Early-life endocrine regulation and neurodevelopmental outcomes

Hollanders, J.J.

2020

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)
Hollanders, J. J. (2020). *Early-life endocrine regulation and neurodevelopmental outcomes*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam].

Early-life endocrine regulation and neurodevelopmental outcomes

Jonneke J. Hollanders

VRIJE UNIVERSITEIT

EARLY-LIFE ENDOCRINE REGULATION AND NEURODEVELOPMENTAL OUTCOMES

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op donderdag 12 maart 2020 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Josephina Jenneke Hollanders
geboren te Delft

TABLE OF CONTENTS

DISCUSSION

Introduction

1

General introduction and outline of thesis

This thesis is centered around the premise that occurrences early in life, or even ante-natally, can have effects in the long-term. This hypothesis is called the Developmental Origins of Health and Disease (DOHaD). An extensive body of evidence has already shown that adverse events in early life can lead to an increased risk of, among others, cardiovascular disease [1] and psychopathology.[2]

Prematurely born infants form a special risk group. Their safe place in the womb is traded for an incubator in the Neonatal Intensive Care Unit (NICU), where they can be subjected to interventions such as ventilation assistance, routine controls, a myriad of medications, surgeries, and other (painful) procedures. Their bodies are also still imma-ture, which hampers their ability to digest milk, to breathe unassisted and to fight off infections, all while being constantly exposed to external risks.

Globally, approximately 10.6% of births occurred prematurely, and 15.4% of those were with a gestational age of <32 weeks.[3] With advancing medical developments in the NICU, such as aggressive feeding strategies, antenatal glucocorticoids and improved ventilation techniques, mortality has decreased significantly in preterm populations.[4-6] Therefore, focus is shifting towards improving long-term outcomes in these children. Preterm infants have previously been shown to be at an increased risk for neurodevel-opmental problems, cardiovascular diseases and deviating growth.[7-9] It is important to know which factors contribute to these adverse outcomes, and it is equally important to study which interventions can improve or prevent the adverse consequences of preterm birth. Treatment of preterm infants is likely to be most successful when normal physiol-ogy is pursued, and a comprehensive understanding of these processes is therefore crucial.

The work presented in this thesis aimed to elucidate both normal physiology, as well as some of the factors contributing to or preventing adverse outcomes in preterm infants, with a focus on early-life endocrine regulation.

PART 1 "EARLY-LIFE GLUCOCORTICOID REGULATION"

A mal-adapted hypothalamus pituitary adrenal (HPA) axis has been implicated as one of the underlying mechanisms behind the DOHaD hypothesis.[10,11] However, not much is known yet about normal fetal and neonatal HPA-axis development, and recognizing aberrant developmental patterns is therefore difficult. In this part, we aimed to shed more light on normal HPA-axis development and its influencing factors.

Glucocorticoids (GCs) can be measured in hair, which offers a retrospective view of HPA-axis activity.[12] We aimed to explore whether this medium provides a reliable insight into fetal HPA-axis activity, and which factors are associated with neonatal hair GC levels

in **Chapter 2**. In **Chapter 3**, we analyzed the association between experienced maternal distress pre- and perinatally and hair GC levels in the neonate and mother.

Exposure to aberrant maternal cortisol levels in utero has been associated with adverse outcomes in the offspring.[13] After birth, infants are still exposed to small amounts of maternal GCs through breastmilk. Several studies have found associations between breastmilk GCs and outcomes in both animal and human studies.[14-21] However, our research group recently reported that GCs in breastmilk follow the diurnal rhythm of maternal HPA-axis activity,[22] and this was not taken into consideration by previous studies. We have reviewed existing evidence concerning breastmilk GCs in **Chapter 4**. Next, we have explored associations between breastmilk GC rhythmicity and (neurodevelopmental) outcomes in the offspring. We assessed the correlation between breastmilk GCs and macronutrients in **chapter 5**, to determine whether associations between breastmilk GCs and outcomes in offspring could actually be attributed to macronutrient variations instead. Subsequently, we described GC rhythmicity in infants at age 1 month and explored associations between this rhythm and breastmilk GC rhythmicity as well as other possible rhythm-influencing factors (**Chapter 6**). Lastly, we researched the associations between breastmilk GC rhythmicity and infant body composition (**Chapter 7**) and behavior and sleep (**Chapter 8**).

PART 2 "GLUCOCORTICOID REGULATION AND SEX"

Sex differences in the production and metabolism of cortisol are present in adults, which have been suggested to arise during puberty under the influence of sex steroids.[23,24] However, sex differences in mortality and short- and long-term morbidity are already present in preterm populations. To explore whether these differences might be partly caused by sex differences in cortisol levels, we performed a systematic review and meta-analysis with regard to basal cortisol levels (**Chapter 9**) as well as a systematic review concerning sex differences in HPA-axis reactivity (**Chapter 10**).

PART 3 "EARLY-LIFE THYROID REGULATION IN PRETERM INFANTS"

Maternal hypothyroxinaemia and congenital hypothyroidism have been associated with adverse neurodevelopmental outcomes (in offspring).[25-27] Transient hypothyroxinaemia of prematurity (THoP), a condition in which circulating T4 concentrations are low due to immature endocrine systems as well as acute illnesses, has also been associated with adverse neurodevelopmental outcomes in infancy and childhood.[28-30] However, it is unclear whether these adverse outcomes persist into adolescence and adulthood. We

therefore used the data of the Project On Preterm and Small-for-gestational-age (POPS) cohort to assess whether THoP was associated with IQ and neuromotor outcomes (**Chapter 11**) as well as behavioral outcomes (**Chapter 12**) at age 19.

PART 4 "EARLY-LIFE GROWTH AND NEURODEVELOPMENT"

Both infants who are born very preterm (VP, i.e., gestational age <32 weeks) and/or who are born with a very low birth weight (VLBW, i.e., birth weight <1,500 grams) require admission to a NICU. Many of these infants are both VP and VLBW, and results of studies in one research population are therefore often applied to the other research population. Nonetheless, previous studies have shown that short-term outcomes differ between infants who are born VP versus those with VLBW.[31] We explored whether long-term outcomes were also different between these two entities. First, using the data of the POPS cohort, we assessed differences in growth and final height between children who were born VP and/or with a VLBW (**Chapter 13**). Next, we also analyzed differences in IQ, neuromotor outcomes, behavior, and functional outcomes at age 19 years between these populations (**Chapter 14**).

   Lastly, we explored whether improved care has led to different growth patterns and long-term growth and neurodevelopmental outcomes in two preterm cohorts, established 20 years apart. We analyzed the occurrence of prenatal and postnatal growth restriction, whether these growth patterns are associated with long-term growth and neurodevelopmental outcomes, and whether these associations changed between cohorts (**Chapter 15**).
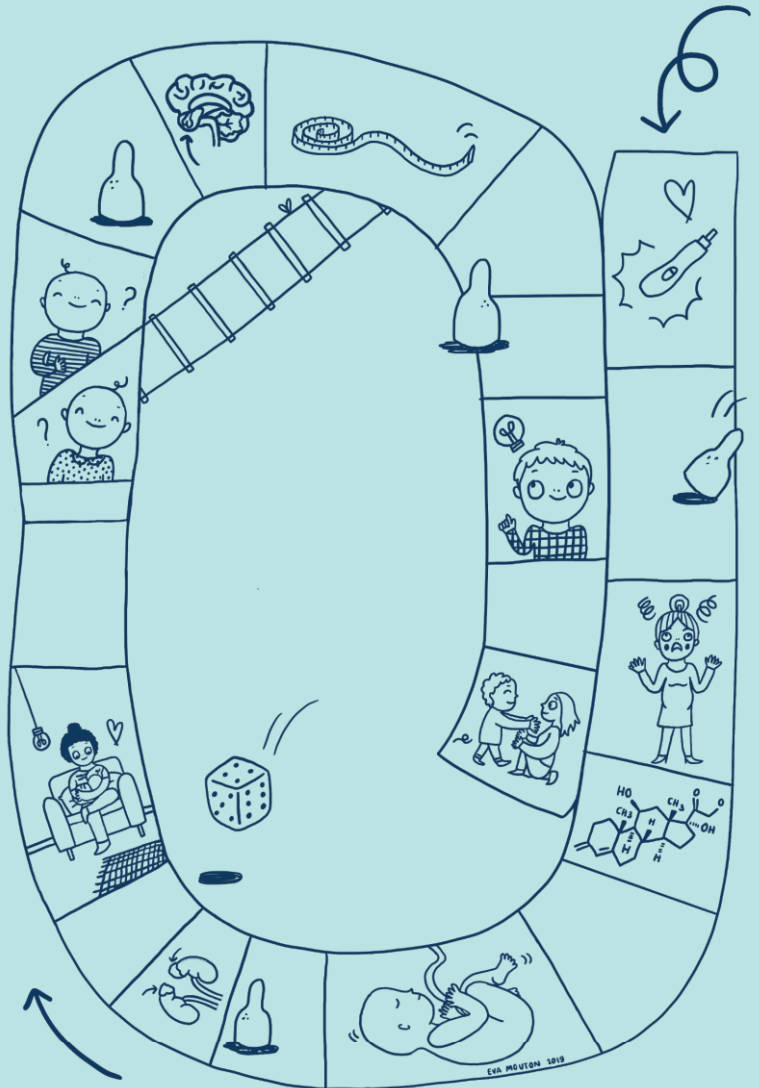
REFERENCES

1. Barker DJ, Osmond C. Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. Lancet 1986; 1:1077-1081
2. O'Donnell KJ, Meaney MJ. Fetal Origins of Mental Health: The Developmental Origins of Health and Disease Hypothesis. Am J Psychiatry 2017; 174:319-328
3. Chawanpaiboon S, Vogel JP, Moller AB, Lumbiganon P, Petzold M, Hogan D, Landoulsi S, Jampathong N, Kongwattanakul K, Laopaiboon M, Lewis C, Rattanakanokchai S, Teng DN, Thinkhamrop J, Watananirun K, Zhang J, Zhou W, Gulmezoglu AM. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. Lancet Glob Health 2019; 7:e37-e46
4. Senterre T, Rigo J. Reduction in postnatal cumulative nutritional deficit and improvement of growth in extremely preterm infants. Acta Paediatr 2012; 101:e64-70
5. Stoelhorst GM, Rijken M, Martens SE, Brand R, den Ouden AL, Wit JM, Veen S. Changes in neonatology: comparison of two cohorts of very preterm infants (gestational age <32 weeks): the Project On Preterm and Small for Gestational Age Infants 1983 and the Leiden Follow-Up Project on Prematurity 1996-1997. Pediatrics 2005; 115:396-405
6. Stoll BJ, Hansen NI, Bell EF, Walsh MC, Carlo WA, Shankaran S, Laptook AR, Sanchez PJ, Van Meurs KP, Wyckoff M, Das A, Hale EC, Ball MB, Newman NS, Schibler K, Poindexter BB, Kennedy KA, Cotten CM, Watterberg KL, D'Angio CT, DeMauro SB, Truog WE, Devaskar U, Higgins RD. Trends in Care Practices, Morbidity, and Mortality of Extremely Preterm Neonates, 1993-2012. Jama 2015; 314:1039-1051
7. Euser AM, de Wit CC, Finken MJ, Rijken M, Wit JM. Growth of preterm born children. Horm Res 2008; 70:319-328
8. Kajantie E, Hovi P. Is very preterm birth a risk factor for adult cardiometabolic disease? Semin Fetal Neonatal Med 2014; 19:112-117
9. Twilhaar ES, Wade RM, de Kieviet JF, van Goudoever JB, van Elburg RM, Oosterlaan J. Cognitive Outcomes of Children Born Extremely or Very Preterm Since the 1990s and Associated Risk Factors: A Meta-analysis and Meta-regression. JAMA Pediatr 2018; 172:361-367
10. Finken MJ, van der Voorn B, Heijboer AC, de Waard M, van Goudoever JB, Rotteveel J. Glucocorticoid Programming in Very Preterm Birth. Horm Res Paediatr 2016; 85:221-231
11. Rosmond R, Bjorntorp P. The hypothalamic-pituitary-adrenal axis activity as a predictor of cardiovascular disease, type 2 diabetes and stroke. J Intern Med 2000; 247:188-197
12. Staufenbiel SM, Penninx BW, Spijker AT, Elzinga BM, van Rossum EF. Hair cortisol, stress exposure, and mental health in humans: a systematic review. Psychoneuroendocrinology 2013; 38:1220-1235
13. Duthie L, Reynolds RM. Changes in the maternal hypothalamic-pituitary-adrenal axis in pregnancy and postpartum: influences on maternal and fetal outcomes. Neuroendocrinology 2013; 98:106-115
14. Dettmer AM, Murphy AM, Guitarra D, Slonecker E, Suomi SJ, Rosenberg KL, Novak MA, Meyer JS, Hinde K. Cortisol in Neonatal Mother's Milk Predicts Later Infant Social and Cognitive Functioning in Rhesus Monkeys. Child Dev 2017;
15. Grey KR, Davis EP, Sandman CA, Glynn LM. Human milk cortisol is associated with infant temperament. Psychoneuroendocrinology 2013; 38:1178-1185
16. Hahn-Holbrook J, Le TB, Chung A, Davis EP, Glynn LM. Cortisol in human milk predicts child BMI. Obesity (Silver Spring) 2016; 24:2471-2474

17.   Hart S, Boylan LM, Border B, Carroll SR, McGunegle D, Lampe RM. Breast milk levels of cortisol and Secretory Immunoglobulin A (SIgA) differ with maternal mood and infant neuro-behavioral functioning. Infant Behav Dev 2004; 27:101-106

18.   Hinde K, Skibiel AL, Foster AB, Del Rosso L, Mendoza SP, Capitanio JP. Cortisol in mother's milk across lactation reflects maternal life history and predicts infant temperament. Behav Ecol 2015; 26:269-281

19.   Sullivan EC, Hinde K, Mendoza SP, Capitanio JP. Cortisol concentrations in the milk of rhesus monkey mothers are associated with confident temperament in sons, but not daughters. Dev Psychobiol 2011; 53:96-104

20.   Catalani A, Casolini P, Cigliana G, Scaccianoce S, Consoli C, Cinque C, Zuena AR, Angelucci L. Maternal corticosterone influences behavior, stress response and corticosteroid receptors in the female rat. Pharmacol Biochem Behav 2002; 73:105-114

21.   Catalani A, Casolini P, Scaccianoce S, Patacchioli FR, Spinozzi P, Angelucci L. Maternal corticosterone during lactation permanently affects brain corticosteroid receptors, stress response and behaviour in rat progeny. Neuroscience 2000; 100:319-325

22.   van der Voorn B, de Waard M, van Goudoever JB, Rotteveel J, Heijboer AC, Finken MJ. Breast-Milk Cortisol and Cortisone Concentrations Follow the Diurnal Rhythm of Maternal Hypothalamus-Pituitary-Adrenal Axis Activity. J Nutr 2016; 146:2174-2179

23.   McCormick CM, Lewis E, Somley B, Kahan TA. Individual differences in cortisol levels and performance on a test of executive function in men and women. Physiol Behav 2007; 91:87-94

24.   Wudy SA, Hartmann MF, Remer T. Sexual dimorphism in cortisol secretion starts after age 10 in healthy children: urinary cortisol metabolite excretion rates during growth. Am J Physiol Endocrinol Metab 2007; 293:E970-976

25.   Finken MJ, van Eijsden M, Loomans EM, Vrijkotte TG, Rotteveel J. Maternal hypothyroxinemia in early pregnancy predicts reduced performance in reaction time tests in 5- to 6-year-old offspring. J Clin Endocrinol Metab 2013; 98:1417-1426

26.   Leger J. Congenital hypothyroidism: a clinical update of long-term outcome in young adults. Eur J Endocrinol 2015; 172:R67-77

27.   Noten AM, Loomans EM, Vrijkotte TG, van de Ven PM, van Trotsenburg AS, Rotteveel J, van Eijsden M, Finken MJ. Maternal hypothyroxinaemia in early pregnancy and school performance in 5-year-old offspring. Eur J Endocrinol 2015; 173:563-571

28.   Delahunty C, Falconer S, Hume R, Jackson L, Midgley P, Mirfield M, Ogston S, Perra O, Simpson J, Watson J, Willatts P, Williams F. Levels of neonatal thyroid hormone in preterm infants and neurodevelopmental outcome at 5 1/2 years: millennium cohort study. J Clin Endocrinol Metab 2010; 95:4898-4908

29.   Meijer WJ, Verloove-Vanhorick SP, Brand R, van den Brande JL. Transient hypothyroxinaemia associated with developmental delay in very preterm infants. Arch Dis Child 1992; 67:944-947

30.   Den Ouden AL, Kok JH, Verkerk PH, Brand R, Verloove-Vanhorick SP. The relation between neonatal thyroxine levels and neurodevelopmental outcome at age 5 and 9 years in a national cohort of very preterm and/or very low birth weight infants. Pediatr Res 1996; 39:142-145

31.   Lapeyre D, Klosowski S, Liska A, Zaoui C, Gremillet C, Truffert P. [Very preterm infant (< 32 weeks) vs very low birth weight newborns (1500 grammes): comparison of two cohorts]. Arch Pediatr 2004; 11:412-416

Part 1

Early-life glucocorticoid regulation

2

Interpretation of glucocorticoids in neonatal hair: a reflection of intra-uterine glucocorticoid regulation?

Jonneke J. Hollanders,
Bibian van der Voorn,
Noera Kieviet,
Koert M. Dolman,
Yolanda B. de Rijke,
Erica L.T. van den Akker,
Joost Rotteveel,
Adriaan Honig,
Martijn J.J. Finken

ABSTRACT

Background

Glucocorticoids (GCs) measured in neonatal hair might reflect intrauterine as well as postpartum GC regulation. We aimed to identify factors associated with neonatal hair GC levels in early life, and their correlation with maternal hair GCs.

Methods

In a single-center observational study, mother-infant pairs (n=108) admitted for >72 hours at the maternity ward of a general hospital were included. At birth and an outpatient visit (OPV, n=72, 44±11 days postpartum), maternal and neonatal hair was analyzed for cortisol and cortisone levels by LC-MS/MS. Data were analyzed regarding: 1) neonatal GC levels postpartum and at the OPV, 2) associations of neonatal GC levels with maternal GC levels, as well as 3) with other perinatal factors.

Results

1) Neonatal GC levels were >5 times higher than maternal levels, with a decrease of ±50% between birth and the OPV for cortisol. 2) Maternal and neonatal cortisol, but not cortisone, levels were correlated both postpartum and at the OPV. 3) Gestational age was associated with neonatal GCs postpartum (log-transformed β [95%CI]: cortisol 0.07 [0.04-0.10]; cortisone 0.04 [0.01-0.06]) and at the OPV (cortisol 0.08 [0.04-0.12]; cortisone 0.00 [-0.04-0.04]), while weaker associations were found between neonatal GCs and other perinatal and maternal factors.

Conclusions

Neonatal hair GCs mainly reflect the third trimester increase in cortisol, which might be caused by the positive feedback loop, a placenta-driven phenomenon, represented by the positive association with GA. Between birth and 1.5 months postpartum, neonatal hair cortisol concentrations decrease sharply, but still appear to reflect both the intra- and extrauterine period.

INTRODUCTION

Prenatal exposure to excessive glucocorticoids (GCs) has been associated with an increased risk of cardiovascular diseases and depressive disorders.[1,2] This might be due to permanent alterations in the settings of the fetal hypothalamic-pituitary-adrenal (HPA) axis, which are protective in the short term, but might pose a risk in the long term.[3]

The development of the fetal HPA-axis is, among other factors, influenced by the placental transfer of maternal glucocorticoids (GCs) throughout pregnancy.[4] During early gestation, maternal GCs are the main supply. By the second half of gestation, the fetal adrenal starts producing its own steroids, predominantly sex steroids (which serve as a substrate for the placental production of estriol) and precursor GCs, since the adrenocortical enzymes are not fully matured yet.[5] Subsequently, during the last 6-8 weeks of pregnancy, the more matured fetal adrenal produces increasing amounts of cortisol and cortisone under the control of corticotropic-releasing hormone (CRH) production in the placenta, which – in contrast to the negative feedback loop between cortisol and CRH under non-pregnant conditions – establishes a positive feedback loop.[6] This increase in cortisol concentration promotes maturation of the fetal lungs as well as of other organs.[7]

Knowledge on the fetal HPA-axis development is mainly based on animal studies,[5,8] as it is difficult to measure fetal HPA-axis activity in humans. Up till now, amniotic fluid GC levels and umbilical cord GC levels have been used to assess intra-uterine GC regulation. Cortisol in amniotic fluid has previously been correlated to maternal cortisol levels [9] and onset of labor.[10] However, the source of amniotic fluid cortisol remains uncertain, although findings point toward fetal production.[11,12] In addition, sampling of amniotic fluid is a stressful occasion and only provides cross-sectional information. Alternatively, umbilical cord blood can be drawn non-invasively, but GC levels are influenced by delivery [13] and might not reflect normal intrauterine HPA-axis activity. GCs measured in scalp hair might offer a solution, as it is used as a measure for HPA-axis activity over time without the disturbing influence of the circadian rhythm. The hair GC concentrations reflect the exposure in the time frame during which the hair grew.[14]

Maternal hair GC levels seem to reflect HPA-axis activity during pregnancy.[15-17] Neonatal hair GC levels have also been associated with pre- and perinatal factors. A recent study by Hoffman et al. (2017) [18] has shown that gestational age as well as birth weight had positive association with cortisol levels in neonatal hair. Neonatal hair GC levels are significantly higher than maternal hair GC levels. This study suggests that features of the fetal adrenal development are represented in neonatal hair GC levels, although these findings are limited due to the fact that this has only been described in one study population, cortisone levels were not taken into account, and the course followed by GC levels in hair postpartum has not been studied.

Therefore, we aimed to describe cortisol and cortisone concentrations in neonatal hair, obtained directly postpartum, and their relation with maternal hair GC levels and pre- and perinatal factors. Lastly, we explored the differences in hair GC concentrations between birth and an outpatient visit (OPV) at approximately 6 weeks postpartum, as well as which factors are of influence on this difference.

METHODS

Population

From February 2012 to August 2013, mother-infant pairs were included in the OLVG West Hospital in Amsterdam, The Netherlands. Subjects were informed of the study before or within 24 hours after delivery. The infant needed to be admitted to the hospital (maternity ward or neonatal care unit) for at least 72 hours for a neonatal or maternal reason, as this was an inclusion criterion for a simultaneous study.[19] Subjects were excluded for the following reasons: 1) insufficient knowledge of the Dutch or English language, 2) mental retardation of one or both parents, 3) multiple pregnancy, 4) use of illicit drugs or regular (>2 IU/week) alcohol use during the last trimester, 5) use of systemic corticosteroids during pregnancy, 6) if participating in this study would interfere with regular care, or 7) use of psychotropic medication.

The study was approved by the medical ethics committees of the OLVG west Hospital and the VU University Medical Center in Amsterdam, the Netherlands. Written informed consent was obtained from all participants.

Determinants

The mother filled in a questionnaire about demographic characteristics.

Information on perinatal characteristics and the reasons for admission to the hospital were obtained from medical records.

Hair cortisol measurements

On the first day postpartum neonatal hair was cut from the posterior vertex of the scalp, as close as possible to the scalp, as this region shows the least variance between different strands.[14] At the outpatient visit (OPV) around 6 weeks postpartum neonatal hair was collected again. The total length of hair directly postpartum was analyzed, with the assumption that it is an indication of GC concentrations during fetal life, while at the OPV only the centimeter of hair closest to the scalp was analyzed, with the assumption that it gives an indication of GC concentrations during the first weeks of life.[20,21]

Maternal hair was also collected on the first day postpartum and at the OPV. Only the centimeter closest to the scalp of maternal hair was analyzed. As, in adults, hair grows

approximately 1cm every month,[17,20,21] the hair measurement postpartum is indicative for the GC levels during the last month of pregnancy.

GC levels (cortisol and cortisone) were measured in hair as previously described.[20] In short, in the presence of deuterium labeled GCs as internal standard, cortisol was extracted using LC-grade methanol at 25°C for 18 hours. These extracts were subsequently centrifuged and cleaned using solid phase extraction. GC concentrations were quantified by liquid chromatography – tandem mass spectrometry LC-MS/MS (Waters XEVO-TQ-S system, Waters Corporation, Milford, MA, USA). GC concentrations were reported as pg per mg hair, and 1.25mg was required for a reliable measurement.

