendocrinology, 115*, Article 104603. https://doi.org/10.1016/j.psyneuen.2020.104603.

Reunamo, J., Sajaniemi, N., Suhonen, E., & Kontu, E. (2012). Cortisol levels and children's orientation in day care. *Early Child Development and Care, 182*(3-4), 363–381. https://doi.org/10.1080/03004430.2011.646727.

Roisman, G. I., Susman, E., Barnett-Walker, K., Booth-LaForce, C., Owen, M. T., Belsky, J., Bradley, R. H., Houts, R., & Steinberg, L. (2009). Early family and child-care antecedents of awakening cortisol levels in adolescence: Early family and child-care antecedents. *Child Development, 80*(3), 907–920. https://doi.org/10.1111/j.1467-8624.2009.01305.x.

Rosen, J. B., & Schulkin, J. (1998). From normal fear to pathological anxiety. *Psychological Review, 105*(2), 325–350. https://doi.org/10.1037/0033-295X.105.2.325.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin,, 86*(3), 638–664. https://doi.org/10.1037/0033-2909.86.3.638.

Sajaniemi, N., Suhonen, E., Kontu, E., Rantanen, P., Lindholm, H., Hyttinen, S., & Hirvonen, A. (2011). Children's cortisol changes and the quality of the early learning environment. *European Early Childhood Education Research Journal, 19*(1), 45–62. https://doi.org/10.1080/1350293x.2011.548938.

Sippell, W. G., Becker, H., Versmold, H. T., Bidlingmaier, F., & Knorr, D. (1978). Longitudinal studies of plasma aldosterone, corticosterone, deoxycosterone, progesterone 17-hydroxyprogesterone, cortisol, and cortisone determined simultaneously in mother and child at birth and during the early neonatal period. I. Spontaneous delivery. *Journal of Clinical Endocrinology and Metabolism, 46*, 971–985. https://doi.org/10.1210/jcem-46-6-971.

Sonia, J. L., Bruce, S. M., Megan, R. G., & Christine, H. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Reviews Neuroscience, 10*(6), 434. https://doi.org/10.1038/nrn2639.

Sumner, M. M. (2009). *Full-day patterns of cortisol production among toddlers at child care* (Doctoral dissertation). *ProQuest*. https://search-proquest-com.ezproxy.leidenuniv.nl:2443/doYiew/305322655/?pq-origsite=primo.

Sumner, M. M., Bernard, K., & Dozier, M. (2010). Young children's full-day patterns of cortisol production on child care days. *Archives of Pediatrics & Adolescent Medicine, 164*(6), 567–571. https://doi.org/10.1001/archpediatrics.2010.85.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon.

Tervahartiala, K., Karlsson, L., Pelto, J., Kortesluoma, S., Hyttinen, S., Ahtola, A., Junttila, N., & Karlsson, H (2019). Toddlers' diurnal cortisol levels affected by out-of-home, center-based childcare and at-home, guardian-supervised childcare: Comparison between different caregiving contexts. *European Child & Adolescent Psychiatry*. https://doi.org/10.1007/s00787-019-01432-3.

Tollenaar, M. S., Jansen, J., Beijers, R., Riksen-Walraven, J. M., & de Weerth, C. (2010). Cortisol in the first year of life: Normative values and intra-individual variability. *Early Human Development, 86*(1), 13–16. https://doi.org/10.1016/j.earlhumdev.2009.12.003.

Tout, K., De Haan, M., Campbell, E. K., & Gunnar, M. R. (1998). Social behavior correlates of cortisol activity at child care: Gender differences and time-of-day effects. *Child Development, 69*(5), 1247–1262. https://doi.org/10.2307/1132263.

Tryphonopoulos, P. D., Letourneau, N., & Azar, R. (2014). Approaches to salivary cortisol collection and analysis in infants. *Biological Research for Nursing, 16*(4), 398–408. https://doi.org/10.1177/1099800413507128.

Vaillancourt, T., Brittain, H., Haltigan, J. D., Ostrov, J. M., & Muir, C. (2018). Cortisol moderates the relation between physical peer victimization and physical aggression in preschoolers attending high-quality child care: Evidence of differential susceptibility across informants. *Merrill-Palmer Quarterly: Journal of Developmental Psychology, 64*(1), 101–134. https://doi.org/10.13110/merrpalmquar1982.64.1.0101.

Vermeer, H. J., & Groeneveld, M. G. (2017). Children's physiological responses to childcare. *Current Opinion in Psychology, 15*, 201–206. https://doi.org/10.1016/j.copsyc.2017.03.006.

Vermeer, H. J., Groeneveld, M. G., Larrea, I., Van IJzendoorn, M. H., Barandiaran, A., & Linting, M. (2010). Child care quality and children's cortisol in Basque country and the Netherlands. *Journal of Applied Developmental Psychology, 31*(4), 339–347. https://doi.org/10.1016/j.appdev.2010.05.001.

Vermeer, H. J., & Van IJzendoorn, M. H. (2006). Children's elevated cortisol levels at daycare: A review and meta-analysis. *Early Childhood Research Quarterly, 21*(3), 390–401. https://doi.org/10.1016/j.ecresq.2006.07.004.

Watamura, S. E., Coe, C. L., Laudenslager, M. L., & Robertson, S. S. (2010). Child care setting affects salivary cortisol and antibody secretion in young children. *Psychoneuroendocrinology, 35*(8), 1156–1166. https://doi.org/10.1016/j.psyneuen.2010.02.001.

Watamura, S. E., Donzella, B., Alwin, J., & Gunnar, M. R. (2003). Morning-to-afternoon increases in cortisol concentrations for infants and toddlers at child care: Age differences and behavioral correlates. *Child Development, 74*(4), 1006–1020. https://doi.org/10.1111/1467-8624.00583.

Watamura, S. E., Kryzer, E. M., & Robertson, S. S. (2009). Cortisol changes at home and child care: Afternoon differences and evening recovery in children attending very high quality full-day center-based child care. *Journal of Applied Developmental Psychology, 30*(4), 475–485. https://doi.org/10.1016/j.appdev.2008.12.027.

Watamura, S. E., Sebanc, A. M., & Gunnar, M. R. (2002). Rising cortisol at childcare: Relations with nap, rest, and temperament. *Developmental Psychobiology, 40*(1), 33–42. https://doi.org/10.1002/dev.10011.