Statistics

Analyses were performed with regard to:

1. Concentrations of GCs in neonatal hair directly postpartum and at the OPV. GC levels were expressed as median (range). Subsequently, GC levels were log-transformed and paired t-tests were performed.

2. The relation between maternal and neonatal (log-transformed) hair GC levels postpartum and at the OPV, was assessed using Pearson correlation coefficients and linear regression.

3. Factors associated with neonatal hair GCs directly postpartum, were assessed using linear regression. Additional analyses were performed to assess the effect of the factors associated with GC levels directly postpartum on the course of GC levels (expressed as delta cortisol and cortisone) and on the GC levels at the OPV, corrected for age at the time of sampling. The following factors, based on literature,[15-18] were taken into consideration:

   a. Perinatal: gestational age, birth weight (in kg and SD-score), sex, mode of delivery, perinatal infection, respiratory distress (meconium-containing amniotic fluid, respiratory insufficiency, respiratory support, PPHN (persistent pulmonary hypertension of the neonate)).

   b. Maternal: age, ethnicity, maternal smoking, parity (primi- vs. multipara), hypertensive disorders (pregnancy-induced hypertension, pre-existent hypertension, pre-eclampsia/HELLP syndrome (Hemolysis, Elevated Liver enzymes and Low Platelet count)).

Results with a P value <0.05 were considered to be statistically significant, although borderline statistically significant results (0.10 > P > 0.05) when found for both cortisol and cortisone were also further explored.

RESULTS

Population

A total of 107 mother-infant pairs donated hair directly postpartum. At the OPV, 72 mother-infants pairs donated hair. The OPV took place at 44±11 days postpartum (range: 22-87 days). Perinatal and demographic characteristics are presented in Table 1.

Table 1: Baseline characteristics of the study population (n=107)

		Mean±SD, median (range) or n (%)
Perinatal	Gestational age	39.5 ± 1.8, 39.5 (33.9-42.1)
	Birth weight	3480 ± 629, 3569 (1806-5290)
	Male	61 (55.5)
	Vaginal delivery	40 (32.5)
	Perinatal infection	34 (30.9)
	Respiratory problems	9 (8.2)
Maternal	Age	33.9 ± 4.8, 34 (21-44)
	Non-Dutch ethnicity	52 (47.3)
	Smoking	2 (1.8)
	Nulliparae	67 (54.5)
	Hypertensive disorders	7 (6.4)

Concentration of GCs in neonatal hair

Results are displayed in Table 2 and Figure 1. Directly postpartum, the median concentration of cortisol was 169 pg/mg (range: 51 – 1294), while the median concentration of cortisone was 85 pg/mg (range: 23 – 597). Maternal GC levels were much lower than neonatal levels, with median concentrations of 5 (range: 0 – 672) and 18 (2 – 87) pg/mg respectively.

Table 2: Concentrations of neonatal and maternal hair glucocorticoid concentrations postpartum and at the outpatient visit

		Postpartum (median,range)	Outpatient visit (median, range)	P-value*
Infant	Cortisol	169, 51 – 1294	71, 2 – 479	<0.001
	Cortisone	85, 23 – 597	91, 30-346	0.99
Maternal	Cortisol	5, 0 – 672	4, 1 – 79	0.001
	Cortisone	18, 2 – 87	18, 8 – 43	0.75

Values expressed as median, range in pg/mg. * Analyzed with a paired t-test, performed with log-transformed GC concentrations
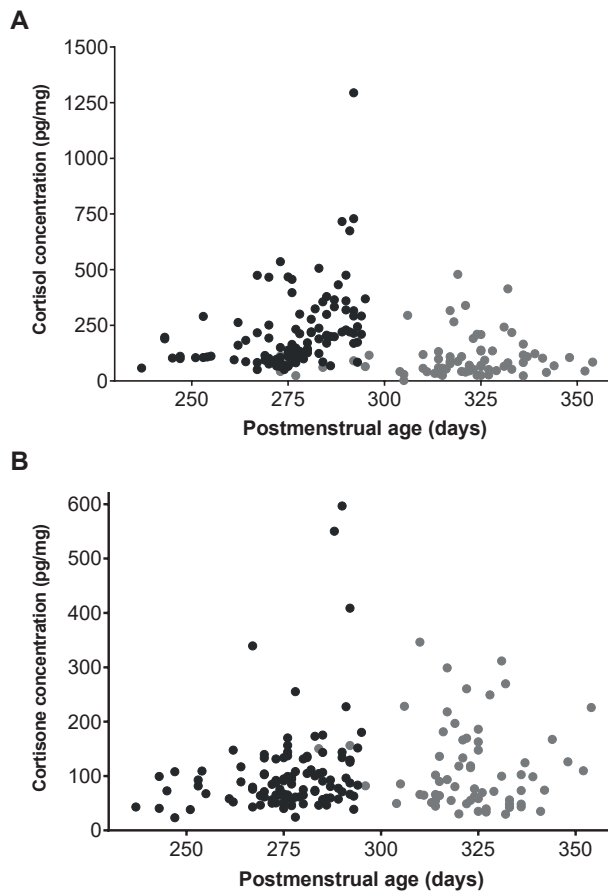
Figure 1: Neonatal hair cortisol (A) and cortisone (B) levels measured directly postpartum (●) and at the OPV (●)

Course of GC levels postpartum

Between birth and the OPV, a steep decrease in cortisol concentrations in infant hair was observed (Table 2 and Figure 1). Maternal hair cortisol levels showed a subtle decrease between birth and the OPV. In contrast, infant and maternal hair cortisone levels remained stable, although a wide range of values was observed. At the OPV, both infant cortisol and cortisone concentrations were still higher than the GC levels in maternal hair. Age of the neonate at the OPV was negatively associated with hair cortisol, but not with cortisone levels (log-transformed β [95%CI]: cortisol -0.01 [-0.02 to -0.001], p=0.03; cortisone: 0.00 [-0.01 to 0.01], p=0.70). Age of the neonate at the OPV was not associated with delta cortisol or cortisone.

Correlations with maternal hair GCs

Directly postpartum, maternal and neonatal hair cortisol were positively associated (n=107, r=0.336, β 0.23 (95%CI: 0.11 – 0.36), p<0.001), while no correlations were found between maternal and infant hair cortisone (p=0.66). At the OPV, the association between maternal and infant hair cortisol was stronger than directly postpartum (n=71, r=0.457, β 0.41 (95%CI: 0.22 – 0.60), p<0.001), and no correlation was found for cortisone (p=0.12).

Table 3: Associations of neonatal hair glucocorticoid concentrations directly postpartum with perinatal and maternal factors

			β (95%CI)	P value
Perinatal factors	Gestational age	Cortisol	0.07 (0.04 – 0.10)	<0.001
		Cortisone	0.04 (0.01 – 0.06)	0.004
	Gestational age (only term pregnancies)	Cortisol	0.11 (0.07 – 0.16)	<0.001
		Cortisone	0.04 (-0.001 – 0.08)	0.06
	Birth weight (kg)	Cortisol	0.09 (-0.003 – 0.17)	0.06
		Cortisone	0.10 (0.03 – 0.17)	0.008
	Birth weight (SD)	Cortisol	0.01 (-0.05 – 0.06)	0.79
		Cortisone	0.03 (-0.02 – 0.07)	0.23
	Male gender	Cortisol	0.10 (-0.02 – 0.21)	0.09
		Cortisone	0.08 (-0.01 – 0.18)	0.07
	Delivery via caesarian section	Cortisol	-0.14 (-0.26 – -0.03)	0.015
		Cortisone	-0.14 (-0.24 – -0.05)	0.003
	Perinatal infection (≥7 days antibiotics)	Cortisol	0.17 (0.06 – 0.29)	0.003
		Cortisone	0.22 (0.13 – 0.31)	<0.001
	Respiratory problems	Cortisol	-0.07 (-0.28 – 0.13)	0.47
		Cortisone	-0.06 (-0.22 – 0.11)	0.52
Maternal factors	Age	Cortisol	-0.01 (-0.02 – 0.01)	0.31
		Cortisone	0.00 (-0.01 – 0.01)	0.82
	Ethnicity	Cortisol	-0.04 (-0.15 – 0.08)	0.55
		Cortisone	-0.09 (-0.16 – 0.01)	0.07
	Maternal smoking	Cortisol	-0.09 (-0.18 – 0.003)	0.06
		Cortisone	-0.09 (-0.43 – 0.26)	0.62
	Parity	Cortisol	-0.26 (-0.36 – -0.16)	<0.001
		Cortisone	-0.07 (-0.17 – 0.02)	0.12
	Hypertensive disorders	Cortisol	-0.02 (-0.25 – 0.21)	0.85
		Cortisone	-0.15 (-0.33 – 0.04)	0.12

Values represent log-transformed β (95% confidence interval) as calculated with linear regression

Factors associated with neonatal hair GCs

The effect of several perinatal and maternal factors on hair GC levels measured directly postpartum were studied (Table 3). Gestational age was strongly associated with both cortisol and cortisone levels, as illustrated in Figure 1, and this association remained significant when only term-born infants (n=98) were studied. Additionally, there was a positive association for both cortisol and cortisone levels with perinatal infection (defined as the need for treatment with antibiotics for ≥7 days).: Birth weight was associated with both cortisol and cortisone when expressed in kg, but this association was lost when birth weight was expressed as SD-score. Moreover, delivery via caesarian section was associated with both lower cortisol and cortisone levels, while multiparity was associated with lower cortisol levels.

Next, the effect of the (borderline) significant factors on the course of GC levels was studied (Table 4). Gestational age was associated with a trend towards a steeper decrease in cortisol and cortisone between birth and the OPV. Additionally, perinatal infection was associated with a steeper decrease in cortisone, while delivery via caesarian section was associated with a smaller decrease in cortisone.

Lastly, the effect of these factors on the GC levels at the OPV was analyzed (Table 4). Gestational age was still positively associated with infant hair cortisol levels, but not with cortisone. Additionally, males had higher cortisol levels in hair at the OPV, but no association was found with cortisone. The other factors were not associated with OPV hair GC levels.

Table 4: Associations of the course of neonatal hair glucocorticoid concentrations with perinatal and maternal factors

			Effect on delta		**Effect on OPV values**	
			β (95%CI)	P value	β (95%CI)	P value
Perinatal factors	Gestational age	Cortisol	-0.02 (-0.05 – 0.00)	0.08	0.08 (0.04 – 0.12)	<0.001
		Cortisone	-0.006 (-0.013 – 0.001)	0.08	0.00 (-0.04 – 0.04)	0.87
	Birth weight (in kg)	Cortisol	-0.05 (-0.12 – 0.02)	0.19	0.13 (0.003 – 0.26)	0.05
		Cortisone	-0.02 (-0.03 – 0.01)	0.15	0.00 (-0.12 – 0.12)	0.96
	Male gender	Cortisol	0.02 (-0.07 – 0.11)	0.68	0.19 (0.02 – 0.36)	0.03
		Cortisone	-0.01 (-0.03 – 0.02)	0.46	0.00 (-0.14 – 0.13)	0.99
	Delivery via caesarian section	Cortisol	0.06 (-0.04 – 0.16)	0.24	-0.06 (-0.25 – 0.13)	0.55
		Cortisone	0.03 (0.01 – 0.05)	0.02	-0.08 (-0.23 – 0.08)	0.33
	Perinatal infection (≥7 days antibiotics)	Cortisol	-0.06 (-0.15 – 0.04)	0.25	-0.04 (-0.22 – 0.14)	0.66
		Cortisone	-0.03 (-0.05 – -0.004)	0.02	0.12 (-0.03 – 0.27)	0.11
Maternal factors	Parity	Cortisol	0.07 (-0.02 – 0.16)	0.11	-0.11 (-0.28 – 0.05)	0.18
		Cortisone	0.01 (-0.01 – 0.03)	0.32	0.01 (-0.13 – 0.14)	0.91

Values represent log-transformed β (95% confidence interval) as calculated with linear regression. All associations were corrected for age at the OPV.

DISCUSSION

In this study, we have described the levels of cortisol and cortisone in neonatal hair, both directly postpartum, as well as at an outpatient visit at 44±11 days postpartum. GC levels in neonatal hair directly postpartum seem to reflect intrauterine GC exposure, they are much higher than maternal levels and appear to be influenced mainly by gestational age, possibly reflecting the normal prenatal increase in endogenous fetal cortisol. After birth, cortisol levels decrease sharply, although at the OPV neonatal levels are still much higher compared to maternal levels. This suggests that at that time point GC levels represent both the intra- and extrauterine period, since GC levels in infants are not markedly different from maternal cortisol levels.[22,23] Additionally, at birth, neonatal hair GC levels are associated with other perinatal factors such as perinatal infection, although to a lesser degree than gestational age.

In our study, we could confirm the association described by Hoffmann et al. [18] between neonatal hair cortisol levels and both gestational age and birth weight directly postpartum. However, we did not find an association with birth weight SDS. Since birth weight and gestational age are correlated, the association with birth weight probably reflects the effect of gestational age rather than of intrauterine growth. The association with gestational age might be indicative of several mechanisms. First, adrenal maturation occurs throughout pregnancy, resulting in increased cortisol production by the fetal adrenal.[5] A higher concentration of GCs in hair might therefore reflect a longer exposure to the maturing HPA-axis. However, since the association between gestational age and neonatal hair GCs is also still present in term neonates, another mechanism appears to be present as well. Induction of labor has been suggested to be partly due to an increase in cortisol, which occurs in all species studied to date and which promotes fetal organ maturation.[7,24] This increase in cortisol is thought to be due to a positive feedback loop established between placenta-derived CRH and cortisol originating from the fetal adrenals,[6,25] which can only be broken by the severance of the umbilical cord. Fetal distress may accelerate this feedback loop,[8] which might explain the increased hair GC levels in neonates who are treated for a perinatal infection.

It is as of yet unknown whether neonatal hair GC levels fully result from fetal cortisol production or whether the transplacental supply of cortisol might also contribute to neonatal GC levels in hair. Previous research has suggested that cortisol is transferred via the placenta to the fetus, although most cortisol is inactivated to cortisone by placental 11B-hydroxysteroid dehydrogenase type 2 (11BHSD2).[8] However, as maternal serum cortisol levels are 10 times higher than fetal serum levels, even small amounts of cortisol could account for about 40% of the variance in fetal concentrations.[4] In our study, we found a positive correlation between maternal and neonatal hair cortisol levels, but not with cortisone. The positive correlation between neonatal and maternal hair cortisol lev-

els might therefore be a reflection of placental transfer. However, this does not explain why the neonatal hair GC levels were much higher compared to maternal levels. We speculate that this may be due to differences in hair growth and structure between the fetus and its mother.

While it is feasible that cortisol in neonatal hair is derived from hair follicles, where it is incorporated after diffusion from blood,[14] cortisol in amniotic fluid might contribute to the GC concentrations measured in hair. Moreover, although hair growth in utero is roughly known, the specifics are still unclear. The first stage of hair growth starts during the 15th week of gestation, and by week 18 to 20 the entire scalp is covered with hair in the primary, anagen stage. Next, between week 24 and 28, the anagen hair converts to telogen hair via a catagen phase.[26] Hair growth, as well as the conversion to more mature hair, is region-specific, and dependent on several biochemical and individual variations.[26-28] Whether hair in the anagen phase already contains GCs, or whether the accumulation of GCs occurs at a later phase, is unknown. Therefore, although it is thought that neonatal hair reflects at least the third trimester of pregnancy,[26] the true time window which is represented by GCs measured in hair is not known. Since perinatal infection and mode of delivery also appear to influence hair GC levels, it is likely that the last stages of pregnancy have a significant contribution to GC levels measured in hair. Future studies should include measurements of growth velocity of neonatal hair.

Our study showed significantly increased GC levels in neonatal hair compared to maternal hair at the OPV (44±11 days postpartum), although a decrease between birth and the OPV was observed in cortisol levels. This suggests that GC levels at the OPV represent a combination of intra- and extrauterine influences, supported by our finding that GC levels at the OPV are still associated with several perinatal factors. However, due to the biochemical and individual variations in hair growth,[26-28] and since hair was only measured twice in this study, the contribution of intrauterine and extrauterine influences on hair GC levels at the OPV is unknown. We recommend to assess at which point in time intra-uterine factors are no longer related to hair GC levels, since this might provide a clear view of early life influences on HPA-axis development. Since hair GC levels appear to be moderately stable in the second half of the first year of life,[29] intrauterine influences on hair GC levels most likely disappear within the first 6 months.

Our study has several strengths and limitations. First, GC analyses were performed using LC-MS/MS which has high sensitivity. Hoffman et al. [18] measured hair cortisol levels with an immunoassay, which might explain the fact that they did not find an association between maternal and neonatal hair cortisol levels, and reported maternal and neonatal cortisol levels much higher compared to our results, since immunoassay are more sensitive to cross-reactivity than LC-MS/MS.[30,31] Cross-reactivity is particularly important to take into account when researching newborns, as they have high concentrations of (precursors of) sex steroids and GCs,[5] which are partly of maternal origin. Additionally,
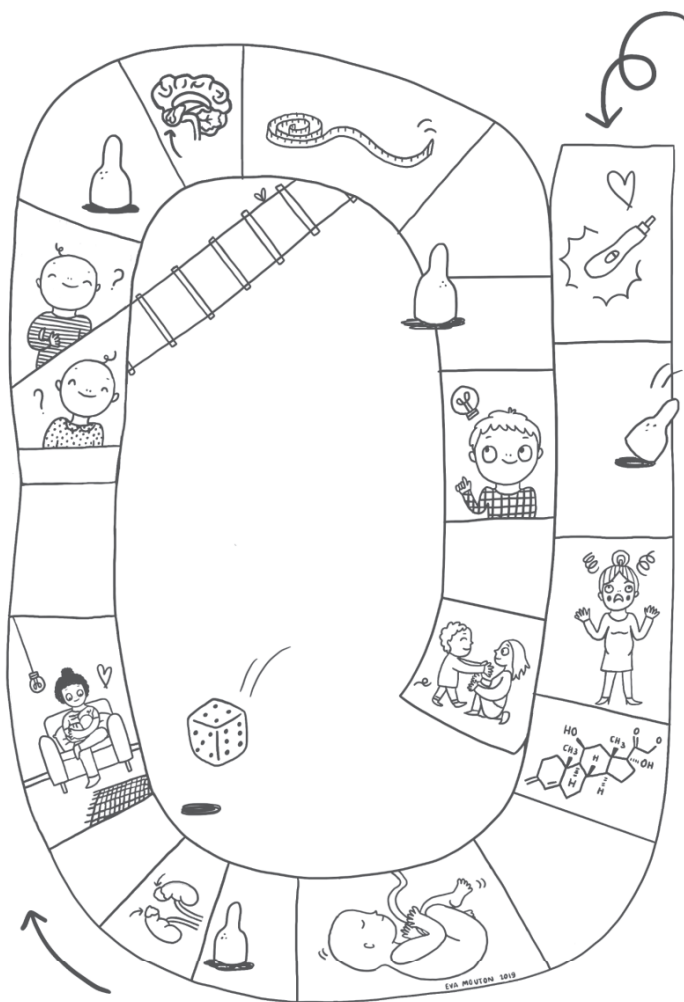
we measured cortisone as well as cortisol, which is valuable knowledge due to the conversion of cortisol to cortisone by placental 11BHSD2.[8] Our database also allowed us to analyze a wide range of pre- and perinatal factors. One of the limitations of our study is that the participants might not represent a normal population, since the participants in our study had to be hospitalized >72 hours. Additionally, there might be selection bias at the OPV measurements due to losses to follow-up, although the number of participants was relatively high (67%). Lastly, there was discrepancy between mother and child in the time frame that the hair measurements represented. In mothers, only the last centimeter of hair was analyzed. As adult hair grows with approximately 1cm per month,[20,21] these analyses are representative of only the last month of pregnancy. Neonatal hair was analyzed in its entirety, and is therefore representative of the intrauterine period during which cortisol can be incorporated in hair. Correlations between GCs in maternal and neonatal hair should therefore be interpreted in this context.

In conclusion, our findings suggest that infant hair GCs reflect the third trimester increase in cortisol, which might be caused by the positive feedback loop, a placenta-driven phenomenon, represented by a positive association with GA. Between birth and 1.5 months postpartum, cortisol concentrations decrease sharply. At this time point, GC levels appear to reflect both the intra- and extrauterine period, since neonatal levels are significantly higher than maternal GC levels. Perinatal complications and maternal HPA-axis activity had minor influences on infant hair GCs.

REFERENCES

1. Cottrell EC, Seckl JR. Prenatal stress, glucocorticoids and the programming of adult disease. Front Behav Neurosci 2009; 3:19

2. Xiong F, Zhang L. Role of the hypothalamic-pituitary-adrenal axis in developmental programming of health and disease. Front Neuroendocrinol 2013; 34:27-46

3. Hanson MA, Gluckman PD. Early developmental conditioning of later health and disease: physiology or pathophysiology? Physiol Rev 2014; 94:1027-1076

4. Gitau R, Cameron A, Fisk NM, Glover V. Fetal exposure to maternal cortisol. Lancet 1998; 352:707-708

5. Brosnan PG. The hypothalamic pituitary axis in the fetus and newborn. Semin Perinatol 2001; 25:371-384

6. McLean M, Smith R. Corticotrophin-releasing hormone and human parturition. Reproduction 2001; 121:493-501

7. Fencl MD, Stillman RJ, Cohen J, Tulchinsky D. Direct evidence of sudden rise in fetal corticoids late in human gestation. Nature 1980; 287:225-226

8. Watterberg KL. Adrenocortical function and dysfunction in the fetus and neonate. Semin Neonatol 2004; 9:13-21

9. Sarkar P, Bergman K, Fisk NM, O'Connor TG, Glover V. Ontogeny of foetal exposure to maternal cortisol using midtrimester amniotic fluid as a biomarker. Clin Endocrinol (Oxf) 2007; 66:636-640

10. Ohana E, Mazor M, Chaim W, Levy J, Sharoni Y, Leiberman JR, Glezerman M. Maternal plasma and amniotic fluid cortisol and progesterone concentrations between women with and without term labor. A comparison. J Reprod Med 1996; 41:80-86

11. Fencl MM, Koos B, Tulchinsky D. Origin of corticosteroids in amniotic fluid. J Clin Endocrinol Metab 1980; 50:431-436

12. Partsch CJ, Sippell WG, MacKenzie IZ, Aynsley-Green A. The steroid hormonal milieu of the undisturbed human fetus and mother at 16-20 weeks gestation. J Clin Endocrinol Metab 1991; 73:969-974

13. Leong MK, Murphy BE. Cortisol levels in maternal venous and umbilical cord arterial and venous serum at vaginal delivery. Am J Obstet Gynecol 1976; 124:471-473

14. Staufenbiel SM, Penninx BW, Spijker AT, Elzinga BM, van Rossum EF. Hair cortisol, stress exposure, and mental health in humans: a systematic review. Psychoneuroendocrinology 2013; 38:1220-1235

15. Braig S, Grabher F, Ntomchukwu C, Reister F, Stalder T, Kirschbaum C, Genuneit J, Rothenbacher D. Determinants of maternal hair cortisol concentrations at delivery reflecting the last trimester of pregnancy. Psychoneuroendocrinology 2015; 52:289-296

16. Dettmer AM, Rosenberg KL, Suomi SJ, Meyer JS, Novak MA. Associations between Parity, Hair Hormone Profiles during Pregnancy and Lactation, and Infant Development in Rhesus Monkeys (Macaca mulatta). PLoS One 2015; 10:e0131692

17. D'Anna-Hernandez KL, Ross RG, Natvig CL, Laudenslager ML. Hair cortisol levels as a retrospective marker of hypothalamic-pituitary axis activity throughout pregnancy: comparison to salivary cortisol. Physiol Behav 2011; 104:348-353

18. Hoffman MC, D'Anna-Hernandez K, Benitez P, Ross RG, Laudenslager ML. Cortisol during human fetal life: Characterization of a method for processing small quantities of newborn hair from 26 to 42 weeks gestation. Dev Psychobiol 2017; 59:123-127

19. Kieviet N, van Keulen V, van de Ven PM, Dolman KM, Deckers M, Honig A. Serotonin and poor neonatal adaptation after antidepressant exposure in utero. Acta Neuropsychiatr 2017; 29:43-53

20. Noppe G, de Rijke YB, Dorst K, van den Akker EL, van Rossum EF. LC-MS/MS-based method for long-term steroid profiling in human scalp hair. Clin Endocrinol (Oxf) 2015; 83:162-166

21. Wennig R. Potential problems with the interpretation of hair analysis results. Forensic Sci Int 2000; 107:5-12

22. Garcia-Blanco A, Vento M, Diago V, Chafer-Pericas C. Reference ranges for cortisol and alpha-amylase in mother and newborn saliva samples at different perinatal and postnatal periods. J Chromatogr B Analyt Technol Biomed Life Sci 2016; 1022:249-255

23. Jonetz-Mentzel L, Wiedemann G. Establishment of reference ranges for cortisol in neonates, infants, children and adolescents. Eur J Clin Chem Clin Biochem 1993; 31:525-529

24. Fowden AL, Li J, Forhead AJ. Glucocorticoids and the preparation for life after birth: are there long-term consequences of the life insurance? Proc Nutr Soc 1998; 57:113-122

25. Challis JR, Hooper S. Birth: outcome of a positive cascade. Baillieres Clin Endocrinol Metab 1989; 3:781-793

26. Gareri J, Koren G. Prenatal hair development: implications for drug exposure determination. Forensic Sci Int 2010; 196:27-31

27. Berger HM, King J, Doughty S, Wharton BA. Nutrition, sex, gestational age, and hair growth in babies. Arch Dis Child 1978; 53:290-294

28. Furdon SA, Clark DA. Scalp hair characteristics in the newborn infant. Adv Neonatal Care 2003; 3:286-296

29. Liu CH, Snidman N, Leonard A, Meyer J, Tronick E. Intra-individual stability and developmental change in hair cortisol among postpartum mothers and infants: Implications for understanding chronic stress. Dev Psychobiol 2016; 58:509-518

30. Ackermans MT, Endert E. LC-MS/MS in endocrinology: what is the profit of the last 5 years? Bioanalysis 2014; 6:43-57

31. Shackleton C. Clinical steroid mass spectrometry: a 45-year history culminating in HPLC-MS/MS becoming an essential tool for patient diagnosis. J Steroid Biochem Mol Biol 2010; 121:481-490

3

Maternal stress during pregnancy is associated with decreased cortisol and cortisone levels in neonatal hair

Bibian van der Voorn,
Jonneke J. Hollanders,
Noera Kieviet,
Koert M. Dolman,
Yolanda B. de Rijke,
Elisabeth F.C. van Rossum,
Joost Rotteveel,
Adriaan Honig,
Martijn J.J. Finken

ABSTRACT

Background

Hair glucocorticoids (GCs) offer a retrospective view on chronic GC exposure. We assessed whether maternal pre- and postnatal stress was associated with neonatal and maternal hair GCs postpartum (pp).

Methods

On the first day pp 172 mother-infant pairs donated hair, of whom 67 had consulted a center of expertise for psychiatric disorders during pregnancy. Maternal stress was scored on the Hospital Anxiety and Depression Scale during the first/second (n = 46), third trimester (n = 57), and pp (n = 172). Hair cortisol and cortisone levels were determined by liquid chromatography-tandem mass spectrometry, and associations with maternal hospital anxiety subscale (HAS) and hospital depression subscale (HDS) scores, and antidepressant use were analyzed with linear regression.

Results

Neonatal hair GCs were negatively associated with elevated HAS-scores during the first/second trimester, log 10 (β [95% CI]) cortisol −0.19 (−0.39 to 0.02) p = 0.07, cortisone −0.10 (−0.25 to 0.05) p = 0.17; third trimester, cortisol −0.17 (−0.33 to 0.00) p = 0.05, cortisone −0.17 (−0.28 to −0.05) p = 0.01; and pp, cortisol −0.14 (−0.25 to −0.02) p = 0.02, cortisone −0.07 (−0.16 to 0.02) p = 0.10. A similar pattern was observed for elevated HDS-scores. Maternal hair GCs were positively associated with elevated HAS-scores pp (cortisol 0.17 [0.01 to 0.32] p = 0.04, cortisone 0.18 [0.06 to 0.31] p = 0.01), but not prenatally or with elevated HDS-scores. Antidepressant use was associated with elevated maternal hair GCs ($p \leq 0.05$), but not with neonatal hair GCs.

Conclusion

Exposure to excessive pre- and perinatal maternal stress was associated with a decrease in neonatal hair GCs, while elevated stress-scores around birth were associated with increased maternal hair GCs and elevated stress-scores earlier in gestation were not associated with maternal hair GCs pp. Further studies are needed to test associations with infant neurodevelopment.

INTRODUCTION

Anxiety or depressive disorders are associated with alterations in hypothalamic-pituitary-adrenal (HPA) axis activity and reactivity, although the evidence is not unequivocal.[1-4] Anxiety and depressive disorders are common in pregnancy, with numbers ranging from 1 in 10 to 1 in 5 pregnant women.[5-7] Although many observational studies described associations between prenatal exposure to maternal stress and neurodevelopmental problems,[8,9] caution must be exercised in the interpretation of some of these findings due to the use of subjective measures of stress, while quantitative indices of HPA axis activity are lacking.[10]

As part of the physiological changes during pregnancy, both maternal and fetal glucocorticoids (GCs) exert a positive feedback effect on the placenta by stimulating the synthesis of placental corticotropin-releasing hormone (CRH). Due to this physiological feed-forward response, maternal cortisol increases during gestation.[11] At the same time, increasing estrogen levels augment the synthesis of corticosteroid-binding globulin,[12] resulting in only a modest increase in free cortisol.[13] Moreover, placental 11β-hydroxysteroid dehydrogenase type 2 (11β-HSD2) converts maternal cortisol to inert cortisone. Accordingly, the fetus is partially protected from overexposure to maternal cortisol.[14] Lower placental 11β-HSD2 activity and, consequently, increased delivery of maternal cortisol to the fetus has been associated with decreased fetal growth.[15] Longer-term consequences of increased fetal exposure to maternal cortisol may include increased HPA axis reactivity and susceptibility to neurodevelopmental problems.[8]

Hair cortisol and cortisone levels represent long-term GC exposure in adults and children above the age of 4 years.[16-18] Accordingly, GC levels in newborn hair might offer a retrospective view on the GC regulation during the last part of pregnancy.[19,20] Kapoor et al. studied hair GC levels in the offspring of rhesus monkeys that were randomized to receive exposure to a startle paradigm for 10 minutes per day, 5 days a week for one-fifth the duration of pregnancy, and found decreased hair cortisol in the exposed offspring, but no difference in hair cortisone.[21] In humans, cortisol in neonatal hair obtained directly postpartum (pp) was higher with advancing gestational age and birth weight.[22] Unfortunately, in this study the impact of maternal stress was not studied. A recent study in humans,[20] testing neonatal hair GC levels in association with maternal hair GC levels and perceived stress, found similar results as Kapoor et al [21] did. However, hair cortisone levels were not taken into account, and hair cortisol levels were measured with an immunoassay technique. Moreover, this study had only included physically and mentally healthy mothers, presumably with low amounts of prenatal stress.

Therefore, in the present study we assessed whether pre- and perinatal exposure to maternal stress is associated with neonatal hair cortisol and cortisone levels directly pp. In addition, we tested associations between maternal stress and GCs in maternal hair

obtained at the same time. To this end, we used data from a cohort in which women with severe distress during pregnancy were overrepresented.

METHODS

Study design and participants

The present study was part of a prospective cohort study that aimed to explore biomarkers, including neonatal hair GCs and 5-hydroxyindoleacetic acid level in urine, for poor neonatal adaptation after prenatal exposure to selective antidepressants (SADs) and maternal stress.[23,24]

A total of 172 mother-infant pairs were recruited at the maternity department, as well as at the psychiatric-obstetric-pediatric (POP) clinic of the OLVG-West Hospital, Amsterdam, The Netherlands, which offers consultation to women with psychiatric disorders before, during, and after pregnancy on an outpatient basis. The reasons for which pregnant women sought advice at the POP clinic were (1) a history of psychiatric disease, and/or (2) symptoms of distress, and/or (3) current or past use of antidepressants. Approximately one-third (n =65) of our sample consisted of women who visited the POP clinic. The other part of the sample (n =107) consisted of mothers admitted pp to the maternity ward for medical reasons in themselves and/or in their infants. Therefore, in this cohort women who experienced severe distress during pregnancy were overrepresented. Among participants, 66 (38%) used SADs, including selective serotonin reuptake inhibitors (SSRIs), serotonin-norepinephrine reuptake inhibitors (SNRIs), noradrenergic or specific serotonin antidepressants (NaSSAs), or a combination of these. Sixty-four of these women sought advise at the POP clinic.

Inclusion and exclusion criteria were similar for both groups. Inclusion criteria were: an expected hospital stay of ≥72 hours after delivery, and willingness to donate hair from themselves and their infants, and to complete the Hospital Anxiety and Depression Scale (HADS) questionnaire directly pp. Exclusion criteria were: use of psychotropic medication other than SADs, use of systemic corticosteroids, non-pharmacologic drugs, or alcohol, smoking during the third trimester of pregnancy, insufficient knowledge of the Dutch or English language, mental impairment of one or both parents, and multiple pregnancies. Parents were informed and written informed consent was obtained within 24 hours after delivery. The study was approved by the Medical Ethics Committees of the OLVG-West Hospital and the VU University Medical Center.

Assessment of maternal stress

As part of standard care at the POP clinic, the HADS [25] was administered as an index of stress experienced in the previous week. The HADS contains 14 items, namely 7 for

anxiety and 7 for depression. From these items, Hospital Anxiety Subscale (HAS) and Hospital Depression Subscale (HDS) scores are derived. A score ≥8 (out of 21 points for each subscale separately) is considered the cut-off for relevant stress.[25,26] The retest reliability of the HADS was found to correlate well with the previous 6 weeks.[26] Furthermore, among pregnant women, the anxiety subscales of the HADS and the Edinburgh Postnatal Depression Scale showed strong correlation.[27]

During the first or second trimester, and/or the third trimester, the HADS was administered only in the women who visited the POP clinic. Directly pp, that is, within 12 to 36 hours, the HADS was administered in all women. At the maternity ward, mothers were asked whether they had used SADs at least during the last two weeks of pregnancy.

Hair glucocorticoid levels

Mother-infant pairs donated hair on the first day pp. A lock of hair was cut from the posterior vertex as close as possible to the scalp. A minimum of 1.25 mg hair is needed for a reliable measurement. Fetal hair growth velocity and the timing of transition from lanugo via vellus into terminal hair strands varies significantly between infants.[28] Therefore, the total length of neonatal hair was analyzed. In mothers, the centimeter of hair closest to the scalp was analyzed, representing the mean levels of cortisol and cortisone during the last month of gestation, as adult hair grows approximately 1 (range: 0.6 – 1.4) cm per month.[29] There is evidence of 11β-HSD2 expression in human eccrine sweat glands and vascular endothelium,[30] raising the possibility of local conversion of blood-borne cortisol to cortisone within skin and/or hair follicles. Therefore, it is unknown which analyte is the best representative of serum cortisol. The sum of hair GCs might be indicative of chronic circulating cortisol and was therefore calculated.

Hair cortisol and cortisone levels were measured as described previously by Noppe et al.[17] In short, hair was washed with isopropanol, and hair GCs were extracted using methanol and solid-phase extraction. Subsequently, cortisol and cortisone concentrations were quantified by liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Waters XEVO-TQ-S system, Waters Corporation, Milford, MA, USA) with positive electrospray ionization, and reported in pg/mg hair. The Lower Limit of Quantitation (LLoQ) of our assay is dependent on the amount of hair extracted. An intra-assay CV of 8.9% was measured at a hair cortisol concentration of 1.8 pg/mg. The intra-assay CV for hair cortisone was 4.4% at a level of 12.5 pg/mg.

Data analysis

Maternal and neonatal characteristics were compared between the pairs whose mothers visited the POP clinic and the pairs admitted for medical reasons, using independent t-tests, Chi Square, or Fisher exact tests (Table 1).

Table 1: Characteristics of mother-infant pairs

		Total group (n =172)		POP mothers (n=67)		Other mothers (n=105)	
Neonatal	Males	92	(53%)	34	(51%)	58	(55%)
	Gestational age wks	39.4	± 1.7	39.2	± 1.6	39.5	± 1.7
	Birth weight g	3,445.5	± 582.6	3358.5	± 511.4	3500.2	± 619.3
	percentile	53.6	± 26.4	49.3	± 24.0	56.3	± 27.6
	Hair cortisol pp pg/mg hair	162.8	(102.8 – 232.2)	155.3	(111.4 – 202.9)	171.3	(96.8 – 291.0)
	Hair cortisone pp pg/mg hair	83.2	(63.1 – 109.8)	79.3	(63.9 – 105.1)	87.2	(61.4 – 128.9)
Maternal	Primiparous	85	(49%)	29	(43%)	56	(53%)
	Age yr	33.8	± 4.7	33.7	± 4.5	34.0	± 4.8
	Ethnicity Dutch	97	(56%)	41	(61%)	56	(53%)
	Caucasian, non-Dutch	16	(9%)	4	(6%)	12	(11%)
	non- Caucasian	59	(35%)	22	(33%)	37	(35%)
	Antidepressants SSRI	45	(26%)	44	(66%)	1[†]	(1%)
	SNRI	7	(4%)	7	(10%)	-[†]	
	NaSSA	9	(5%)	8	(12%)	1[†]	(1%)
	Combination*	5	(3%)	5	(8%)	-[†]	
	HADS score pp HAS score ≥8	30	(17%)	19	(28%)	11[†]	(11%)
	HDS score ≥8	19	(11%)	11	(16%)	8	(8%)
	Hair cortisol pp pg/mg hair	5.4	(3.6 – 10.6)	6.9	(4.4 – 12.0)	4.8[†]	(3.4 – 9.9)
	Hair cortisone pp pg/mg hair	19.5	(14.5 – 31.2)	21.7	(15.4 – 46.9)	18.2[†]	(12.9 – 26.8)

Data are presented as mean ± SD, median (interquartile range), or n (%). Abbreviations: pp = postpartum; HADS = Hospital Anxiety and Depression Scale; HAS = Hospital Anxiety Scale; HDS = Hospital Depression Scale; SSRI = selective serotonin reuptake inhibitors; SNRI = serotonin-norepinephrine reuptake inhibitors, NaSSA = noradrenergic or specific serotonin antidepressants.
* These women were treated with a combination of SSRI with NaSSA (n =4), or NaSSA with SNRI (n =1) [†] Different from POP mothers, P < 0.05

Hair cortisol and cortisone levels were skewed to the right and therefore logarithmically transformed prior to analysis. Linear regression was used to assess associations between HADS scores and hair GC levels. Associations with maternal stress were assessed with hair GC level as dependent factor, and HAS or HDS score as continuous or dichotomous (with a score of ≥8 points as cut-off for elevated stress) independent factor.[25] Among infants whose mothers visited the POP clinic, the relative contributions of pre- and perinatal stress exposure were tested by using combinations of (1) low prenatal and low perinatal (reference), (2) low prenatal and high perinatal, (3) high prenatal and low perinatal, and (4) high prenatal and high perinatal levels of stress exposure. Low prenatal stress exposure was defined as low HAS and HDS scores in both the first/second and the third trimester, while high prenatal stress exposure was defined as a score ≥8 on one or both subscales in the first/second and/or the third trimester. Likewise, low peri-

natal stress exposure was defined as low HAS and HDS scores pp, while high perinatal stress exposure was defined as a score ≥8 on one or both subscales pp. Associations with maternal SAD use were analyzed with hair GC level as dependent factor, and SAD use as dichotomous independent factor.

Confounders were selected a priori, based on the literature.[21-23,31] Sex, birth weight percentile, gestational age, and primiparity were added to the multivariable model, one by one. Subsequently, based on statistical impact (i.e., a >10% change in beta) the final model was created. When a confounder was found to have a statistical impact on more than 50% of the associations being analyzed, we also explored the univariate association with the outcome. In addition, similar to Kapoor et al.,[21] interaction between perinatal stress (HADS scores pp) and sex on neonatal hair GC levels was tested.

RESULTS

The characteristics of participants are shown in Table 1. A total of 67 women visited the POP clinic, of whom 98% reported SAD use, 28% had an elevated HAS score, and 16% had an elevated HDS score. For mothers admitted pp for medical reasons in themselves and/or in their infants (n =105), these numbers (2%, 11%, and 8%, respectively) were similar to previously reported prevalence rates in the normal population.[5-7,32] Sex distribution, gestational age, birth weight, parity, maternal age, and ethnicity did not differ between the groups. Neonatal hair cortisone levels were significantly lower in female neonates (median [IQR]: 75.1 [59.8 – 99.7] pg/mg for females and 92.1 [65.4 – 129.2] pg/mg for males, p=0.049). Neonatal hair cortisol levels did not differ significantly between the sexes.

The characteristics of mother-infant pairs by time point are shown in Supplementary Table 1. The great majority of women who visited the POP clinic during the first or second, and/or the third trimester, used SADs: 44 out of 46 (96%) and 54 out of 57 (95%) respectively. Those who visited the POP clinic during the first or second trimester, and/or third trimester, more often used SADs during the entire pregnancy: 41 out of 46 (89%) and 43 out of 57 (75%) respectively.

The association between maternal stress and neonatal hair GCs

HADS scores during pregnancy were only known for the mothers who visited the POP clinic (n=65), namely, 46 during the first or second trimester, and 57 during the third trimester. As part of the routine follow-up, 38 of them were seen on both occasions – 8 only during the first or second trimester, and 19 only during the third trimester (Supplementary Figure 1).

Table 2: Maternal stress in association with neonatal hair cortisol and cortisone levels

		Postpartum (n =166)		3rd trimester (n =57)		1st – 2nd trimester (n =45)	
		crude	adjusted	crude	adjusted	crude	adjusted
Cortisol	**Anxiety**	HAS score					
		-0.02 (-0.03; 0) *	-0.01 (-0.02; 0)	-0.02 (-0.03; 0) *	-0.01 (-0.03; 0)	-0.02 (-0.04; 0) *	-0.03 (-0.04; -0.01) *
		Elevated HAS					
		-0.14 (-0.26; -0.03) *	-0.09 (-0.20; 0.01)	-0.17 (-0.33; 0) *	-0.10 (-0.26; 0.07)	-0.19 (-0.39; 0.02)	-0.22 (-0.40; -0.03) *
	Depression	HDS score					
		-0.01 (-0.03; 0)	-0.01 (-0.02; 0)	-0.01 (-0.03; 0)	-0.01 (-0.03; 0.01)	-0.03 (-0.05; -0.01) *	-0.03 (-0.05; -0.01) *
		Elevated HDS					
		-0.12 (-0.26; 0.02)	-0.10 (-0.23; 0.02)	-0.19 (-0.38; -0.01) *	-0.12 (-0.31; 0.07)	-0.35 (-0.57; -0.12) *	-0.32 (-0.53; -0.11) *
Cortisone	**Anxiety**	HAS score					
		-0.01 (-0.02; 0) *	-0.01 (-0.02; 0)	-0.01 (-0.03; 0) *	-0.01 (-0.02; 0.01)	-0.01 (-0.03; 0.01)	-0.01 (-0.03; 0)
		Elevated HAS					
		-0.09 (-0.18; 0.00) *	-0.07 (-0.16; 0.02)	-0.17 (-0.28; -0.05) *	-0.12 (-0.23; 0.00)	-0.10 (-0.25; 0.05)	-0.12 (-0.26; 0.01)
	Depression	HDS score					
		-0.01 (-0.02; 0) *	-0.01 (-0.02; 0) *	-0.01 (-0.02; 0)	-0.01 (-0.02; 0.01)	-0.01 (-0.02; 0.01)	-0.01 (-0.02; 0.01)
		Elevated HDS					
		-0.10 (-0.21; 0.01)	-0.09 (-0.19; 0.02)	-0.17 (-0.31; -0.03) *	-0.11 (-0.25; 0.03)	-0.09 (-0.27; 0.09)	-0.07 (-0.24; 0.10)

Data are presented as Log10-transformed β (95% CI). The adjusted model is corrected for primiparity.

Abbreviations: HAS = Hospital Anxiety Scale; Elevated HAS = score ≥8; HDS = Hospital Depression Scale; Elevated HDS = score ≥8

* p <0.05

We were able to collect enough hair (median [IQR]: 5.1 [3.2 – 8.9] mg, while at least 1.25 mg is required) in 166 newborns (97%). In crude analyses, maternal anxiety and depression experienced during pregnancy or pp were negatively associated with the neonatal hair cortisol level (Table 2). Similar, but weaker, associations were found for the neonatal hair cortisone level. Associations with the sum of neonatal hair GCs resembled those with neonatal hair cortisol alone (data not shown). Correction for sex, gestational age, or birth weight percentile did not change these associations. Correction for primiparity strengthened the associations with first or second trimester stress scores. Associations with maternal stress during the third trimester or pp became non-significant or disappeared when parity was factored in (Table 2). When tested univariately, primiparity was associated with higher neonatal hair cortisol and cortisone levels (log10-transformed β [95% CI]: 0.24 [0.16; 0.32] p <0.001, and 0.10 [0.04; 0.17] p <0.001, respectively). There was no evidence for interaction between dichotomous HADS, HAS, or HDS scores and sex on neonatal hair GC levels.

Among the infants whose mothers visited the POP clinic, persistent exposure to elevated maternal stress was associated with the largest decrease in neonatal hair GC levels (Table 3).

Table 3: Maternal stress over time among the women who visited the POP clinic, in association with neonatal hair cortisol and cortisone levels.

	Exposure category	n	Cortisol		Cortisone	
			Beta	(95% CI)	Beta	(95% CI)
HAS	Low prenatally & low perinatally	36	ref		ref	
	High prenatally & low perinatally	11	-0.09	(-0.30; 0.12)	-0.14	(-0.29; 0.02)
	Low prenatally & high perinatally	6	-0.03	(-0.29; 0.23)	0.01	(-0.18; 0.20)
	High prenatally & high perinatally	12	-0.19	(-0.39; 0.01)	-0.15	(-0.29;- 0.01) *
HDS	Low prenatally & low perinatally	50	ref		ref	
	High prenatally & low perinatally	2	-0.09	(-0.36; 0.19)	-0.18	(-0.38; 0.03)
	Low prenatally & high perinatally	5	-0.13	(-0.55; 0.29)	-0.14	(-0.45; 0.17)
	High prenatally & high perinatally	8	-0.27	(-0.49;- 0.05) *	-0.17	(-0.33; 0.00) *

Data are presented as Log10-transformed β (95% CI). Abbreviations: HAS = Hospital Anxiety Scale; HDS = Hospital Depression Scale score * p <0.05

The association between maternal stress and maternal hair GCs

Maternal anxiety experienced directly pp, but not in gestation, was positively associated with maternal hair cortisol and cortisone levels pp (Table 4). Similar findings were obtained with the sum of maternal hair GCs (data not shown). Correction for sex, gestational age, birth weight percentile, or primiparity did not change these associations.

Table 4: Maternal stress in association with maternal hair cortisol and cortisone levels

			Postpartum (n= 169)		3rd trimester (n= 56)		1st – 2nd trimester (n= 45)	
			crude	adjusted	crude	adjusted	crude	adjusted
Cortisol	**Anxiety**	HAS score	0.01 (-0.01; 0.03)	0.01 (0; 0.03)	-0.01 (-0.04; 0.01)	-0.01 (-0.03; 0.01)	0 (-0.03; 0.03)	0 (-0.03; 0.03)
		Elevated HAS	0.15 (0; 0.30)	0.18 (0.03; 0.33) *	-0.14 (-0.36; 0.08)	-0.08 (-0.30; 0.15)	-0.03 (-0.30; 0.25)	-0.05 (-0.33; 0.22)
	Depression	HDS score	0 (-0.02; 0.02)	0 (-0.02; 0.02)	0 (-0.03; 0.02)	0 (-0.02; 0.02)	0.01 (-0.03; 0.04)	0.01 (-0.02; 0.04)
		Elevated HDS	0.02 (-0.17; 0.20)	0.03 (-0.15; 0.21)	-0.08 (-0.33; 0.17)	-0.01 (-0.26; 0.25)	0.18 (-0.14; 0.50)	0.21 (-0.11; 0.53)
Cortisone	**Anxiety**	HAS score	0.02 (0; 0.03) *	0.02 (0.01; 0.03) *	-0.01 (-0.03; 0.01)	-0.01 (-0.03; 0.02)	0 (-0.02; 0.03)	0 (-0.02; 0.02)
		Elevated HAS	0.17 (0.05; 0.29) *	0.19 (0.07; 0.31) *	-0.15 (-0.35; 0.04)	-0.11 (-0.31; 0.10)	-0.04 (-0.27; 0.18)	-0.06 (-0.29; 0.17)
	Depression	HDS score	0.01 (-0.01; 0.02)	0.01 (-0.01; 0.02)	0 (-0.02; 0.02)	0 (-0.02; 0.02)	0 (-0.03; 0.03)	0 (-0.02; 0.03)
		Elevated HDS	0.10 (-0.05; 0.25)	0.11 (-0.04; 0.25)	-0.01 (-0.23; 0.21)	0.06 (-0.17; 0.29)	0.11 (-0.16; 0.39)	0.13 (-0.15; 0.41)

Data are presented as Log10-transformed β (95% CI). The adjusted model is corrected for primiparity.

Abbreviations: HAS = Hospital Anxiety Scale; Elevated HAS = score ≥8; HDS = Hospital Depression Scale; Elevated HDS = score ≥8

* p<0.05

The association between SAD use and neonatal and maternal hair glucocorticoid levels

Sixty-six women (38%) used SADs. Forty-five women were on SSRIs, 7 on SNRIs, 9 on NaSSAs, and 5 on a combination of these, including SSRIs with NaSSAs (in 4) and NaSSAs with SNRIs (n=1).

The use of SADs was not associated with neonatal hair GC levels, but it was positively associated with maternal hair cortisol and cortisone levels pp (log-10 transformed β [95% CI]: 0.14 [0.02, 0.26] p =0.02, and 0.18 [0.09; 0.27] p <0.001 respectively).

DISCUSSION

In this study, maternal stress during pregnancy and pp was associated with decreased neonatal hair GC levels, with the lowest values seen in the infants of mothers with persistent stress. In addition, maternal distress pp, but not in gestation, was associated with increased maternal hair GC levels pp. SAD use during pregnancy was unlikely to explain these associations.

We recently published data suggesting that neonatal hair GC levels are influenced by the third-trimester increase in HPA axis activity.[19] This phenomenon might offer an explanation for the high neonatal hair GC levels observed in our study, although evidence is lacking concerning which part of intrauterine GC regulation is reflected in neonatal hair. It has been hypothesized that neonatal hair GC levels might reflect amniotic-fluid GC levels.[21] However, this must be balanced against evidence from studies in adults suggesting that hair GCs reflect the body's HPA axis activity, with hair GCs being associated with long-term integrated salivary cortisol.[33] It is conceivable that neonatal hair GCs reflect a combination of maternal and fetal GCs, determined by placental factors like the third-trimester increase in placental CRH and 11β-HSD2 activity and fetal adrenal maturation. The balance between cortisol and cortisone in hair may be different than in blood, as 11β-HSD2 is widely expressed in epithelial tissues.[30]

It is unclear why the HAS score, but not the HDS score pp, was positively associated with maternal hair GCs pp. In general, in patients with psychiatric symptomatology anxiety and depressive symptoms are highly overlapping. Findings from studies investigating associations between HADS scores and indices of HPA axis activity in matrices other than hair, like saliva, are highly contradictory. One study found that the HAS score, but not the HDS score, correlated positively with the cortisol awakening response among patients with coronary artery disease.[34] Another study among patients with coronary artery disease found that the HDS score correlated positively with salivary cortisol in men, but not in women, while in women, but not in men, the HAS score correlated negatively with salivary cortisol.[35] In patients with low back pain, HADS scores were unrelated to

the diurnal cortisol rhythm measured in saliva,[36] while in the normal population, HADS scores were inversely associated with peak salivary cortisol during a psychological stress protocol.[37]

Previous studies have shown that maternal stress is associated with reduced placental 11β-HSD2 activity,[38-41] thereby allowing a larger proportion of maternal cortisol to reach the fetus. In contrast, we found that higher maternal stress scores were associated with lower neonatal hair GCs levels, in spite of evidence of increased maternal HPA axis activity. There are several explanations possible for these discrepant findings. First, maternal HPA axis activity may be lower in those pregnant women with chronic stress, so that, despite reduced placental 11β-HSD2 activity, a smaller amount of cortisol is able to cross the placenta. Some studies have shown blunted HPA axis activity in chronic stress,[4] but it is unclear whether their findings could be extrapolated to pregnancy, when HPA activity is regulated to a large extent by placental CRH. However, the stressed women in our cohort had elevated hair GC levels, representing increased HPA axis activity in the last trimester. Second, the third-trimester surge in the production of GCs by the fetal adrenal probably contributes much more to the fetal GC pool than maternal HPA axis activity, at least in the last part of gestation. Still, the decrease in placental 11β-HSD2 activity as observed in chronic distress could contribute significantly to the fetal GC pool at earlier stages of gestation, when placental 11β-HSD2 activity might be lower.[42] We speculate that increased exposure to maternal cortisol earlier in gestation might lead to a long-lasting suppression of fetal HPA axis activity. Analysis of the entire neonatal hair, which we did in our study, probably reflects the endocrine milieu during the last trimester, as it may take 20 weeks until the scalp is fully covered with anagen phase hair, turning into the catagen phase between week 24 and 28.[43]

In rat pups, GC receptor density in the forebrain was found to develop in a sex-specific manner, with females having more GC receptors than males.[44] It could therefore be hypothesized that prenatal maternal stress might have sex-specific effects on neonatal HPA axis activity. However, we as well as others did not find evidence for sex differences in, or sex-specificity of the effects of prenatal maternal stress on, neonatal hair GC levels.[20-22]

Fetal exposure to maternal stress or to excessive maternal GCs has been associated with neurodevelopmental problems as well as alterations in HPA axis settings.[8,9] The mechanisms behind these associations are not yet fully elucidated. Therefore, long-term follow-up of our cohort is warranted to explore associations with HPA axis development, including aspects like stress reactivity and development of diurnal rhythmicity, in addition to neurodevelopment.[12]

The major strength of this study is the unique sample of women experiencing a wide range of stress levels during pregnancy and pp, including an overrepresentation of severely distressed women. In addition, to the best of our knowledge, this is the first translational study that measured hair GC levels in both infants and their mothers who

experienced a reliably quantified amount of stress during pregnancy and pp.[45] A limitation of our study is that HADS scores during pregnancy were known only for the women who sought consultation at the POP clinic. Other limitations are the lack of information on neonatal hair growth (and, thereby, the lack of standardization on newborn hair length), and the self-report of SAD use. In addition, the small sample size did not allow us to test associations by type of SAD. Lastly, ideally other tests of HPA axis activity, relating to HPA axis reactivity and rhythmicity, should have been included too.[46]

In conclusion, we found that maternal stress was associated with increased maternal hair GC levels and decreased neonatal hair GC levels, with the lowest values seen in children of mothers with persistent distress throughout pregnancy. It is suggested that maternal stress during pregnancy increased intra-uterine GC exposure, thereby suppressing fetal HPA axis activity.

REFERENCES

1. Knorr U, Vinberg M, Kessing LV, Wetterslev J. Salivary cortisol in depressed patients versus control persons: a systematic review and meta-analysis. Psychoneuroendocrinology 2010; 35:1275-1286
2. Lewis EJ, Yoon KL, Joormann J. Emotion regulation and biological stress responding: associations with worry, rumination, and reappraisal. Cogn Emot 2017:1-12
3. Vreeburg SA, Zitman FG, van Pelt J, Derijk RH, Verhagen JC, van Dyck R, Hoogendijk WJ, Smit JH, Penninx BW. Salivary cortisol levels in persons with and without different anxiety disorders. Psychosom Med 2010; 72:340-347
4. Zorn JV, Schur RR, Boks MP, Kahn RS, Joels M, Vinkers CH. Cortisol stress reactivity across psychiatric disorders: A systematic review and meta-analysis. Psychoneuroendocrinology 2017; 77:25-36
5. Austin MP, Hadzi-Pavlovic D, Priest SR, Reilly N, Wilhelm K, Saint K, Parker G. Depressive and anxiety disorders in the postpartum period: how prevalent are they and can we improve their detection? Arch Womens Ment Health 2010; 13:395-401
6. Dennis CL, Falah-Hassani K, Shiri R. Prevalence of antenatal and postnatal anxiety: systematic review and meta-analysis. Br J Psychiatry 2017; 210:315-323
7. Shakeel N, Eberhard-Gran M, Sletner L, Slinning K, Martinsen EW, Holme I, Jenum AK. A prospective cohort study of depression in pregnancy, prevalence and risk factors in a multi-ethnic population. BMC Pregnancy Childbirth 2015; 15:5
8. Duthie L, Reynolds RM. Changes in the maternal hypothalamic-pituitary-adrenal axis in pregnancy and postpartum: influences on maternal and fetal outcomes. Neuroendocrinology 2013; 98:106-115
9. Sandman CA, Davis EP, Buss C, Glynn LM. Exposure to prenatal psychobiological stress exerts programming influences on the mother and her fetus. Neuroendocrinology 2012; 95:7-21
10. Talge NM, Neal C, Glover V, Early Stress TR, Prevention Science Network F, Neonatal Experience on C, Adolescent Mental H. Antenatal maternal stress and long-term effects on child neurodevelopment: how and why? J Child Psychol Psychiatry 2007; 48:245-261
11. Glynn LM, Davis EP, Sandman CA. New insights into the role of perinatal HPA-axis dysregulation in postpartum depression. Neuropeptides 2013; 47:363-370
12. Trainer PJ. Corticosteroids and pregnancy. Semin Reprod Med 2002; 20:375-380
13. Demey-Ponsart E, Foidart JM, Sulon J, Sodoyez JC. Serum CBG, free and total cortisol and circadian patterns of adrenal function in normal pregnancy. J Steroid Biochem 1982; 16:165-169
14. Watterberg KL. Adrenocortical function and dysfunction in the fetus and neonate. Semin Neonatol 2004; 9:13-21
15. Kajantie E, Dunkel L, Turpeinen U, Stenman UH, Wood PJ, Nuutila M, Andersson S. Placental 11 beta-hydroxysteroid dehydrogenase-2 and fetal cortisol/cortisone shuttle in small preterm infants. J Clin Endocrinol Metab 2003; 88:493-500
16. D'Anna-Hernandez KL, Ross RG, Natvig CL, Laudenslager ML. Hair cortisol levels as a retrospective marker of hypothalamic-pituitary axis activity throughout pregnancy: comparison to salivary cortisol. Physiol Behav 2011; 104:348-353
17. Noppe G, de Rijke YB, Dorst K, van den Akker EL, van Rossum EF. LC-MS/MS-based method for long-term steroid profiling in human scalp hair. Clin Endocrinol (Oxf) 2015; 83:162-166
18. Noppe G, Van Rossum EF, Koper JW, Manenschijn L, Bruining GJ, de Rijke YB, van den Akker EL. Validation and reference ranges of hair cortisol measurement in healthy children. Horm Res Paediatr 2014; 82:97-102

19. Hollanders JJ, van der Voorn B, Kieviet N, Dolman KM, de Rijke YB, van den Akker ELT, Rotteveel J, Honig A, Finken MJJ. Interpretation of glucocorticoids in neonatal hair: a reflection of intrauterine glucocorticoid regulation? Endocr Connect 2017; 6:692-699

20. Romero-Gonzalez B, Caparros-Gonzalez RA, Gonzalez-Perez R, Delgado-Puertas P, Peralta-Ramirez MI. Newborn infants' hair cortisol levels reflect chronic maternal stress during pregnancy. PLoS One 2018; 13:e0200279

21. Kapoor A, Lubach GR, Ziegler TE, Coe CL. Hormone levels in neonatal hair reflect prior maternal stress exposure during pregnancy. Psychoneuroendocrinology 2016; 66:111-117

22. Hoffman MC, D'Anna-Hernandez K, Benitez P, Ross RG, Laudenslager ML. Cortisol during human fetal life: Characterization of a method for processing small quantities of newborn hair from 26 to 42 weeks gestation. Dev Psychobiol 2017; 59:123-127

23. Kieviet N, de Groot S, Noppe G, de Rijke YB, van Rossum EF, van den Akker EL, Dolman KM, Honig A. Is poor neonatal adaptation after exposure to antidepressant medication related to fetal cortisol levels? An explorative study. Early Hum Dev 2016; 98:37-43

24. Kieviet N, van Keulen V, van de Ven PM, Dolman KM, Deckers M, Honig A. Serotonin and poor neonatal adaptation after antidepressant exposure in utero. Acta Neuropsychiatr 2017; 29:43-53

25. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983; 67:361-370

26. Herrmann C. International experiences with the Hospital Anxiety and Depression Scale--a review of validation data and clinical results. J Psychosom Res 1997; 42:17-41

27. Matthey S, Valenti B, Souter K, Ross-Hamid C. Comparison of four self-report measures and a generic mood question to screen for anxiety during pregnancy in English-speaking women. J Affect Disord 2013; 148:347-351

28. Furdon SA, Clark DA. Scalp Hair Characteristics in the Newborn Infant. Advances in Neonatal Care 2003; 3:286-296

29. Pragst F, Balikova MA. State of the art in hair analysis for detection of drug and alcohol abuse. Clin Chim Acta 2006; 370:17-49

30. Smith RE, Maguire JA, Stein-Oakley AN, Sasano H, Takahashi K, Fukushima K, Krozowski ZS. Localization of 11 beta-hydroxysteroid dehydrogenase type II in human epithelial tissues. J Clin Endocrinol Metab 1996; 81:3244-3248

31. Dettmer AM, Rosenberg KL, Suomi SJ, Meyer JS, Novak MA. Associations between Parity, Hair Hormone Profiles during Pregnancy and Lactation, and Infant Development in Rhesus Monkeys (Macaca mulatta). PLoS One 2015; 10:e0131692

32. Meunier MR, Bennett IM, Coco AS. Use of antidepressant medication in the United States during pregnancy, 2002-2010. Psychiatr Serv 2013; 64:1157-1160

33. Short SJ, Stalder T, Marceau K, Entringer S, Moog NK, Shirtcliff EA, Wadhwa PD, Buss C. Correspondence between hair cortisol concentrations and 30-day integrated daily salivary and weekly urinary cortisol measures. Psychoneuroendocrinology 2016; 71:12-18

34. Merswolken M, Deter HC, Siebenhuener S, Orth-Gomer K, Weber CS. Anxiety as predictor of the cortisol awakening response in patients with coronary heart disease. Int J Behav Med 2013; 20:461-467

35. Norris CM, Ljubsa A, Hegadoren KM. Gender as a determinant of responses to a self-screening questionnaire on anxiety and depression by patients with coronary artery disease. Gend Med 2009; 6:479-487

36. Harris A, Endresen Reme S, Tangen T, Hansen AM, Helene Garde A, Eriksen HR. Diurnal cortisol rhythm: Associated with anxiety and depression, or just an indication of lack of energy? Psychiatry Res 2015; 228:209-215

37. de Rooij SR, Schene AH, Phillips DI, Roseboom TJ. Depression and anxiety: Associations with biological and perceived stress reactivity to a psychological stress protocol in a middle-aged population. Psychoneuroendocrinology 2010; 35:866-877

38. O'Donnell KJ, Bugge Jensen A, Freeman L, Khalife N, O'Connor TG, Glover V. Maternal prenatal anxiety and downregulation of placental 11beta-HSD2. Psychoneuroendocrinology 2012; 37:818-826

39. Seth S, Lewis AJ, Saffery R, Lappas M, Galbally M. Maternal Prenatal Mental Health and Placental 11beta-HSD2 Gene Expression: Initial Findings from the Mercy Pregnancy and Emotional Wellbeing Study. Int J Mol Sci 2015; 16:27482-27496

40. Togher KL, O'Keeffe MM, Khashan AS, Gutierrez H, Kenny LC, O'Keeffe GW. Epigenetic regulation of the placental HSD11B2 barrier and its role as a critical regulator of fetal development. Epigenetics 2014; 9:816-822

41. Togher KL, Treacy E, O'Keeffe GW, Kenny LC. Maternal distress in late pregnancy alters obstetric outcomes and the expression of genes important for placental glucocorticoid signalling. Psychiatry Res 2017; 255:17-26

42. Schoof E, Girstl M, Frobenius W, Kirschbaum M, Repp R, Knerr I, Rascher W, Dotsch J. Course of placental 11beta-hydroxysteroid dehydrogenase type 2 and 15-hydroxyprostaglandin dehydrogenase mRNA expression during human gestation. Eur J Endocrinol 2001; 145:187-192

43. Gareri J, Koren G. Prenatal hair development: implications for drug exposure determination. Forensic Sci Int 2010; 196:27-31

44. Slotkin TA, Seidler FJ, Wood CR, Lau C. Development of glucocorticoid receptor regulation in the rat forebrain: implications for adverse effects of glucocorticoids in preterm infants. Brain Res Bull 2008; 76:531-535

45. Liu CH, Snidman N, Leonard A, Meyer J, Tronick E. Intra-individual stability and developmental change in hair cortisol among postpartum mothers and infants: Implications for understanding chronic stress. Dev Psychobiol 2016; 58:509-518

46. Flom M, St John AM, Meyer JS, Tarullo AR. Infant hair cortisol: associations with salivary cortisol and environmental context. Dev Psychobiol 2017; 59:26-38

At the POP-clinic

At the maternity Ward

1st – 2nd trimester

46 completed
the HADS

3rd trimester

8 with no
HADS data

38 completed
the HADS

19 completed
the HADS

n = 57

Postpartum

n = 67

105 completed
the HADS

n = 172
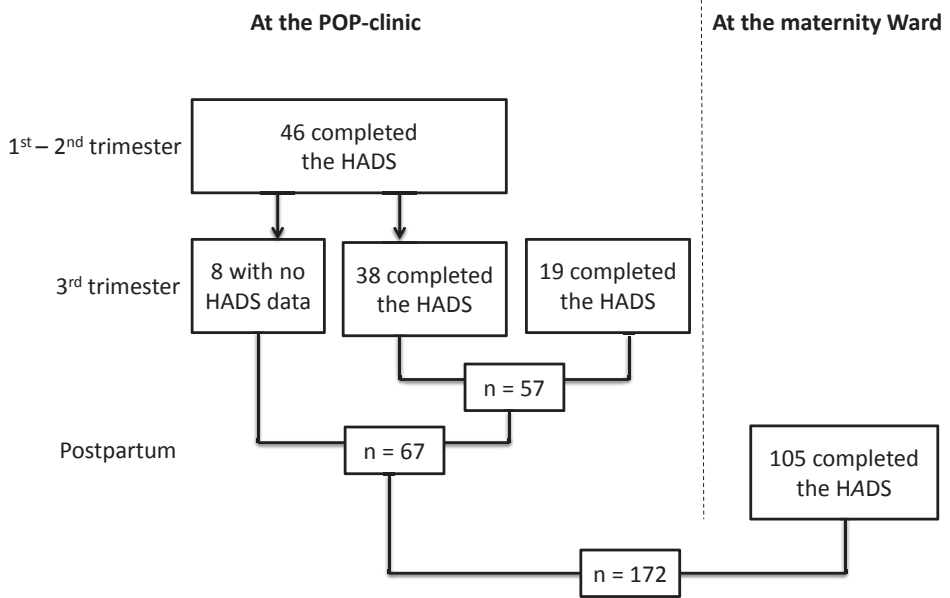
Supplementary Figure 1: Flowchart of pre- and postnatal assessments.

Supplementary Table 1: Characteristics of mother-infant pairs by time point

			Pairs with available HADS data in 1st-2nd trimester		Pairs with available HADS data in 3rd trimester		Pairs with available HADS data pp	
		N =	**46**		**57**		**172**	
Neonatal	Gestational age	wks	39.1	± 1.5	39.3	± 1.6	39.4	± 1.7
	Birth weight	g	3,328	± 540	3,370	± 497	3,446	± 583
		percentile	48.7	± 24.7	48.4	± 23.6	53.6	± 26.4
	Hair cortisol pp	pg/mg hair	156.3	(114.0 – 222.9)	154.3	(110.3 – 204.6)	162.8	(102.8 – 232.2)
	Hair cortisone pp	pg/mg hair	84.7	(64.5 – 109.1)	79.7	(63.9 – 106.3)	83.2	(63.1 – 109.8)
Maternal	Type of SAD	SSRI	33	(72%)	37	(65%)	45	(26%)
		SNRI	6	(13%)	5	(9%)	7	(4%)
		NaSSA	2	(4%)	8	(14%)	9	(5%)
		Combination*	3	(7%)	4	(7%)	5	(3%)
	Timing of SAD use	Entire pregnancy	41	(89%)	43	(75%)	54	(31%)
		≥3rd trimester	2	(4%)	3	(5%)	3	(2%)
		<3rd trimester	1	(2%)	8	(14%)	9	(5%)
	HADS scores 1st-2nd trimester	HAS score ≥8	17	(37%)	15	(26%)	17	(10%)
		HDS score ≥8	9	(20%)	8	(14%)	9	(5%)
	HADS scores 3rd trimester	HAS score ≥8	11	(24%)	23	(40%)	23	(13%)
		HDS score ≥8	7	(15%)	14	(25%)	14	(8%)
	HADS scores pp	HAS score ≥8	13	(28%)	17	(30%)	30	(17%)
		HDS score ≥8	7	(15%)	9	(16%)	19	(11%)
	Hair cortisol pp	pg/mg hair	6.9	(3.9 – 13.9)	7.1	(4.6 – 13.9)	5.4	(3.6 – 10.6)
	Hair cortisone pp	pg/mg hair	23.1	(16.9 – 62.1)	25.3	(17.5 – 54.8)	19.5	(14.5 – 31.2)

Data are expressed as mean ± SD, median (interquartile range), or n (%). Abbreviations: pp = postpartum; SAD = Selective antidepressant; HADS = Hospital Anxiety and Depression Scale; HAS = Hospital Anxiety Scale; HDS = Hospital Depression Scale; SSRI = selective serotonin reuptake inhibitors; SNRI = serotonin-norepinephrine reuptake inhibitors, NaSSA = noradrenergic or specific serotonin antidepressants.
*These subjects were treated with SSRI and NaSSA (n =4), or with NaSSA and SNRI (n =1).

EVA MOUTON 2019

4

Nutritional programming by glucocorticoids in breast milk: Targets, mechanisms and possible implications

Jonneke J. Hollanders,
Annemieke C. Heijboer,
Bibian van der Voorn,
Joost Rotteveel,
Martijn J.J. Finken

ABSTRACT

Vertical transmission of glucocorticoids via breast milk might pose a mechanism through which lactating women could prepare their infants for the postnatal environment. The primary source of breast-milk glucocorticoids is probably the systemic circulation. Research from our group showed that milk cortisol and cortisone concentrations follow the diurnal rhythm of maternal hypothalamus-pituitary-adrenal axis activity, with a higher abundance of cortisone compared to cortisol. Measurement of breast-milk glucocorticoid concentrations is challenging due to possible cross-reactivity with progestagens and sex steroids, which are severely elevated during pregnancy and after parturition. This requires precise methods that are not hindered by cross reactivity, such as LC-MS/MS. There are some data suggesting that breast-milk glucocorticoids could promote intestinal maturation, either locally or after absorption into the systemic circulation. Breast-milk glucocorticoids might also have an effect on the intestinal microbiome, although this has not been studied thus far. Findings from studies investigating the systemic effects of breast-milk glucocorticoids are difficult to interpret, since none took the diurnal rhythm of glucocorticoids in breast milk into consideration, and various analytical methods were used. Nevertheless, glucocorticoids in breast milk might offer a novel potential pathway for signal transmission from mothers to their infants.

INTRODUCTION

Numerous studies suggest that adversities occurring in early life could predispose to later diseases such as cardiovascular diseases, type 2 diabetes mellitus and neuropsychiatric diseases. The mechanisms that could explain these associations relate to the concept of early-life programming, stating that insults early in life could persistently alter the body's structure and/or function. These alterations, although adaptive in nature, might become deleterious with age.[1]

Glucocorticoids are known for their programming effects on metabolism and the brain.[2] The fetal cortisol hypothesis postulates that a lower activity of the placental barrier enzyme 11β-hydroxysteroid dehydrogenase (11β-HSD) type 2 allows a larger proportion of maternal cortisol to reach the fetus, leading to permanent alterations in hypothalamus-pituitary-adrenal (HPA) axis settings, and, hence, predisposition to cardiometabolic and neuropsychiatric diseases in offspring.[3]

Emerging data suggest that postnatal nutrition could play a role in early-life programming.[4] Breast feeding has been associated with improved health outcomes, including reduced risks of infections and obesity.[4,5] Although glucocorticoids were recovered in breast milk already in the early 1970s,[6] only few studies have addressed their effects in offspring. The recent discovery of a diurnal rhythm in the secretion of glucocorticoids into breast milk has opened new avenues for the study of the postnatal programming effects of maternal glucocorticoids.

GLUCOCORTICOIDS, PREGNANCY AND THE MAMMARY GLAND

Cortisol is produced by the zona fasciculata of the adrenal cortex. Its synthesis is regulated by adrenocorticotropin hormone (ACTH) from the anterior pituitary gland. The release of ACTH, in turn, is under the control of corticotropin releasing hormone (CRH) from the hypothalamus. When cortisol is present in adequate amounts, a negative feedback system operates on the pituitary gland and hypothalamus. The hypothalamus also receives input from multiple brain areas involved in the stress response. In healthy individuals, the secretion of cortisol follows a diurnal rhythm, with a peak in the early morning, followed by gradual decline over the day, and a nadir at midnight.

Pregnancy-induced changes in HPA axis activity

Maternal cortisol increases sharply in the last part of gestation due to an exponential rise in the secretion of CRH from placental origin, which stimulates the release of ACTH from the pituitary.[7] In contrast to the inhibitory effect of glucocorticoids on the secretion of CRH by the hypothalamus, glucocorticoids stimulated the expression of the CRH

gene in cultures of human placenta.[8] In the fetal compartment a similar rise in cortisol occurs, which is necessary for the maturation of several organs, such as the lungs and the liver.[9] The cortisol rise has also been implicated to play a pivotal role in the onset of parturition.[9]

Glucocorticoids as hormonal regulators of lactation

Glucocorticoids seem to be involved in the lobulo-alveolar development of the mammary gland during the last stage of pregnancy.[10] Although glucocorticoid receptors were detectable in the lactating mouse mammary gland,[11] the increase in circulating glucocorticoids at parturition is probably not the primary trigger of lactogenesis .[12] The effects of glucocorticoids on lactogenesis are probably more permissive, by enhancing the effects of prolactin and spermidine on the synthesis of α-lactalbumin and casein.[12] However, in bovine mammary epithelium glucocorticoids were found to have direct effects on the sodium transport.[13] Progesterone could antagonize the actions of glucocorticoids at the level of the glucocorticoid receptor.[14] Therefore, the effects of glucocorticoids on the mammary gland may become more evident after parturition, when progesterone levels fall.

Determinants of breast-milk glucocorticoid concentrations

Owing to their lipophilic structure, glucocorticoids are able to cross mammary epithelia through simple diffusion in the direction of their concentration gradient. The primary source of breast-milk glucocorticoids is probably the systemic circulation. This assumption is based on observations showing that breast-milk cortisol was strongly correlated with plasma cortisol,[15] with the concentration in breast milk being about 1-13% of the circulating level.[15-17] Another possible source of breast-milk glucocorticoids is the skin, which was recently shown to be capable of glucosteroidogenesis,[18] although it is currently unclear whether dermal glucocorticoids could penetrate the mammary gland in significant amounts.

The concentration of breast-milk cortisol was relatively high in the last part of gestation, compatible with the rise in plasma cortisol, followed by a >50% decline within 2 days after delivery.[16] Mothers who delivered prematurely had lower levels of breast-milk glucocorticoids,[19] either due to an earlier disruption of the positive feedback loop in the secretion of placental CRH or immaturity of the mammary gland.

We studied lactating women who provided frequent sample collections over a 24-hr period, when their children had reached the age of 1 month, and found that levels of cortisol and cortisone in breast milk peaked at 7:00 am, mirroring those in saliva obtained at the same time.[19] The early-morning peak was approximately five times as high as the nadir. This typical diurnal rhythm was replicated by another group.[20] We also found that cortisone was much more abundant in breast milk than cortisol, despite the observation that serum cortisone constitutes only a minor fraction of circulating glucocorticoids.

These patterns could probably be attributed to a high expression of 11β-HSD type 2 in the mammary gland, analogous to the salivary gland.[21]

MEASUREMENT OF GLUCOCORTICOID CONCENTRATIONS IN BREAST MILK

Immunoassay versus LC-MS/MS

For the study of glucocorticoids in breast milk, reliable analytical methods are necessary. Steroid hormones can be measured with immunoassays or chromatography, possibly coupled to mass spectrometry. In general, immunoassays are prone to cross reactivity with compounds that share the general structure with the hormone of interest, necessitating the use of more specific analytical methods.[22] This is especially important during pregnancy and after parturition, when concentrations of steroid hormones like progestagens and sex steroids, as well as their precursors and metabolites, are severely elevated. The most specific method today is chromatography coupled to mass spectrometry, with liquid chromatography tandem mass spectrometry (LC-MS/MS) being regarded as the method of choice. Furthermore, LC-MS/MS analysis carries the advantage of multiple steroid hormone measurements during the same run, enabling the simultaneous measurement of cortisol and cortisone.[23]

Binding to proteins

In serum or plasma, glucocorticoids are for >90% bound to corticosteroid binding globulin (CBG) or albumin.[24] However, albumin-bound and free fractions increase, whereas the CBG-bound fraction decrease, at increasing levels of total cortisol.[25] Steroids in breast milk are also protein-bound, with CBG or a CBG-like protein having been identified in human breast milk already in 1976.[26] This protein was found to decline rapidly after parturition.[26] The presence of corticosteroid binding proteins in breast milk has implications for the choice of the analytical method. Measurement of the total cortisol concentration requires that cortisol is displaced from CBG and other corticosteroid binding proteins. Many immunoassays suffer from ineffective displacement of cortisol from its binding protein, leading to falsely lower cortisol concentrations in the presence of higher CBG concentrations.[27] The improper release of a hormone from its binding protein has also been identified in immunoassays for testosterone and 25-hydroxy-vitamin D3,[28,29] illustrating that this is a widespread analytical problem. Underestimation of the total cortisol concentration is unlikely to occur with LC-MS/MS analysis, where the organic solvents used allow a proper release of cortisol from its binding protein. Alternatively, free glucocorticoid concentrations in breast milk could be measured, yet this requires even more sophisticated methods, such as equilibrium dialysis or ultrafiltration, prior to

LC-MS/MS analysis. Moreover, the very low concentrations might challenge the sensitivity of current LC-MS/MS analyzers.

Steroid conjugates

Many steroid hormones form conjugates with sulfate or glucuronic acid, which renders them inactive. Ninety-five percent of steroids recovered from breast milk were conjugated, predominantly to sulfate.[17] Steroid conjugates cannot be absorbed by the intestines, which requires hydrolysis by sulfatases or glucuronidases.[30] In the human gut, these enzymes are absent at birth and they increase with age due to bacterial colonization.[30] Therefore, incubation of the breast milk with conjugate enzymes could yield much higher levels that do not reflect the biologically available, clinically relevant fraction.

Comparisons between studies

In view of these analytical issues, comparing results of studies concerning breast-milk glucocorticoids is difficult. In addition, differences in standardization between methods used for steroid hormone analysis may contribute to these issues.[31-33] We have previously reviewed human breast-milk cortisol concentrations as reported in the literature, and found a wide range of concentrations, varying from 0 to 1,700 nmol/L.[23] A large variety of methods was used by these studies, the majority of which used immunoassays, sometimes preceded by a deconjugation step. However, the wide range of concentrations, notably those reported by Groer et al,[34] could not wholly be explained by the variety in methods.

Given all these uncertainties, it is highly important for future studies to carefully select and validate the method that will be used so that proper conclusions are drawn. We have developed and extensively validated an isotope-diluted LC/MS-MS method for the measurement of total cortisol and cortisone in human breast milk without use of enzymatic deconjugation.[23] Reference ranges of our assay are reported in Table 1.

Table 1: Reference ranges of breast-milk cortisol and cortisone measured with our LC-MS/MS assay.

Sampling time (hrs)	N	Cortisol (nmol/L)	Cortisone (nmolL)
0:00-6:00	46	3.3 (1.4-6.8)	17.6 (10.8-23.6)
6:00-12:00	64	8.3 (5.1-12.2)	29.6 (24.1-35.0)
12:00-18:00	63	2.8 (2.1-4.5)	18.9 (14.9-31.3)
18:00-24:00	64	1.1 (0.8-1.9)	10.1 (6.8-12.7)

Concentrations are reported as median (interquartile range).

TARGETS OF GLUCOCORTICOIDS IN THE DEVELOPING GUT

Fate of breast-milk glucocorticoids in the neonatal gut

There is some evidence showing that breast-milk glucocorticoids are absorbed by the developing gut into the systemic circulation. Experiments in rats showed that milk-ingested labeled corticosterone – the principal glucocorticoid in rats – was able to cross the pups' intestinal epithelial barrier and was, subsequently, detectable in their plasma and brains.[35] Additionally, corticosterone was detectable in the serum of adrenalecto-mized pups fed with their own mother's milk.[36] In humans, a 40% higher salivary cortisol level was found among infants who were breastfed during one year in comparison with infants who were formula-fed.[37] Moreover, another study found that salivary cortisol concentration in infants correlated positively with the level in their mothers only in those who were breast-fed.[38]

Effects of breast-milk glucocorticoids on intestinal maturation

There is overwhelming evidence, mostly from studies in rats, indicating that glucocorticoids are important for intestinal maturation. Rat pups can be compared to preterm infants until postnatal day 20, at which stage their development resembles that of term infants. Therefore, rat pups are ideal for the study of gut development. Glucocorticoids administered systemically were found to increase the activities of lactase and sucrase, and fucosylation, and to decrease sialylation,[39] which are all indicative of intestinal matu-ration. Other studies have assessed the effect of enterally supplemented glucocorticoids on gut maturation. Yeh et al.[36] studied the effect of enterally supplemented corticoste-rone on the digestive capacity of the intestine of rats that were adrenalectomized at day 12. Adrenalectomized pups were fed with formula supplemented with either 0, 0.1, 0.5, 1.0, 5.0, 10.0 or 50.0 µg/ml corticosterone, or with their own mother's milk. Intact pups, serving as controls, were either mother-fed or fed with unsupplemented formula. They found a dose-dependent increase in the activities of sucrase and maltase with increased glucocorticoid supplementation. In comparison, mother-fed pups, both intact and adrenalectomized, had maltase activity comparable to pups fed with formula supple-mented with 0-0.5 µg/ml corticosterone, and undetectable sucrase activity, indicating that the maturational effect of corticosterone might only occur when administered at supraphysiological doses. However, systemic glucocorticoid levels were detectable in mother-fed adrenalectomized pups and all of these pups survived, compared to only 11% of the adrenalectomized pups that did not receive corticosterone-supplemented formula. It appeared, however, that when glucocorticoids were administered systemi-cally, they were more effective in inducing maturation than enterally administered glu-cocorticoids, and it was thus suggested that enteral glucocorticoids probably influence gut maturation after absorption into the systemic circulation.[36]

Teichberg et al.[40] studied the effect of supplementing formula with corticosterone in early-weaned (i.e., postnatal day 17, whereas normal weaning occurs at age 21 days) rats. They found that supplying formula with corticosterone at levels normally found in the breast milk of dams (i.e., 0.26 μmol/L, by use of a local immunoassay) prevented the delay in jejunal closure to macromolecule uptake usually seen with early weaning. These pups were already relatively mature, and therefore these findings might be extrapolated to term-born infants.[40] However, contrary to rats, normal gut development in humans does not seem to rely on the uptake of macromolecules, and the amount of macromolecule uptake appears to be lower in humans than in rats.[41]

Studies in humans, or in ex-vivo human tissues, were all designed to investigate the effects on the developing gut of glucocorticoid treatment given to mothers presenting with impending preterm delivery. Nanthakumar et al.[42] studied human intestinal xenografts implanted in mice treated with a single injection of cortisone acetate. They found an increase in lactase activity, as well as a decrease in the immune response to both endogenous and exogenous inflammatory stimuli, in immature (20-week-old) intestine. These results were not seen in more mature (30-week-old) intestine. They concluded that glucocorticoids could accelerate intestinal maturation, but that this effect was restricted to earlier pregnancy. Villa et al.[43] studied the effect of hydrocortisone on in vitro cultured human intestine from 10 fetuses with gestational age 14-20 weeks and from one term newborn. Hydrocortisone induced a 2-fold increase in lactase activity after 5 days of culture, although it did not affect lactase mRNA, suggestive of a possible posttranscriptional effect on lactase activity.[43]

Only few in vivo studies have been conducted in humans, none of them studying the effect of enteral glucocorticoids. Costalos et al.[44] addressed the effect of antenatal glucocorticoids on the secretion of gastrointestinal peptides in preterm infants (gestational age <34 weeks), both immediately after birth and after the initiation of enteral feeding. They found higher gastrin levels from birth onwards in infants who had received antenatal glucocorticoids (n=28), and higher motilin levels only after the initiation of enteral feeding compared to infants who did not receive antenatal glucocorticoids (n=17). Antenatal glucocorticoids did not influence vasointestinal peptide. Gastrin promotes gastric secretion, while motilin accelerates gastric emptying, suggesting that antenatal treatment with glucocorticoids might aid in improving digestion. Watkins et al.[45] studied 9 preterm infants (gestational age: 32-36 weeks), of whom four had received dexamethasone (n=3) or phenobarbital (n=1) antenatally. In these infants, the bile salt pool and production were markedly increased, approaching levels of full-term infants, implying a more mature liver and possibly gastrointestinal tract.[45]

To summarize, glucocorticoids, whether administered enterally or systemically, seem to have a trophic effect on the intestine in both animals and humans, although this effect appears to be restricted to a limited time window. Whether cortisol in breast

milk has trophic effects on the intestines of term-born human neonates has yet to be determined.

Effects of breast-milk glucocorticoids on the gut microbiota

Glucocorticoids might play a role in the composition of the microbiome. Although no studies have been conducted aimed at the development of the microbiome after the enteral administration of glucocorticoids, several studies have researched the interaction between HPA axis activity and the microbiome, which is part of the "gut-brain axis hypothesis".[46,47]

Zijlmans et al.[48] studied maternal prenatal stress, defined as elevated basal salivary cortisol or as reported by mothers, and the development of the neonatal microbiome. They found that increased maternal prenatal stress was associated with alterations in the microbiome during the first 110 days after birth, with an increased abundance of Proteobacteria such as Escherichia, and a lower abundance of lactic acid bacteria and Actinobacteria. Moreover, this pattern was associated with an increase in maternally reported gastrointestinal and allergic symptoms.[48] In animals, maternal separation early in life has been shown to alter the diversity and composition of the microbiome,[49] while treatment with Lactobacillus during separation appears to normalize basal corticosterone levels in rats.[50] Additionally, stress induced by maternal separation was associated with increased bacterial translocation in 10-day-old rat pups.[51] The role of the microbiome on HPA axis activity has also been demonstrated in mice. Germ-free mice had an elevated stress response, while mono-colonization with Bifidobacterium infantis was associated with an attenuated stress response and mono-colonization with Escherichia coli with an even stronger response to stressful insults,[52] implicating an effect of the microbiome on HPA axis activity.

Gut bacteria were found to have steroid-converting properties. Not only have they been implicated to play a role in the metabolism of cortisol, they probably also metabolize glucocorticoid precursors and metabolites.[53-57] Some of these compounds enter the enterohepatic circulation, but their biological effect is unknown.

Therefore, although the developmental timing for local, trophic effects of glucocorticoids on the intestine is probably before birth, the interaction between glucocorticoids and the gut microbiome (i.e., the "gut-brain axis") might be a possible mechanism through which glucocorticoids in breast milk could exert systemic effects in the infant.

Table 2: Summary of studies addressing the systemic effects of breast-milk glucocorticoids in offspring.

Study	N	Species	Age at breast-milk sampling	Timing of breast-milk sampling (hrs)	Outcomes	Main results
Primates						
Sullivan et al. (2011)	44	Rhesus monkey	3-4 mo.	NR	Behavioral observations over 25-hr period	Mothers of males had higher cortisol concentrations in milk than did mothers of females, and cortisol concentrations in maternal milk were related to a confident temperament factor in sons, but not daughters.
Hinde et al. (2015)	108	Rhesus monkey	1 mo. and 3.5 mo.	11.30–13.00 after 3.5–4 h of milk accumulation	Behavioral observations over 25-hr period	Milk cortisol was positively associated with a more nervous, less confident temperament in offspring, albeit with few differences between genders. Milk cortisol was positively associated with weight gain in offspring.
Dettmer et al. (2017)	34	Rhesus monkey	7 and 21 d, or 14 and 30 d	13.45-14.00	Social behavior and cognitive function (inhibitory control, impulsivity)	Milk cortisol during the first month post-partum was positively associated with impulsivity on a cognitive task, but not global social behaviors, months later.
Rodents						
Catalani et al. (2000)	NR	Wistar rats (m)	NA	NA	Restraint stress, learning test (conditioned avoidance learning), tests of anxiety (elevated plus-maze test, dark-light test) and hippocampal corticosteroid receptor expression	CORT-exposed pups exhibited a lower adrenocortical response to stress, improvements in learning, reduced fearfulness in anxiogenic situations and greater numbers of hippocampal corticosteroid receptors.
Catalani et al. (2002)	NR	Wistar rats (f)	NA	NA	Restraint stress, learning tests (Morris water maze, aquatic T maze, conditioned avoidance learning), tests of anxiety (elevated plus-maze test, dark-light test, conditioned suppression of drinking) and hippocampal corticosteroid receptor expression	CORT-exposed pups exhibited a lower adrenocortical response to stress, improvements in learning and reduced fearfulness in anxiogenic situations. The numbers of hippocampal corticosteroid receptors was no different from unexposed pups.

Table 2: Summary of studies addressing the systemic effects of breast-milk glucocorticoids in offspring. (continued)

Study	N	Species	Age at breast-milk sampling	Timing of breast-milk sampling (hrs)	Outcomes	Main results
Man						
Hart et al. (2004)	32	-	7-11 d	Mid-morning	Neonatal Behavioral Assessment Scale	Milk cortisol was positively correlated with the Autonomic Stability cluster.
Grey et al. (2013)	52	-	3 mo.	NR	Infant Behavioral Questionnaire	Milk cortisol was positively associated with negative affectivity.
Hahn-Holbrook et al. (2016)	51	-	3 mo.	11.30-16.00	Length and weight until age 2	Milk cortisol was inversely associated with BMI gains in the first 2 years of life.

Abbreviations: CORT = corticosterone; f = female; m = male; NA = not applicable; NR = not reported

SYSTEMIC EFFECTS OF BREAST-MILK GLUCOCORTICOIDS IN OFFSPRING

Surprisingly few studies have investigated associations between breast-milk glucocorticoids and systemic effect in offspring. Here, we review the evidence from studies in both animals and man. The findings from these studies are summarized in Table 2.

Studies in primates

The few animal studies on this topic have nearly all been conducted in Rhesus macaques. The choice for Rhesus monkeys was based on data showing marked variation in behavior and temperament as well as HPA axis vulnerability to early-life experiences in this species.[58]

Sullivan et al.[58] investigated 44 mothers and their offspring at 3-4 months of age. Cortisol was measured in breast milk, and in the plasma of mothers and their infants. Infant temperament was assessed during a 25-hr behavioral observation. Milk cortisol was correlated with maternal plasma cortisol (r = 0.586, P = 0.017) as well as with protein (r = 0.441, P = 0.003) and fat contents in milk (r = 0.398, P = 007). Milk cortisol was higher in the mothers of males (p = 0.043). In males, milk cortisol was correlated with confident behavior (r = 0.669, P = 0.002), but not with the other temperament factors vigilance, gentleness or nervousness. In females, milk cortisol was not correlated with temperament.

Hinde et al.[59] assessed associations between milk cortisol, available milk energy (i.e., the product of milk yield and milk energy density), infant temperament and weight gain in 108 mother-infant dyads. Breast milk was analyzed at 1 month and at 3.5 months

post-partum. Milk cortisol was positively associated with milk fat (P = 0.038) and protein contents (P = 0.0064). A higher milk yield was associated with a lower milk cortisol concentration (P = 0.0014), attributable to dilution. Lower-parity mothers produced milk with a higher cortisol concentration. There were no effects of social ranking or age at sampling on the milk cortisol concentration, and the effect of maternal body weight was negligible. Milk cortisol, independent of available milk energy, predicted a more nervous, less confident temperament, with few differences between genders. It was positively associated with infant weight gain.

Dettmer et al.[60] investigated 34 mother-infant dyads from birth to 8 months of age. Milk collections for determination of cortisol occurred twice in the first month post-partum, and the mean concentration was calculated. At 8 months, social behavior and cognitive function were measured. Mean milk cortisol was positively associated with impulsivity on a cognitive task, and, in females only, with total frequency of play.

Studies in rodents

There are few data suggesting that adult offspring from rat mothers who had free access to 200 µg/ml corticosterone hemisuccinate solution from day 1 post-partum until weaning exhibited longer-term sequelae. These sequelae included a better performance during learning tests, reduction of fearfulness in anxiogenic situations, lower stress-induced corticosterone secretion, and, in males only, greater numbers of hippocampal corticosteroid receptors, as compared to unexposed animals.[61,62] Corticosterone was not measured in milk during these experiments, although previous data showed subsequent elevations in maternal plasma and milk as well as in the gastric contents, plasma and brains of offspring.[35]

Studies in man

Hart et al.[63] investigated associations between maternal mood, milk cortisol, and neonatal behavior in 32 mother-infant pairs. Maternal depressive symptoms were assessed at 7-11 days post-partum using the Beck Depression Inventory (BDI), the Profile of Mood States (POMS) and the State Trait Anxiety Inventory (STAI). At the same time, neonatal behavior was assessed using the Neonatal Behavioral Assessment Scale (NBAS). Milk cortisol was correlated with the POMS Hostility scale (r = 0.35, P < 0.05), suggesting that mothers' greater hostility coincided with higher concentrations of milk cortisol, but not with scores on the other POMS subscales, or on the BDI or the STAI. Milk cortisol was correlated with the Autonomic Stability cluster of the NBAS (r = 0.40, P < 0.05). Due to lack of follow-up data, the sustainability of these associations could not be tested.

Grey et al.[64] tested whether milk cortisol is associated with infant temperament in 52 mother-infant pairs. Milk cortisol and infant temperament, by maternal report of the Infant Behavior Questionnaire (IBQ), were assessed when the infants had reached the

age of 3 months. Analyses revealed a positive association between milk cortisol and the negative affectivity dimension of the IBQ (r = 0.37, P < 0.01). No correlation was found between milk cortisol and the surgery/extraversion or the orienting/regulation dimensions of the IBQ. Again, the sustainability of observations was not tested. In the same cohort, Hahn-Holbrook et al.[65] tested associations between milk cortisol and change in body mass index (BMI) across time. They found that a higher milk cortisol was associated with lower gains in BMI in the first 2 years of life (P = 0.046).

For several reasons, the findings obtained from these studies were highly contradictory. All of them had measured cortisol (but not cortisone) by use of immunoassays, which are prone to cross reactivity. Moreover, studies differed in sampling procedures. None of them had obtained multiple samples across the day, and only one adjusted the analyses for timing of sampling. However, we have previously argued that, in spite of correction for time of collection, vertical transmission of glucocorticoids cannot be determined from a single milk specimen.[66] This reasoning is based on observations showing that there was a large inter-individual variability in the diurnal rhythmicity of salivary cortisol.[67] Therefore, frequent sample collections over a 24-hr period are necessary to obtain meaningful indices of glucocorticoid exposure, such as peak, nadir, diurnal variability and area under the curve.

CONCLUSIONS AND FUTURE PROSPECTS

Given the methodological shortcomings of the studies conducted thus far, at present there is no compelling evidence for persistent effects of breast-milk glucocorticoids in offspring. Future studies should use sound methodologies to test associations between breast-milk glucocorticoids and infant developmental pathways, including frequent sampling and LC-MS/MS analysis with simultaneous measurement of cortisol and cortisone.

Recent evidence from studies in adults suggests that flattening of HPA axis rhythmicity predicts the onset of type 2 diabetes mellitus and cardiovascular mortality.[68,69] Such patterns have previously been associated with major depressive disorder and other psychopathologies, albeit not unequivocally.[70,71] If associations between flattened HPA axis rhythmicity in breast milk and adverse infant outcomes are demonstrated, lactating women should be requested not to express their milk, or to give their infants expressed milk that was obtained at the same time of the day. Such results could also have implications for donor milk banks, which provide donor human milk to vulnerable infants when own mother's milk is not available. In clinical practice, preterm infants are often fed with pooled donor milk, which lacks a diurnal rhythm. Our group has recently demonstrated

that pasteurization, necessary to secure the safety of donor milk, does not affect gluco-corticoid contents in milk.[72]

Future studies should also explore the mechanisms, including possible sexually dimorphic effect as observed in animals, that could underpin associations between breast-milk glucocorticoids and infant outcomes. Studies conducted to date have mainly focused on the effects glucocorticoids might have on intestinal maturation. The effects on the gut microbiota remain to be explored. Gut bacteria were found to be sensitive to disruptions in circadian clock rhythms,[73] but the role of disturbances in HPA axis rhythmicity has never been studied. Furthermore, the steroid-converting properties of gut bacteria require further study. The implication of microbial 11β-reductase activity, if present, is that cortisol could be produced in the gut from inert cortisone. However, the net effect will be unclear, as intestinal epithelia were found to express 11β-HSD type 2, which catalyzes the reverse reaction.[21]

In conclusion, breast-milk glucocorticoids follow the diurnal rhythm of maternal HPA axis activity probably through simple diffusion from the systemic circulation. Breast-milk glucocorticoids might have (direct or indirect) effects in infants, both locally (e.g., intestinal maturation, microbiome) and systemically (e.g., growth, body composition, neurodevelopment). In order to investigate this, the diurnal rhythm of glucocorticoids should be taken into account and the analytical methods should be chosen carefully. Nevertheless, glucocorticoids in breast milk might offer a novel potential pathway for signal transmission from mothers to their infants.

REFERENCES

1. Gluckman PD, Hanson MA, Pinal C. The developmental origins of adult disease. Matern Child Nutr 2005; 1:130-141
2. Cottrell EC, Seckl JR. Prenatal stress, glucocorticoids and the programming of adult disease. Front Behav Neurosci 2009; 3:19
3. Edwards CR, Benediktsson R, Lindsay RS, Seckl JR. Dysfunction of placental glucocorticoid barrier: link between fetal environment and adult hypertension? Lancet 1993; 341:355-357
4. Patro-Golab B, Zalewski BM, Kolodziej M, Kouwenhoven S, Poston L, Godfrey KM, Koletzko B, van Goudoever JB, Szajewska H. Nutritional interventions or exposures in infants and children aged up to 3 years and their effects on subsequent risk of overweight, obesity and body fat: a systematic review of systematic reviews. Obes Rev 2016; 17:1245-1257
5. Kramer MS, Kakuma R. Optimal duration of exclusive breastfeeding. Cochrane Database Syst Rev 2012:CD003517
6. Tucker HA, Schwalm JW. Glucocorticoids in mammary tissue and milk. J Anim Sci 1977; 45:627-634
7. Glynn LM, Davis EP, Sandman CA. New insights into the role of perinatal HPA-axis dysregulation in postpartum depression. Neuropeptides 2013; 47:363-370
8. Robinson BG, Emanuel RL, Frim DM, Majzoub JA. Glucocorticoid stimulates expression of cortico-tropin-releasing hormone gene in human placenta. Proc Natl Acad Sci U S A 1988; 85:5244-5248
9. Braun T, Challis JR, Newnham JP, Sloboda DM. Early-life glucocorticoid exposure: the hypotha-lamic-pituitary-adrenal axis, placental function, and long-term disease risk. Endocr Rev 2013; 34:885-916
10. Topper YJ, Freeman CS. Multiple hormone interactions in the developmental biology of the mam-mary gland. Physiological reviews 1980; 60:1049-1106
11. Shyamala G. Specific cytoplasmic glucocorticoid hormone receptors in lactating mammary glands. Biochemistry 1973; 12:3085-3090
12. Neville Me. Lactation: Physiology, Nutrition, and Breast-Feeding. New York: Plenum Press.
13. Quesnell RR, Han X, Schultz BD. Glucocorticoids stimulate ENaC upregulation in bovine mam-mary epithelium. Am J Physiol Cell Physiol 2007; 292:C1739-1745
14. Shyamala G, Dickson C. Relationship between receptor and mammary tumour virus production after stimulation by glucocorticoid. Nature 1976; 262:107-112
15. Patacchioli F, Cigliana G. Maternal plasma and milk free cortisol during the first 3 days of breast-feeding following spontaneous delivery or elective cesarean section. Gynecolog Obstet Invest 1992; 34:159-163
16. Kulski JK, Hartmann PE. Changes in the concentration of cortisol in milk during different stages of human lactation. Aust J Exp Biol Med Sci 1981; 59:769-778
17. Sahlberg BL, Axelson M. Identification and quantitation of free and conjugated steroids in milk from lactating women. Journal of steroid biochemistry 1986; 25:379-391
18. Slominski AT, Manna PR, Tuckey RC. On the role of skin in the regulation of local and systemic steroidogenic activities. Steroids 2015; 103:72-88
19. van der Voorn B, de Waard M, van Goudoever JB, Rotteveel J, Heijboer AC, Finken MJ. Breast-Milk Cortisol and Cortisone Concentrations Follow the Diurnal Rhythm of Maternal Hypothalamus-Pituitary-Adrenal Axis Activity. J Nutr 2016; 146:2174-2179

20. Pundir S, Wall CR, Mitchell CJ, Thorstensen EB, Lai CT, Geddes DT, Cameron-Smith D. Variation of Human Milk Glucocorticoids over 24 hour Period. J Mammary Gland Biol Neoplasia 2017; 22:85-92

21. Smith RE, Maguire JA, Stein-Oakley AN, Sasano H, Takahashi K, Fukushima K, Krozowski ZS. Localization of 11 beta-hydroxysteroid dehydrogenase type II in human epithelial tissues. J Clin Endocrinol Metab 1996; 81:3244-3248

22. Ackermans MT, Endert E. LC-MS/MS in endocrinology: what is the profit of the last 5 years? Bioanalysis 2014; 6:43-57

23. van der Voorn B, Martens F, Peppelman NS, Rotteveel J, Blankenstein MA, Finken MJ, Heijboer AC. Determination of cortisol and cortisone in human mother's milk. Clin Chim Acta 2015; 444:154-155

24. Nguyen PT, Lewis JG, Sneyd J, Lee RS, Torpy DJ, Shorten PR. Development of a formula for estimating plasma free cortisol concentration from a measured total cortisol concentration when elastase-cleaved and intact corticosteroid binding globulin coexist. J Steroid Biochem Mol Biol 2014; 141:16-25

25. Dorin RI, Pai HK, Ho JT, Lewis JG, Torpy DJ, Urban FK, 3rd, Qualls CR. Validation of a simple method of estimating plasma free cortisol: role of cortisol binding to albumin. Clin Biochem 2009; 42:64-71

26. Payne DW, Peng LH, Pearlman WH. Corticosteroid-binding proteins in human colostrum and milk and rat milk. J Biol Chem 1976; 251:5272-5279

27. Vos MJ, Bisschop PH, Deckers MML, Endert E. The cortisol-CBG ratio affects cortisol immunoassay bias at elevated CBG concentrations. Clin Chem Lab Med 2017;

28. Heijboer AC, Blankenstein MA, Kema IP, Buijs MM. Accuracy of 6 routine 25-hydroxyvitamin D assays: influence of vitamin D binding protein concentration. Clin Chem 2012; 58:543-548

29. Heijboer AC, Savelkout E, Kruit A, Endert E, Blankenstein MA. Inaccurate First-Generation Testosterone Assays Are Influenced by Sex Hormone–Binding Globulin Concentrations. The Journal of Applied Laboratory Medicine 2016; 2:194-201

30. Kent TH, Fischer LJ, Marr R. Glucuronidase activity in intestinal contents of rat and man and relationship to bacterial flora. Proc Soc Exp Biol Med 1972; 140:590-594

31. Buttler RM, Martens F, Ackermans MT, Davison AS, van Herwaarden AE, Kortz L, Krabbe JG, Lentjes EG, Syme C, Webster R, Blankenstein MA, Heijboer AC. Comparison of eight routine unpublished LC-MS/MS methods for the simultaneous measurement of testosterone and androstenedione in serum. Clin Chim Acta 2016; 454:112-118

32. Buttler RM, Martens F, Fanelli F, Pham HT, Kushnir MM, Janssen MJ, Owen L, Taylor AE, Soeborg T, Blankenstein MA, Heijboer AC. Comparison of 7 Published LC-MS/MS Methods for the Simultaneous Measurement of Testosterone, Androstenedione, and Dehydroepiandrosterone in Serum. Clin Chem 2015; 61:1475-1483

33. El-Farhan N, Pickett A, Ducroq D, Bailey C, Mitchem K, Morgan N, Armston A, Jones L, Evans C, Rees DA. Method-specific serum cortisol responses to the adrenocorticotrophin test: comparison of gas chromatography-mass spectrometry and five automated immunoassays. Clin Endocrinol (Oxf) 2013; 78:673-680

34. Groer MW, Humenick S, Hill PD. Characterizations and psychoneuroimmunologic implications of secretory immunoglobulin A and cortisol in preterm and term breast milk. J Perinat Neonatal Nurs 1994; 7:42-51

35. Angelucci L, Patacchioli FR, Scaccianoce S, Di Sciullo A, Cardillo A, Maccari S. A model for later-life effects of perinatal drug exposure: maternal hormone mediation. Neurobehav Toxicol Teratol 1985; 7:511-517

36. Yeh KY, Yeh M, Holt PR. Induction of intestinal differentiation by systemic and not by luminal corticosterone in adrenalectomized rat pups. Endocrinology 1989; 124:1898-1904

37. Cao Y, Rao SD, Phillips TM, Umbach DM, Bernbaum JC, Archer JI, Rogan WJ. Are breast-fed infants more resilient? Feeding method and cortisol in infants. The Journal of pediatrics 2009; 154:452-454

38. Benjamin Neelon SE, Stroo M, Mayhew M, Maselko J, Hoyo C. Correlation between maternal and infant cortisol varies by breastfeeding status. Infant Behav Dev 2015; 40:252-258

39. Mahmood A, Torres-Pinedo R. Effect of hormone administration on the sialylation and fucosylation of intestinal microvillus membranes of suckling rats. Pediatr Res 1985; 19:899-902

40. Teichberg S, Wapnir RA, Moyse J, Lifshitz F. Development of the neonatal rat small intestinal barrier to nonspecific macromolecular absorption. II. Role of dietary corticosterone. Pediatr Res 1992; 32:50-57

41. Axelsson I, Jakobsson I, Lindberg T, Polberger S, Benediktsson B, Raiha N. Macromolecular absorption in preterm and term infants. Acta Paediatr Scand 1989; 78:532-537

42. Nanthakumar NN, Young C, Ko JS, Meng D, Chen J, Buie T, Walker WA. Glucocorticoid responsiveness in developing human intestine: possible role in prevention of necrotizing enterocolitis. Am J Physiol Gastrointest Liver Physiol 2005; 288:G85-92

43. Villa M, Menard D, Semenza G, Mantei N. The expression of lactase enzymatic activity and mRNA in human fetal jejunum. Effect of organ culture and of treatment with hydrocortisone. FEBS Lett 1992; 301:202-206

44. Costalos C, Gounaris A, Sevastiadou S, Hatzistamatiou Z, Theodoraki M, Alexiou EN, Constandellou E. The effect of antenatal corticosteroids on gut peptides of preterm infants--a matched group comparison: corticosteroids and gut development. Early Hum Dev 2003; 74:83-88

45. Watkins JB, Szczepanik P, Gould JB, Klein P, Lester R. Bile salt metabolism in the human premature infant. Preliminary observations of pool size and synthesis rate following prenatal administration of dexamethasone and phenobarbital. Gastroenterology 1975; 69:706-713

46. Cryan JF, Dinan TG. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. Nat Rev Neurosci 2012; 13:701-712

47. Foster JA, McVey Neufeld KA. Gut-brain axis: how the microbiome influences anxiety and depression. Trends Neurosci 2013; 36:305-312

48. Zijlmans MA, Korpela K, Riksen-Walraven JM, de Vos WM, de Weerth C. Maternal prenatal stress is associated with the infant intestinal microbiota. Psychoneuroendocrinology 2015; 53:233-245

49. O'Mahony SM, Marchesi JR, Scully P, Codling C, Ceolho AM, Quigley EM, Cryan JF, Dinan TG. Early life stress alters behavior, immunity, and microbiota in rats: implications for irritable bowel syndrome and psychiatric illnesses. Biol Psychiatry 2009; 65:263-267

50. Gareau MG, Jury J, MacQueen G, Sherman PM, Perdue MH. Probiotic treatment of rat pups normalises corticosterone release and ameliorates colonic dysfunction induced by maternal separation. Gut 2007; 56:1522-1528

51. Moussaoui N, Braniste V, Ait-Belgnaoui A, Gabanou M, Sekkal S, Olier M, Theodorou V, Martin PG, Houdeau E. Changes in intestinal glucocorticoid sensitivity in early life shape the risk of epithelial barrier defect in maternal-deprived rats. PLoS One 2014; 9:e88382

52. Sudo N, Chida Y, Aiba Y, Sonoda J, Oyama N, Yu XN, Kubo C, Koga Y. Postnatal microbial coloniza-tion programs the hypothalamic-pituitary-adrenal system for stress response in mice. J Physiol 2004; 558:263-275

53. Bokkenheuser VD, Winter J, Dehazya P, de Leon O, Kelly WG. Formation and metabolism of tetra-hydrodeoxycorticosterone by human fecal flora. J Steroid Biochem 1976; 7:837-843

54. Feighner SD, Hylemon PB. Characterization of a corticosteroid 21-dehydroxylase from the intesti-nal anaerobic bacterium, Eubacterium lentum. J Lipid Res 1980; 21:585-593

55. Winter J, Bokkenheuser VD. 21-dehydroxylation of corticoids by anaerobic bacteria isolated from human fecal flora. J Steroid Biochem 1978; 9:379-384

56. Winter J, Bokkenheuser VD, Ponticorvo L. Bacterial metabolism of corticoids with particular refer-ence to the 21-dehydroxylation. J Biol Chem 1979; 254:2626-2629

57. Winter J, Cerone-McLernon A, O'Rourke S, Ponticorvo L, Bokkenheuser VD. Formation of 20 beta-dihydrosteroids by anaerobic bacteria. J Steroid Biochem 1982; 17:661-667

58. Sullivan EC, Hinde K, Mendoza SP, Capitanio JP. Cortisol concentrations in the milk of rhesus monkey mothers are associated with confident temperament in sons, but not daughters. Dev Psychobiol 2011; 53:96-104

59. Hinde K, Skibiel AL, Foster AB, Del Rosso L, Mendoza SP, Capitanio JP. Cortisol in mother's milk across lactation reflects maternal life history and predicts infant temperament. Behav Ecol 2015; 26:269-281

60. Dettmer AM, Murphy AM, Guitarra D, Slonecker E, Suomi SJ, Rosenberg KL, Novak MA, Meyer JS, Hinde K. Cortisol in Neonatal Mother's Milk Predicts Later Infant Social and Cognitive Functioning in Rhesus Monkeys. Child Dev 2017;

61. Catalani A, Casolini P, Cigliana G, Scaccianoce S, Consoli C, Cinque C, Zuena AR, Angelucci L. Maternal corticosterone influences behavior, stress response and corticosteroid receptors in the female rat. Pharmacol Biochem Behav 2002; 73:105-114

62. Catalani A, Casolini P, Scaccianoce S, Patacchioli FR, Spinozzi P, Angelucci L. Maternal corticos-terone during lactation permanently affects brain corticosteroid receptors, stress response and behaviour in rat progeny. Neuroscience 2000; 100:319-325

63. Hart S, Boylan LM, Border B, Carroll SR, McGunegle D, Lampe RM. Breast milk levels of cortisol and Secretory Immunoglobulin A (SIgA) differ with maternal mood and infant neuro-behavioral functioning. Infant Behavior & Development 2004; 27:101-106

64. Grey KR, Davis EP, Sandman CA, Glynn LM. Human milk cortisol is associated with infant tempera-ment. Psychoneuroendocrinology 2013; 38:1178-1185

65. Hahn-Holbrook J, Le TB, Chung A, Davis EP, Glynn LM. Cortisol in human milk predicts child BMI. Obesity (Silver Spring) 2016; 24:2471-2474

66. Finken MJJ, van der Voorn B, Hollanders JJ, Dijkstra LR, Toorop AA, Rotteveel J. Cortisol in human milk: The good, the bad, or the ugly? Obesity (Silver Spring) 2017; 25:1153

67. Stone AA, Schwartz JE, Smyth J, Kirschbaum C, Cohen S, Hellhammer D, Grossman S. Individual differences in the diurnal cycle of salivary free cortisol: a replication of flattened cycles for some individuals. Psychoneuroendocrinology 2001; 26:295-306

68. Hackett RA, Kivimaki M, Kumari M, Steptoe A. Diurnal Cortisol Patterns, Future Diabetes, and Impaired Glucose Metabolism in the Whitehall II Cohort Study. J Clin Endocrinol Metab 2016; 101:619-625

69. Kumari M, Shipley M, Stafford M, Kivimaki M. Association of diurnal patterns in salivary cortisol with all-cause and cardiovascular mortality: findings from the Whitehall II study. J Clin Endocrinol Metab 2011; 96:1478-1485

70. Heim C, Ehlert U, Hellhammer DH. The potential role of hypocortisolism in the pathophysiology of stress-related bodily disorders. Psychoneuroendocrinology 2000; 25:1-35

71. Zorn JV, Schur RR, Boks MP, Kahn RS, Joels M, Vinkers CH. Cortisol stress reactivity across psychiatric disorders: A systematic review and meta-analysis. Psychoneuroendocrinology 2017; 77:25-36

72. van der Voorn B, de Waard M, Dijkstra LR, Heijboer AC, Rotteveel J, van Goudoever JB, Finken MJJ. Stability of Cortisol and Cortisone in Human Breast Milk During Holder Pasteurization. J Pediatr Gastroenterol Nutr 2017;

73. Oster H, Challet E, Ott V, Arvat E, Ronald de Kloet E, Dijk DJ, Lightman S, Vgontzas A, Van Cauter E. The Functional and Clinical Significance of the 24-Hour Rhythm of Circulating Glucocorticoids. Endocr Rev 2017; 38:3-45

EVA MOUTON 2019

5

The association between breastmilk glucocorticoid concentrations and macronutrient contents throughout the day

Jonneke J. Hollanders[*],
Stefanie M.P. Kouwenhoven[*],
Bibian van der Voorn,
Johannes B. van Goudoever,
Joost Rotteveel,
Martijn J.J. Finken
[*] Authors contributed equally to this manuscript

ABSTRACT

Background

Glucocorticoids (GCs) in breastmilk follow the maternal hypothalamus-pituitary-adrenal axis activity and may affect the offspring's growth and neurodevelopment. There is some evidence suggesting that macronutrients in breastmilk also fluctuate throughout the day. We aimed to research whether GCs and macronutrients are correlated in multiple breastmilk samples obtained over a 24-h period.

Methods

A total of 10 mothers provided 45 breastmilk samples collected over a 24-h period. Cortisol and cortisone levels were determined by LC-MS/MS, and macronutrients were measured with mid-infrared spectroscopy. Correlations between breastmilk GCs and macronutrients were assessed with Pearson correlations and linear mixed models.

Results

No associations were found between breastmilk GCs and macronutrients (cortisol: β-0.1 (95% confidence interval: -1.0 to 0.7), -4.9 (-12.9 to 3.1) for fat, protein, and carbohydrates, respectively; and -0.3 (-5.6 to 5.0) and cortisone: 0.0 (-2.5 to 2.5), -17.4 (-39.8 to 5.0), and -2.7 (-17.7 to 12.3)) for fat, protein, and carbohydrates, respectively. Adjusting for the time of collection to account for GC rhythmicity did not change the results.

Conclusion

We found no associations between GCs and macronutrients in human breastmilk. The excretion of GCs in breastmilk and the effects of breastmilk GCs on offspring are, therefore, likely independent of the excretion and effects of the macronutrients.

INTRODUCTION

It has been well established that breastmilk is the preferred nutrition for neonates. The composition of breastmilk shows a wide variability between mothers.[1,2] Moreover, the nutritional composition of milk changes throughout the lactation period,[2,3] during the day,[4,5] and within the same feed,[3,5] to meet the requirements of the infant. In short, protein concentrations decrease throughout the lactation period, while lactose and fat remain mostly stable, although some variation in fat concentrations has been observed.[2] Over the day, fat concentrations show a diurnal rhythm, with higher concentrations during the day compared with the morning, evening and night. In contrast, protein and lactose concentrations remain stable throughout the day. Additionally, during a feed, fat concentrations increase depending on the degree of breast emptying,[6] while no effects on lactose or protein [5,7] have been observed. The excretion of milk macronutrients is a complex and active process, including transcellular transport as well as intracellular synthesis.[8]

The beneficial effects of breastmilk result not only from the personalized nutritive composition but also from non-nutritive bioactive factors,[9] such as glucocorticoids (GCs). In breastmilk, unlike serum, cortisone is much more abundant than cortisol, probably due to the expression of 11β-HSD type 2 in the mammary gland, analogous to the salivary gland.[10] Breast-milk GCs have been associated with growth and neurodevelopment in both animal and human studies (as reviewed in [11]). However, the results were contradictory: both a more and less confident temperament was reported with higher cortisol levels in rhesus macaques, while in rats, better stress resilience and less fearfulness was found. In man, cortisol has been associated with autonomic stability as measured with the Neonatal Behavioral Assessment Scale, and with the negative affectivity dimension of the Infant Behavior Questionnaire. Additionally, in rhesus macaques, cortisol has been associated with weight gain in offspring, while in man, an inverse correlation with body mass index (BMI) was found.

Studies of the breastmilk of rhesus monkeys have found a positive correlation between GC levels and milk protein as well as fat concentrations.[12,13] Additionally, although GCs are lipophilic, 70-85% of the GCs measured in a single sample of cow's milk were shown to be associated with the skim fraction of the milk.[14] Although associations with nutritional content have not yet been studied in humans, these findings are confirmed by our liquid chromatography–tandem mass spectrometry (LC–MS/MS) method for breastmilk GC analyses, which showed that the removal of undesired lipids by hexane washings did not influence the milk's GC content.[15] This might be due to the binding of GCs to proteins, such as corticosteroid binding globulin (CBG), CBG-like proteins,[16] and albumin.[17] The excretion of GCs in milk is currently not fully understood, although they are possibly subject to passive diffusion, owing to their lipophilic structure.[11]

We have recently shown that the GCs cortisol and cortisone in breastmilk show a diurnal rhythm, which is correlated to the maternal hypothalamic-pituitary-adrenal (HPA) axis activity,[18] with high concentrations in the early morning which decline to a nadir at night. These findings were replicated by another group.[19] It is currently unknown whether this diurnal rhythm exerts any effects on the offspring. Additionally, it is not known whether this rhythm is correlated with milk macronutrients. If milk GCs appear to be correlated with macronutrients, their effects on the offspring may be (partly or wholly) explained by the nutritional composition of breastmilk. We therefore measured GCs, as well as macronutrients, in samples obtained before every feed over a 24-hour period, to determine whether GC concentrations are correlated with macronutrients. We hypothesized that GC concentrations, due to protein binding, are correlated with milk protein concentrations, but not with other macronutrients.

METHODS

Study population

The subjects in this study were a subset of mothers who participated in the Cortisol in Mother's Milk (CosMos) study, which included mother-infants pairs between March 2016 and July 2017 from both the Onze Lieve Vrouwe Gasthuis (OLVG) and the Amsterdam University Medical Center, location VU University Medical Center (VUmc) in Amsterdam. The primary aim of the study was to research the effect of breastmilk GC rhythmicity on the infant's own HPA-axis activity, behavior, and body composition. Women included at the OLVG were monitored at the Psychiatric Obstetric Pediatric (POP) outpatient-clinic due to previous or current psychopathologic complaints. We included these mothers to obtain a diverse population, because psychopathology is often accompanied by changes in HPA-axis activity.[20,21] Mother-infant pairs were eligible for inclusion if the infant was born at term age (37-42 weeks), with a normal birth weight (-2 to +2 SD score), and if the mothers had the intention of exclusively breastfeeding the infants for ≥3 months. Subjects were excluded from the study for the following reasons: major congenital malformations, multiple pregnancy, pre-eclampsia, maternal alcohol consumption of >7 International Unit (IU)/wk, and/or a fever (temperature >38.5°C) at 1 month postpartum. Medication use other than "over the counter" drugs was also an exclusion criterion, except for anti-depressant use in the mothers included at the OLVG. The approval of the Medical Ethics Committee of the VUmc was obtained (protocol number 2015.524), and written informed consent was obtained from all the participating mothers.

Milk sampling

At 1 month postpartum (±5 days), the mothers were asked to collect a milk sample before every feed during a 24-hour period. The milk was collected using a breast pump or via manual expression. The mothers were asked to collect 1-2 mL of milk. Milk samples were eligible for macronutrient analysis when the mothers collected >10 mL of milk for several or all of the samplings. This resulted in 45 random samples, which were used for this study.

The milk was stored in the mother's freezer until transportation to the laboratory, where it was stored at -20 °C. The samples were thawed twice – once for the GC analyses, and once more for the macronutrient analyses – and were stored at -20 °C in between analyses for several months.

At the same time, the participants were asked to fill in the Hospital Anxiety and Depression Scale (HADS), which is an index to measure clinically relevant anxiety and/or depression symptoms in patients from non-psychiatric hospitals.[22] A score of ≥8 on the depression (HDS) or anxiety (HAS) subscale is indicative of an elevated psychological stress level.

Laboratory

GC concentrations

The total cortisol and cortisone concentrations in the breastmilk were determined by isotope dilution liquid chromatography–tandem mass spectrometry (LC–MS/MS) as previously published.[15] In short, the breastmilk samples were washed 3 times with 2 mL of hexane to remove lipids after adding internal standards (13C3-labeled cortisol and 13C3-labeled cortisone). Then, samples were extracted and analyzed by XLC-MS/MS13, a Symbiosis online SPE system (Spark Holland, Emmen, The Netherlands) coupled to a Quattro Premier XE tandem mass spectrometer (Waters Corp., Milford, MA, USA). The intra-assay coefficients of variation (CV%) were 4 and 5% for cortisol levels of 7 and 23 nmol/L and 5% for cortisone levels of 8 and 33 nmol/L for the LC-MS/MS measurements, while the inter-assay CV% was <9% and the lower limit of quantitation was 0.5 nmol/L for both cortisol and cortisone.

Macronutrient analysis

First, the milk samples were homogenized with an ultrasonic processor (VCX130, Sonics & Materials Inc, Newtown, CT, USA; 98% amplitude for 7.0 s) after heating the sample to 40 °C. Next, the fat, protein, and carbohydrate concentrations (in g/100 mL), as well as the total solids and energy, were analyzed simultaneously with a human milk analyzer (MIRIS, Uppsala, Sweden) through the use of mid-infrared spectroscopy. After 10 samples, the milk analyzer was recalibrated. The milk samples were analyzed in duplicate or, when possible, triplicate (n=4) and were subsequently averaged. The protein concentrations

were expressed as crude or true protein. Crude protein is based on the total amount of nitrogen in the sample, which also includes non-protein nitrogen compounds (20-25% of the total nitrogen in human milk). The true protein levels were calculated by the software as 80% of the crude protein concentrations. For our analyses, the true protein concentrations were used. The energy content (kcal/100 mL) was calculated with the formula: $9.25 \times$ fat $+ 4.40 \times$ crude protein $+ 3.95 \times$ carbohydrate.

Statistics

First, GC, fat, carbohydrate, and protein concentrations were plotted for subjects with >5 samples to visualize the data (n = 4).

Next, the correlations between the milk macronutrients and milk GCs were analyzed with Pearson correlations. Additionally, to account for the repeated measurements within each subject, linear mixed models were performed with cortisol or cortisone concentration as the dependent variable and the macronutrient concentration (fat, protein, or carbohydrates) or total energy as the fixed effect, with the subject added as a random effect. The analyses were repeated while adjusting for the time of collection as a fixed effect.

RESULTS

Study population

A random subset of 10 mothers collected a total of 45 samples with >10 mL of milk; the sampling took place between 26 and 36 days postpartum. None of the mothers used antidepressants, and only one mother had an elevated HAS score (11 points). The mothers gave birth to 5 boys and 5 girls, with a gestational age between 37+1 and 41+2 weeks and a birth weight between 2726 and 4030 g. A total of 5 mothers delivered their babies via caesarian section.

Individual plots

Figure 1 shows the individual plots for 4 mothers, who had >5 samples of >10 mL of breastmilk. While in all the mothers a diurnal rhythm of cortisol and cortisone can be seen, no rhythm appeared to be present for fat, carbohydrates, and protein.

Correlations

The cortisol and cortisone concentrations were highly correlated (Table 1; p<0.001), while the GC concentrations were not correlated with fat, protein, or carbohydrate milk contents of the milk or with the total solids or milk energy.
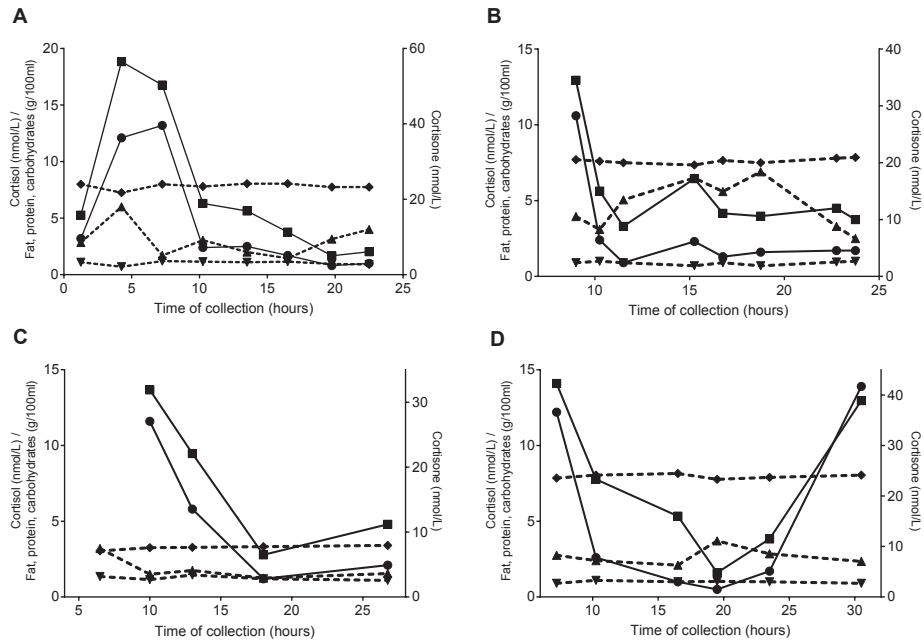
Figure 1: Panels A-D represent the individual plots for milk cortisol (●), cortisone (■), fat (▲), protein (▼), and carbohydrate (◆) concentrations for four mothers.

Left y-axis: cortisol concentrations (nmol/L) and milk macronutrient concentrations (g/100 mL); right y-axis: cortisone concentrations (nmol/L); x-axis: time of collection (hours); continuous lines: GC concentrations; and dotted lines: macronutrient concentrations

Table 1: Correlations between milk glucocorticoids and macronutrients

	Cortisol	Cortisone	Fat	Protein	Carbohydrates	Total solids	Energy
Cortisol	1						
Cortisone	0.827[*]	1					
Fat	-0.096	-0.050	1				
Protein	-0.071	-0.119	-0.455[*]	1			
Carbohydrates	-0.016	-0.050	-0.639[*]	-0.090	1		
Total solids	-0.127	-0.088	-0.988[*]	-0.404[*]	-0.564[*]	1	
Energy	-0.109	-0.068	-0.998[*]	-0.432[*]	-0.615[*]	0.995[*]	1

Values represent Pearson's r; [*] P-value <0.001

The fat concentrations, carbohydrate concentrations, total solids, and energy were highly correlated. The protein levels were correlated with fat concentrations, total solids, and energy, but not with the carbohydrate concentrations.

Linear Mixed Models

No significant associations were found between the milk GC concentrations and the milk macronutrients, total solids, or milk energy (Table 2). Adjusting for the time of collection did not alter these results.

Using crude protein concentrations instead of true protein levels did not change the results (data not shown).

Table 2: Associations between milk glucocorticoids and macronutrients

	Cortisol				Cortisone			
	Crude analyses		**Adjusted for time of collection**		**Crude analyses**		**Adjusted for time of collection**	
	β (95% CI)	P value	β (95% CI)	P value	β (95% CI)	P value	β (95% CI)	P value
Fat	-0.3 (-1.3 to 0.7)	0.54	-0.1 (-1.0 to 0.7)	0.75	-0.4 (-3.2 to 2.3)	0.75	0.0 (-2.5 to 2.5)	0.998
Protein	-2.0 (-11.0 to 6.9)	0.65	-4.9 (-12.9 to 3.1)	0.22	-9.6 (-34.6 to 15.4)	0.44	-17.4 (-39.8 to 5.0)	0.13
Carbohydrates	-0.3 (-6.3 to 5.7)	0.92	-0.3 (-5.6 to 5.0)	0.90	-2.7 (-19.4 to 14.1)	0.75	-2.7 (-17.7 to 12.3)	0.72
Total solids	-0.5 (-1.7 to 0.7)	0.41	-0.3 (-1.4 to 0.8)	0.55	-1.0 (-4.3 to 2.4)	0.57	-0.5 (-3.6 to 2.6)	0.74
Energy	0.0 (-0.2 to 0.1)	0.48	0.0 (-0.1 to 0.1)	0.66	-0.1 (-0.4 to 0.3)	0.66	0.0 (-0.3 to 0.3)	0.88

Values represents β (95% CI), analyzed with linear mixed models.

DISCUSSION

To the best of our knowledge, this study is the first to assess the correlations between the diurnal pattern of cortisol and cortisone in breastmilk and the milk macronutrients fat, protein and carbohydrates, as well as energy in humans. We did not find any correlations between breastmilk GCs and macronutrients, even after taking the diurnal rhythm of breastmilk GCs into account. The excretion of GCs in breastmilk and the effects of GCs on offspring are, therefore, likely to be independent of macronutrients.

Previously, studies in rhesus macaques showed correlations between milk glucocorticoids and both the fat and protein contents of milk [12,13] This discrepancy with our study might be due to several reasons. First, both Sullivan et al. [13] and Hinde et al. [12] emptied the mammary gland entirely, while for ethical reasons, the mothers in our study collected only a small portion of milk before feeding their infants. Although previous research has shown that GC concentrations are similar in fore- and hindmilk [23] and that lactose and protein levels also do not change during a feeding, [5,7] fat levels do increase during a feed. [6] Additionally, both studies only took one sample of milk per subject, and only Hinde et al. (2015) [12] minimized the time window of the sample col-

lection (between 11:30-13:00, after 3.5-4 h of milk accumulation). This sampling period most likely did not measure the peak cortisol levels, because cortisol concentrations in rhesus macaques reach maximum levels around 8:00,[24] and the GC concentrations are, therefore, likely to be less variable in those studies as compared with ours. However, when we reanalyzed our data with only the samples collected after 11:30, we still found no association between milk GCs and macronutrients [data not shown]. Schwalm et al. (1978)[14] researched cow's milk and concluded that no positive correlation was found between milk glucocorticoids and fat, while there was also no correlation found with milk protein. No previous study has researched the correlations between milk GCs and macronutrients in human milk.

We did not find a correlation between GCs and milk protein, as hypothesized. Although 70-85% of GCs in cow's milk are associated with the skim fraction of the milk,[14] the concentration of GCs in milk is low. Moreover, only a small proportion of the protein fraction in milk harbors CBG, CBG-like proteins, or albumin.[16,25] It is, therefore, still possible that GCs are correlated with specific proteins, but not with the total protein concentration. Nevertheless, CBG-binding activity in milk was also possibly found to exhibit a diurnal rhythm,[26] with a peak in the evening, which did not coincide with the GC peak. This, however, should not have hampered our GC measurements, as our laboratory method assesses total GC concentrations.[11] Lastly, it is unlikely that GCs could be correlated with fat concentrations, because the removal of undesired lipids by hexane washings during our LC-MS/MS method did not influence the breastmilk's GC content,[15] although the free fraction of breastmilk GCs could hypothetically be correlated with the lipid fraction.

Our findings could have several implications. First, it appears that the excretion of GCs in milk is not dependent on the macronutrients and vice versa. Indeed, due to the lipophilic structure of GCs, they are probably subject to passive diffusion from the systemic circulation into the breastmilk. This assumption was corroborated by our previous research, showing that GCs in breastmilk closely follow the maternal HPA-axis activity.[18] In contrast, the secretion of macronutrients into breastmilk is an active and complex process, with both transcellular transport and intracellular synthesis. Proteins and lactose are transported in secretory vehicles and secreted via exocytosis, while lipids are released in milk fat globules.[8] The lack of association between breast-milk macronutrients and GCs may suggest that the latter do not influence these transport mechanisms, at least not acutely. Second, several studies have been published recently suggesting that the GCs in breastmilk affect neurodevelopment and growth in humans.[27-29] Because milk GCs and macronutrients are not associated, it is unlikely that these effects could have been falsely attributed to GCs. However, the underlying mechanism is not yet understood fully, but might be due to a direct influence of breastmilk GCs on the infant's HPA-axis activity and development, or through an interaction with the gut microbiome (i.e., the "gut-brain axis hypothesis") .[11,30,31] Therefore, further research to assess these

associations is still warranted, especially because these studies did not take the diurnal rhythm of GCs in breastmilk into account.

Our study has several strengths and limitations. First, this is the first study to research the correlation between GCs and macronutrients in human milk. Additionally, the diurnal rhythm of GCs in breastmilk was taken into account by sampling multiple times over the day, allowing for a more precise analysis of the correlation between breastmilk GCs and macronutrients. However, our study also had limitations. Because mothers were asked to collect milk prior to each feeding, only foremilk was collected. Preferably, a portion of a completely pumped feed would be analyzed, which would also include the hindmilk. Next, the CosMos study aimed to research the association between milk GCs and effects on offspring. The current study used leftover samples from mothers who collected >10 mL of milk, rather than the requested 1-2 mL. The sample size of this study was therefore limited. Moreover, the samples were thawed twice – once for the GC analyses and once more for the macronutrient analyses – and were stored at -20 °C in between analyses for several months. While GCs in milk are highly stable,[15] macronutrients are probably more vulnerable to degradation, with studies showing a decrease in fat concentrations of up to 10%, while other macronutrients remained stable or showed minimal degradation (up to 2%); these effects were already observed after 2-7 days of freezing.[32-37] It is therefore possible that the macronutrient concentrations were initially different than what was measured during the analyses. However, this would lead to a systematic error rather than a random error, and it is therefore likely that the correlations would still be absent even if the macronutrient analyses had been performed with fresh breastmilk. The results of our study should, therefore, be interpreted with caution due to all of these factors.

In conclusion, no associations between GCs and macronutrients in human milk were found in this study. The excretion of GCs in breastmilk and the effects of breastmilk GCs on the offspring are, therefore, likely independent of the excretion and effects of the macronutrients.

REFERENCES

1. Gidrewicz DA, Fenton TR. A systematic review and meta-analysis of the nutrient content of pre-term and term breast milk. BMC Pediatr 2014; 14:216

2. Mitoulas LR, Kent JC, Cox DB, Owens RA, Sherriff JL, Hartmann PE. Variation in fat, lactose and protein in human milk over 24 h and throughout the first year of lactation. Br J Nutr 2002; 88:29-37

3. Saarela T, Kokkonen J, Koivisto M. Macronutrient and energy contents of human milk fractions during the first six months of lactation. Acta Paediatr 2005; 94:1176-1181

4. Gunther M, Stanier JE. Diurnal variation in the fat content of breast-milk. Lancet 1949; 2:235-237

5. Khan S, Hepworth AR, Prime DK, Lai CT, Trengove NJ, Hartmann PE. Variation in fat, lactose, and protein composition in breast milk over 24 hours: associations with infant feeding patterns. J Hum Lact 2013; 29:81-89

6. Daly SE, Di Rosso A, Owens RA, Hartmann PE. Degree of breast emptying explains changes in the fat content, but not fatty acid composition, of human milk. Exp Physiol 1993; 78:741-755

7. Thurl S, Henker J, Taut H, Tovar K, Sawatzki G. Variations of neutral oligosaccharides and lactose in human milk during the feeding. Z Ernahrungswiss 1993; 32:262-269

8. Truchet S, Honvo-Houeto E. Physiology of milk secretion. Best Pract Res Clin Endocrinol Metab 2017; 31:367-384

9. Bernt KM, Walker Wa. Human milk as a carrier of biochemical messages. Acta paediatrica Suppl 1999; 88:27-41

10. Smith RE, Maguire JA, Stein-Oakley AN, Sasano H, Takahashi K, Fukushima K, Krozowski ZS. Localization of 11 beta-hydroxysteroid dehydrogenase type II in human epithelial tissues. J Clin Endocrinol Metab 1996; 81:3244-3248

11. Hollanders JJ, Heijboer AC, van der Voorn B, Rotteveel J, Finken MJJ. Nutritional programming by glucocorticoids in breast milk: Targets, mechanisms and possible implications. Best Pract Res Clin Endocrinol Metab 2017; 31:397-408

12. Hinde K, Skibiel AL, Foster AB, Del Rosso L, Mendoza SP, Capitanio JP. Cortisol in mother's milk across lactation reflects maternal life history and predicts infant temperament. Behav Ecol 2015; 26:269-281

13. Sullivan EC, Hinde K, Mendoza SP, Capitanio JP. Cortisol concentrations in the milk of rhesus monkey mothers are associated with confident temperament in sons, but not daughters. Dev Psychobiol 2011; 53:96-104

14. Schwalm JW, Tucker HA. Glucocorticoids in mammary secretions and blood serum during reproduction and lactation and distributions of glucocorticoids, progesterone, and estrogens in fractions of milk. J Dairy Sci 1978; 61:550-560

15. van der Voorn B, Martens F, Peppelman NS, Rotteveel J, Blankenstein MA, Finken MJ, Heijboer AC. Determination of cortisol and cortisone in human mother's milk. Clin Chim Acta 2015; 444:154-155

16. Payne DW, Peng LH, Pearlman WH. Corticosteroid-binding proteins in human colostrum and milk and rat milk. J Biol Chem 1976; 251:5272-5279

17. Nguyen PT, Lewis JG, Sneyd J, Lee RS, Torpy DJ, Shorten PR. Development of a formula for estimating plasma free cortisol concentration from a measured total cortisol concentration when elastase-cleaved and intact corticosteroid binding globulin coexist. J Steroid Biochem Mol Biol 2014; 141:16-25

18. van der Voorn B, de Waard M, van Goudoever JB, Rotteveel J, Heijboer AC, Finken MJ. Breast-Milk Cortisol and Cortisone Concentrations Follow the Diurnal Rhythm of Maternal Hypothalamus-Pituitary-Adrenal Axis Activity. J Nutr 2016; 146:2174-2179

19. Pundir S, Wall CR, Mitchell CJ, Thorstensen EB, Lai CT, Geddes DT, Cameron-Smith D. Variation of Human Milk Glucocorticoids over 24 hour Period. J Mammary Gland Biol Neoplasia 2017; 22:85-92

20. Jarcho MR, Slavich GM, Tylova-Stein H, Wolkowitz OM, Burke HM. Dysregulated diurnal cortisol pattern is associated with glucocorticoid resistance in women with major depressive disorder. Biol Psychol 2013; 93:150-158

21. Vreeburg SA, Hoogendijk WJ, DeRijk RH, van Dyck R, Smit JH, Zitman FG, Penninx BW. Salivary cortisol levels and the 2-year course of depressive and anxiety disorders. Psychoneuroendocrinology 2013; 38:1494-1502

22. Herrmann C. International experiences with the Hospital Anxiety and Depression Scale--a review of validation data and clinical results. J Psychosom Res 1997; 42:17-41

23. Patacchioli F, Cigliana G. Maternal plasma and milk free cortisol during the first 3 days of breast-feeding following spontaneous delivery or elective cesarean section. Gynecolog Obstet Invest 1992; 34:159-163

24. Umberkoman-Wiita B, Hansen S, Herbert J, Moore GF. Circadian rhythms in serum and CSF cortisol of rhesus monkeys, and their modulation by timed injections of L-5-hydroxytryptophan. Brain Res 1981; 222:235-252

25. Nagasawa T, Kiyosawa I, Fukuwatari Y, Kitayama T, Uechi M. Alpha-lactalbumin and serum albumin in human milk. J Dairy Sci 1973; 56:177-180

26. Agrimonti F, Frairia R, Fornaro D, Torta M, Borretta G, Trapani G, Bertino E, Angeli A. Circadian and circaseptan rhythmicities in corticosteroid-binding globulin (CBG) binding activity of human milk. Chronobiologia 1982; 9:281-290

27. Grey KR, Davis EP, Sandman CA, Glynn LM. Human milk cortisol is associated with infant temperament. Psychoneuroendocrinology 2013; 38:1178-1185

28. Hahn-Holbrook J, Le TB, Chung A, Davis EP, Glynn LM. Cortisol in human milk predicts child BMI. Obesity (Silver Spring) 2016; 24:2471-2474

29. Hart S, Boylan LM, Border B, Carroll SR, McGunegle D, Lampe RM. Breast milk levels of cortisol and Secretory Immunoglobulin A (SIgA) differ with maternal mood and infant neuro-behavioral functioning. Infant Behav Dev 2004; 27:101-106

30. Cryan JF, Dinan TG. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. Nat Rev Neurosci 2012; 13:701-712

31. Foster JA, McVey Neufeld KA. Gut-brain axis: how the microbiome influences anxiety and depression. Trends Neurosci 2013; 36:305-312

32. Berkow SE, Freed LM, Hamosh M, Bitman J, Wood DL, Happ B, Hamosh P. Lipases and Lipids in Human-Milk - Effect of Freeze-Thawing and Storage. Pediatr Res 1984; 18:1257-1262

33. Garcia-Lara NR, Escuder-Vieco D, Garcia-Algar O, De la Cruz J, Lora D, Pallas-Alonso C. Effect of Freezing Time on Macronutrients and Energy Content of Breastmilk. Breastfeed Med 2012; 7:295-301

34. Ahrabi AF, Handa D, Codipilly CN, Shah S, Williams JE, McGuire MA, Potak D, Aharon GG, Schanler RJ. Effects of Extended Freezer Storage on the Integrity of Human Milk. J Pediatr 2016; 177:140-143

35. Peters MD, McArthur A, Munn Z. Safe management of expressed breast milk: A systematic review. Women Birth 2016; 29:473-481

36.   de Waard M, Mank E, van Dijk K, Schoonderwoerd A, van Goudoever JB. Holder-Pasteurized Human Donor Milk: How Long Can It Be Preserved? J Pediatr Gastroenterol Nutr 2018; 66:479-483

37.   Chang YC, Chen CH, Lin MC. The macronutrients in human milk change after storage in various containers. Pediatr Neonatol 2012; 53:205-209

6

Biphasic glucocorticoid rhythm in one month old infants: reflection of a developing HPA-axis?

Jonneke J. Hollanders,
Bibian van der Voorn,
Paul de Goede,
Alyssa A. Toorop,
Lisette R. Dijkstra,
Adriaan Honig,
Joost Rotteveel,
Koert M. Dolman,
Andries Kalsbeek,
Martijn J.J. Finken

ABSTRACT

Context
The hypothalamus-pituitary-adrenal (HPA) axis displays a diurnal rhythm. However, little is known about its development in early life.

Objective
To describe HPA-axis activity and study possible influencing factors in 1-month-old infants.

Design
Observational.

Setting
Amsterdam UMC, location VUMC, and OLVG, Amsterdam.

Participants
Fifty-five mother-infant pairs.

Interventions
Collection of breastmilk and infants' saliva 1 month postpartum for analysis of glucocorticoids (GCs; i.e., cortisol and cortisone) using LC-MS/MS.

Main outcome measure
GC rhythm in infants' saliva, and associations with vulnerability for maternal psychological distress (increased Hospital Anxiety and Depression Scale (HADS) score or consultation at the Psychiatric Obstetric Pediatric (POP) clinic), season at sampling, sex and breastmilk GC rhythmicity, analyzed with Sigmaplot and regression analyses.

Results
A significant biphasic GC rhythm was detected in infants, with peaks at 6:53±1:01 (mean±SEM) and 18:36±1:49 for cortisol, and at 8:50±1:11 and 19:57±1:13 for cortisone. HADS-score, POP-consultation, season at sampling and sex were not associated with the infants' GC rhythm. Breastmilk cortisol maximum was positively associated with infants' cortisol area-under-the-curve (AUC) increase and maximum. Higher breastmilk cortisone AUCincrease, AUCground and maximum were associated with an earlier maximum in infants. Breastmilk and infant GC concentrations were associated between 6:00-9:00.

Conclusions

A biphasic GC rhythm, peaking in the morning and evening, was seen in 1-month-old infants at a group level. Breastmilk GC parameters might be associated with the infants' GC rhythm, possibly caused by a signaling effect of breastmilk GCs, or as an associative effect of increased mother-infant synchrony. These results contribute to an increased understanding of early-life HPA-axis development.

INTRODUCTION

In adults, the hypothalamus pituitary adrenal (HPA) axis displays a diurnal rhythm, peaking in the morning and with a nadir at night. However, it is not exactly clear when this adult-type rhythm is established in children, with studies reporting ages ranging from 2 weeks to 9 months in healthy infants.[1-9]

A rhythm in HPA-axis activity might already be present in the human fetus. Term neonates born in the afternoon through elective caesarian section appeared to have increased cortisol concentrations compared to neonates born during other times throughout the day.[10] Additionally, maternal estriol levels, partly reflecting dehydroandrostenedione (DHEAS) production by the fetal zone of the adrenal cortex, display a 24-hour rhythm during pregnancy inversely related to maternal cortisol levels.[11] Furthermore, several studies have shown data suggesting that a diurnal glucocorticoid (GC) rhythm is present from birth onward. Iwata et al. (2013) [12] described a diurnal cortisol rhythm peaking in the afternoon in newborns 2-11 days postpartum, while Spangler (1991) [8] found a biphasic pattern in neonates 2-7 days postpartum.

It is conceivable that an HPA-axis rhythm emerges prenatally, and continues to develop into an adult-type rhythm after birth, with a shift from a peak in the afternoon towards a morning peak. Currently, it is not clear which factors drive the development of an adult-type diurnal rhythm of the HPA-axis. Diurnal rhythms in general are mostly regulated by the suprachiasmatic nuclei (SCN), located in the anterior hypothalamus,[13,14] and its entrainment is predominantly dependent on exogenous time cues.[15] Indeed, light-dark cycles are an important regulator of SCN rhythmicity,[16] and have also been shown to influence infants' activity levels.[17] Moreover, maternal depressive disorders prior to or during pregnancy were associated with sleeping problems in infants,[18] while in adults psychopathology has been linked to changes in HPA-axis activity.[19,20] A twin study has previously concluded that environmental factors outweigh the genetic contribution to the development of an HPA-axis rhythm.[9] However, which exogenous factors influence the development of an HPA-axis rhythm has not been studied yet.

Maternal activity has been associated with infant activity independent of exposure to light,[21] while formula milk with day/night nutrient levels in synchrony with the environment, appeared to affect sleep patterns in infants.[22] Breastfeeding mothers have been shown to exhibit more touching and gazing behavior towards their infants, suggestive of more interactive behavior,[23] and these associations appear to be partly influenced by infant sex.[24] Breastmilk as well as breastfeeding itself might therefore also act as a possible contributor to the development of an HPA-axis rhythm. Additionally, breastmilk itself contains components which might aid in the development of an adult-type GC rhythm. For example, melatonin exhibits a strong diurnal pattern in breastmilk.[25] Similarly, our research group has previously shown that a diurnal rhythm of cortisol and cortisone is

present in breastmilk, mirroring maternal HPA-axis activity.[26] In rats GCs were able to cross the intestinal epithelial barrier,[27] and earlier research in humans has shown that serum cortisol levels were 40% higher in infants who were breastfed.[28] Moreover, cortisol levels in maternal and infant saliva were significantly correlated in breastfed infants, but not in formula-fed infants.[29] Accordingly, glucocorticoids in breastmilk might influence the process of HPA-axis rhythm development.

We therefore performed an exploratory study, aimed at assessing how some exogenous factors are associated with GC rhythmicity in infants, with a focus on the association with GC rhythmicity in breastmilk. Mother with a medical history of psychopathology were oversampled in an attempt to increase the range of maternal HPA-axis activity, since depression and anxiety have previously been shown to impact the GC circadian rhythm.[19,20] GC levels in the infants and breastmilk were sampled at one month postpartum, since HPA-axis development into an adult-type rhythm appears to still be in progress at that time-point, while intra-uterine influences are likely to have disappeared. Both cortisol and cortisone were determined, since cortisone levels are higher compared to cortisol levels in both saliva and breastmilk, probably due to local conversion by 11β-hydroxysteroid dehydrogenase type 2 (11β-HSD2),[30] and are therefore less likely to have a concentration below the lower limit of detection. Moreover, cortisone seems to be a more reliable biomarker compared to cortisol, at least in saliva and hair.[31,32] GCs in breastmilk, season at time of sampling, maternal psychopathology and infant sex were explored as possible influencing factors, with use of specialized rhythm analysis software.

METHODS

Study population

Between March 2016 and July 2017, mother-infant pairs were included from the general hospital OLVG as well as the academic Amsterdam UMC, location VUMC, both located in Amsterdam, as part of the Cortisol in Mother's Milk (CosMos) study. The primary aim of the CosMos study was to research the associations between breastmilk GC rhythmicity and the infant's own HPA-axis activity, behavior, and body composition. Women included at the OLVG were recruited at the Psychiatric Obstetric Pediatric (POP) outpatient clinic where they were monitored because of an increased risk for psychopathologic complaints. Inclusion criteria were: 1) born at term age (37-42 weeks), 2) normal birth weight (-2 to +2 SD), and 3) the intention to exclusively breastfeed for ≥ 3 months. Mother-infant pairs were excluded due to 1) major congenital anomalies, 2) multiple pregnancy, 3) pre-eclampsia or HELLP, 4) maternal alcohol consumption of >7 IU/week and/or 5) a fever (temperature >38.5°C) at time of GC sampling. Additionally, mothers

were also excluded if they used medication other than "over the counter" drugs, except for anti-depressant use in the mother-infants pairs included at the OLVG.

Approval of the Medical Ethics Committee of the VUMC was obtained (protocol number 2015.524), and written informed consent was obtained from all participating mothers.

Data collection

Peripartum

Shortly after inclusion, within the first week postpartum, mothers filled in a questionnaire pertaining to their pregnancy and birth, as well as anthropometric and demographic data.

One month postpartum

One month postpartum (±5 days), mothers collected a portion of breastmilk (1-2 ml) before every feeding moment during a 24-hour period, with the use of a breast pump or via manual expression. In order to minimize intra-individual differences, we requested mothers to use the same method for all their samples. Simultaneously, before feeding, they also collected their infant's saliva, using a SalivaBio Infant's Swab (exclusively from Salimetrics, State College, PA).

Milk and saliva was stored in the mother's freezer, and subsequently in the laboratory at -20°C for less than 3 months prior to analysis.

At time of sampling, mothers were also asked to fill in the Hospital Anxiety and Depression Scale (HADS) questionnaire, which assessed self-reported levels of depression and anxiety symptoms.[33,34] It contains 14 questions, with seven questions concerning depressive symptoms (hospital depressive subscale; HDS) and seven anxiety symptoms (hospital anxiety subscale; HAS). Items are scored 0-3, and a score of ≥8 on one of the two subscales (HDS/HAS) is indicative of clinically relevant depression and/or anxiety symptoms.

Laboratory

Total cortisol and cortisone concentrations in breast milk were determined by isotope dilution liquid chromatography–tandem mass spectrometry (LC–MS/MS) as previously published.[35] In short, hexane washing was done thrice to remove lipids, after adding internal standards (13C3-labeled cortisol and 13C3-labeled cortisone). Then, samples were extracted using Isolute plates (Biotage, Uppsala, Sweden) and analyzed by LC-MS/MS (Acquity with Quattro Premier XE, Milford MA, USA, Waters Corporation). The intra-assay coefficients of variation (CV%) were 4 and 5% for cortisol levels of 7 and 23 nmol/L, and 5% for cortisone levels of 8 and 33 nmol/L for LC-MS/MS measurements, while the inter-assay CV% was <9% and the Lower Limit of Quantitation was 0.5 nmol/L for both

cortisol and cortisone. Cortisol and cortisone concentrations in saliva were determined with the same method as that for breast milk, but without the hexane-washing procedure.

Statistics

Regression analyses

In total, 63 mother-infant pairs were included, of whom 55 pairs had valid GC levels for both mother and infant. Due to extremely high GC levels, one infant was excluded for cortisol analyses (n=54), and another infant was excluded for cortisone analyses (n=54).

GC levels were visualized by calculating mean (95% confidence interval (CI)) in 2-hour time windows. Additionally, linear mixed models (LMMs), which allow correcting for intra-individual measurements, were used to determine the slope of the increasing (i.e., 00:00-7:00) and decreasing (i.e., 7:01-23:59) part of the diurnal rhythm, in line with our previous study.[26]

Next, linear regression analyses were performed, for which seven additional mother-infant pairs were excluded because total sampling time was <8 hours and/or no samples were collected between 5:00-10:00 (i.e., sample collection around the morning peak), because this could interfere with the interpretation of the rhythm parameters. This resulted in 47 included mother-infant pairs.

Infant saliva and breastmilk cortisol and cortisone data were converted into rhythm parameters which, when taken into account together, will allow a full overview of HPA-axis rhythmicity:[36]

1. Area Under the Curve (AUC) with respect to the ground (AUCg) as well as increase (AUCi) were calculated by using the trapezoid rule as described by Pruessner et al. (2003).[37] AUC calculations were corrected for the total sampling time, since this differed between mothers. AUCg provides information on total GC exposure over the sampling time, while the AUCi is a measure of GC variability.
2. The maximum concentration was determined and used as a proxy for peak concentrations.
3. The time at which maximum concentration was measured was used as a proxy for time of peak.

Associations between infant saliva rhythm parameters and increased HDS/HAS score, consultation at the POP outpatient clinic (as a proxy for vulnerability for maternal psychological distress), season at time of sampling (divided into two 4-month windows: 21/4 to 21/8 (summer) and 21/10 to 21/2 (winter); used as a parameter for light-dark exposure), and sex were analyzed. Season at time of sampling analyses were repeated with time windows of 3 and 6 months, as well as by determining season at birth, divided into 3- , 4- and 6-month windows.

Additionally, the associations between breastmilk and infant saliva rhythm parameters were determined.

Lastly, the associations between maternal and infant raw GC levels, split up in 3 hour time intervals, were analyzed (n=54), by using LMMs.

Interactions with POP-clinic attendance were tested. No effect modification was found, and the data were therefore not stratified. Instead, as sensitivity analyses, analyses assessing the possible influencing factors were repeated while only including mothers who did not attend the POP-clinic (n=40).

Sigmaplot analyses

Daily rhythmicity of cortisone and cortisol in the infants' saliva were assessed using Gaussian peak regression with Sigmaplot 14.0 software (Systat Software, San Jose, CA, USA). The data were best fitted (i.e., most optimal *P* value and adjusted R-squared, least residuals and dependent on the least amount of variables) to the following regression formula, after testing single, double and triple peak formulas (Supplementary Table 1): $y = a_1*exp(-.5*((x-x_1)/b_1)^2) + a_2*exp(-.5*((x-x_2)/b_2)^2)$, where a_1 and a_2 represent the estimates for the first and second peak heights respectively; b_1 and b_2 represent the estimates for the full width at half maximum of the first and second peak respectively (i.e. a measure of the broadness of the peak); x_1 and x_2 represents the estimates for the location (i.e. the timing along the 24h cycle) of the first peak and second peak respectively. Intra-individual values were taken into account and grouped together through using the "shared parameters" function for the regressions.

GC rhythmicity was assessed separately for the following possible influencing factors: HADS-score (HDS and/or HAS < or ≥8), POP-clinic consultation (yes/no), season at sampling (21/4 to 21/8 and 21/10 to 21/2), sex (male/female), breastmilk AUCi (< or > p50), breastmilk AUCg (< or > p50). Subsequently, T-tests were used to calculated *P* values for differences in the timing of peaks (x_1 and x_2). The estimates a and b were not compared since results were often found to be unreliable, in contrast to the x-estimate.

RESULTS

Population description

Table 1 shows the population characteristics. Increased HDS and/or HAS scores were found in 10 mothers, half of whom were included at the Amsterdam UMC, location VUMC. Fifteen mothers were included at the POP-clinic, of whom 33.3% had clinically relevant depression (HDS) and/or anxiety (HAS) symptoms.

Table 1: Perinatal and maternal characteristics of the study population (n=55)

	n (%) or mean±SD
Gestational age (weeks)	39.7±1.3
Birth weight	
- grams	3550±467
- SDS	0.2±0.9
Male sex	35 (63.6)
HAS/HDS >8	10 (18.2)
- Amsterdam UMC	5 (12.5)
- OLVG hospitals, POP clinic	5 (33.3)
Consulted POP outpatient clinic	15 (27.3)
Season of birth:	
- between 4/21 and 8/21	17 (30.9)
- between 10/21 and 2/21	14 (25.5)

Infant GC rhythm

Regression analyses

Figures 1A and 1B show the infants' and breastmilk cortisol and cortisone levels over the day. A clear diurnal rhythm can be distinguished for both infant salivary and breastmilk GC levels. LMM analyses revealed that infant salivary as well as breastmilk GC levels significantly increased between 00:00-7:00 and significantly decreased between 7:01-23:59 [all P values <0.003].

Sigmaplot analyses

A significant biphasic cortisol and cortisone rhythm could be detected in the infants' saliva. The P values for overall fit as well as the placement of the peaks were $P<0.0001$ for both cortisol and cortisone. For cortisol, the first peak occurred at 6:53±1:01 (mean±SEM) and the second peak at 18:36±1:49. For cortisone, the first peak occurred at 8:50±1:11 and the second peak at 19:57±1:13. Analyses at an individual level could not be performed due to data constraints. It could therefore not be ruled out whether results represent a true biphasic rhythm in infants, or two separate groups of infants with a single morning or evening peak. Figure 1C and 1D show the cortisone rhythm for two individual infants, one with an adult-type rhythm (1C), and the other with a clear biphasic rhythm (1D).

Rhythm influencing factors

Regression analyses

Table 2 shows the associations between rhythm parameters and possible influencing factors. Male sex was associated with a lower salivary cortisol AUCi and a lower salivary
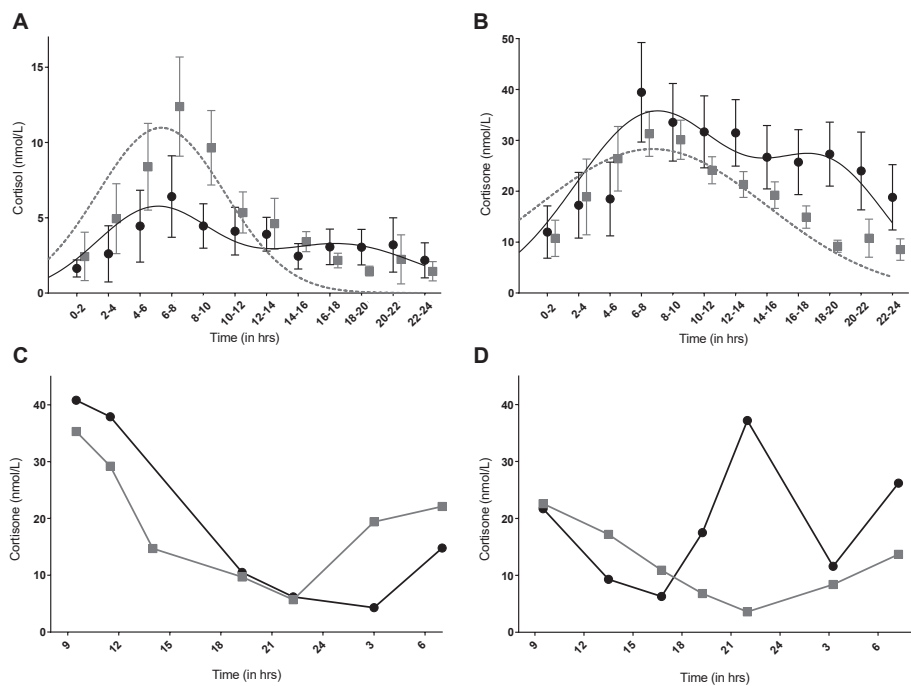
Figure 1: Cortisol (A) and cortisone (B, C & D) rhythms at a group level (A & B) and for two individuals (C & D), in infant's saliva (●) and breastmilk (■). The formula for cortisol (A) and cortisone (B) rhythms in infant's saliva as calculated by Sigmaplot is plotted as a continuous black line, the rhythm in breastmilk is plotted as a dotted grey line.

maximum cortisol concentration. However, these associations were not present when salivary cortisone rhythm parameters were analyzed. No associations were found between rhythm parameters and other factors. Repeating season at time of sampling analyses with the other time windows did not reveal any associations either [data not shown].

Breastmilk maximum cortisol levels were positively associated with salivary cortisol AUCi and maximum levels in the infant (Table 3). Additionally, higher breastmilk cortisone AUCi, AUCg and maximum concentrations were associated with an earlier time of salivary maximum cortisone in the infant. No other associations were found.

Table 4 shows the associations between raw data of breastmilk and infant salivary GC concentrations, divided into 3-hour time intervals. A positive association was found for both cortisol and cortisone between 6:00-9:00, while no associations were found in the other time windows.

Repeated analyses with Ln-transformed GC levels found similar results [data not shown].

Table 2: Associations between GC rhythm parameters in the infants' saliva and possible influencing factors

		AUCi		AUCg		Maximum		Time of maximum	
	n	β (95%CI)	P	β (95%CI)	P	β (95%CI)	P	β (95%CI)	P
Cortisol									
Increased HADS-score	47	0.3 (-1.1 to 1.6)	0.68	0.1 (-1.5 to 1.7)	0.90	1.9 (-3.3 to 7.1)	0.47	-1.1 (-5.3 to 3.0)	0.59
POP-clinic consultation	46	-0.7 (-1.9 to 0.6)	0.28	-0.7 (-2.2 to 0.7)	0.31	-1.0 (-5.7 to 3.7)	0.67	-0.3 (-4.1 to 3.5)	0.88
Season at sampling	29	-0.2 (-1.6 to 1.2)	0.75	-0.4 (-2.1 to 1.3)	0.62	-1.6 (-6.8 to 3.5)	0.52	-2.8 (-7.2 to 1.5)	0.19
Male sex	47	-1.2 (-2.2 to -0.2)	0.02	-1.1 (-2.3 to 0.1)	0.07	-6.4 (-10.0 to -2.9)	0.001	-0.4 (-3.6 to 2.9)	0.83
Cortisone									
Increased HADS-score	47	4.9 (-0.7 to 10.5)	0.08	3.7 (-3.8 to 11.2)	0.32	3.5 (-12.8 to 19.9)	0.67	-0.2 (-4.1 to 3.8)	0.94
POP-clinic consultation	46	-1.4 (-6.7 to 3.9)	0.59	-2.0 (-9.0 to 5.0)	0.57	-12.3 (-26.8 to 2.2)	0.09	-2.4 (-6.0 to 1.2)	0.19
Season at sampling	30	-4.5 (-10.0 to 1.0)	0.11	-5.5 (-12.8 to 1.9)	0.14	-8.8 (-26.1 to 8.4)	0.30	-0.3 (-4.3 to 3.7)	0.88
Male sex	47	-0.9 (-5.4 to 3.6)	0.68	-0.9 (-6.9 to 5.0)	0.76	-5.5 (-18.1 to 7.1)	0.38	2.1 (-1.0 to 5.1)	0.18

Values represent β (95% CI) as analyzed with linear regression

Increased HADS-score: ≥8 on the HDS and/or HAS subscore

Season at sampling was divided into 4-month windows: 21/4 to 21/8 (summer) and 21/10 to 21/2 (winter)

Table 3: Associations between GC rhythm parameters in the infants' saliva and breastmilk

		Infants' saliva							
		AUCi		AUCg		Maximum		Time of maximum	
		β (95%CI)	P	β (95%CI)	P	β (95%CI)	P	β (95%CI)	P
Breastmilk	**Cortisol** AUCi	0.2 (-0.04 to 0.4)	0.10	0.2 (-0.1 to 0.5)	0.19	0.7 (-0.2 to 1.6)	0.13	-0.4 (-1.1 to 0.3)	0.26
	AUCg	0.2 (-0.02 to 0.4)	0.08	0.2 (-0.1 to 0.4)	0.16	0.6 (-0.1 to 1.4)	0.11	-0.4 (-1.0 to 0.3)	0.27
	Maximum	0.1 (0.01 to 0.12)	0.02	0.1 (-0.01 to 0.1)	0.12	0.2 (0.02 to 0.4)	0.03	-0.2 (-0.3 to 0.02)	0.07
	Time of maximum	0.0 (-0.2 to 0.2)	0.90	0.1 (-0.1 to 0.4)	0.27	-0.2 (-1.0 to 0.5)	0.57	0.4 (-0.3 to 1.0)	0.25
	Cortisone AUCi	0.4 (-0.2 to 1.0)	0.18	0.3 (-0.4 to 1.1)	0.41	1.5 (-0.1 to 3.1)	0.07	-0.4 (-0.8 to -0.1)	0.03
	AUCg	0.1 (-0.3 to 0.5)	0.66	0.1 (-0.5 to 0.7)	0.76	0.9 (-0.3 to 2.1)	0.13	-0.3 (-0.6 to -0.1)	0.02
	Maximum	0.2 (-0.1 to 0.4)	0.18	0.0 (-0.3 to 0.3)	0.91	0.4 (-0.2 to 1.1)	0.20	-0.2 (-0.4 to -0.04)	0.01
	Time of maximum	0.2 (-0.7 to 1.2)	0.65	0.8 (-0.4 to 2.1)	0.20	0.7 (-2.1 to 3.4)	0.63	0.4 (-0.3 to 1.0)	0.28

Values represent β (95% CI) as analyzed with linear regression

AUCi: area under the curve increase, representing GC variability

AUCg: area under the curve ground, representing total GC exposure

Table 4: Associations between breastmilk and infants' saliva GC concentrations per 3-hour time interval

	Cortisol		Cortisone	
Time interval (n)	β (95%CI)	*P*	β (95%CI)	*P*
0:00-3:00 (20/21)	-0.1 (-0.4 to 0.1)	0.35	-0.3 (-0.9 to 0.2)	0.21
3:00-6:00 (23)	0.2 (-0.1 to 0.5)	0.23	0.3 (-0.1 to 0.7)	0.14
6:00-9:00 (42)	0.2 (0.03 to 0.5)	0.03	0.7 (0.1 to 1.3)	0.03
9:00-12:00 (54/52)	0.0 (-0.3 to 0.3)	0.96	0.0 (-0.6 to 0.6)	0.94
12:00-15:00 (44)	0.1 (-0.04 to 0.3)	0.11	0.6 (-0.1 to 1.3)	0.09
15:00-18:00 (43)	-0.1 (-0.7 to 0.5)	0.78	-0.1 (-0.8 to 0.7)	0.88
18:00-21:00 (39)	0.1 (-0.7 to 1.0)	0.78	-0.2 (-1.4 to 1.0)	0.69
21:00-24:00 (35)	-0.3 (-1.4 to 0.7)	0.51	-0.3 (-1.6 to 1.0)	0.63

Values represent β (95% CI) as analyzed with linear mixed models, while adjusting for intra-individual measurements

When repeating the analyses while excluding the mother-infant pairs who attended the POP-clinic (n=15), small changes were found: two associations became significant, whereas three associations became non-significant (Supplementary Tables 2 and 3). Additionally, the association between breastmilk and infant GC concentrations collected between 6:00-9:00 also disappeared (Supplementary Table 4).

Sigmaplot analyses

Table 5 shows the mean differences in time of peak for the studied possible influencing factors. Infants of mothers who attended the POP-clinic had a significantly earlier time of the second salivary cortisol peak. Time of the first salivary cortisol peak was earlier in infants with a breastmilk AUCi and AUCg >p50, and time of the second salivary cortisol

Table 5: Mean differences (in hours) in time of peak for possible influencing factors

	Cortisol		Cortisone	
Influencing factors	Peak 1	Peak 2	Peak 1	Peak 2
HADS-score	0:18±2:00	6:30±3:48	-0:54±6:06	0:36±6:36
POP-clinic consultation	-1:36±1:12	-5:12±1:42**	-1:42±2:12	-2:00±2:18
Season at sampling	-2:06±1:24	-2:30±2:48	-1:54±4:12	-2:12±3:42
Sex	-0:12±2:42	1:00±3:06	-1:54±3:12	-0:24±3:18
AUCi breastmilk	-3:12±1:30*	-3:12±3:18	-1:00±4:42	-0:18±5:18
AUCg breastmilk	-2:24±0:54*	-6:00±1:42***	-1:18±3:30	0:18±3:42

Values represent mean differences ± SEM in hours as tested with t-tests

* *P* value <0.05, ** *P* value <0.01, *** *P* value <0.001

HADS-score was dichotomized as <8 or ≥8 on the anxiety and/or depression subscore

Season at sampling was divided into 4-month windows: 21/4 to 21/8 (summer) and 21/10 to 21/2 (winter)

AUCi and AUCg of breastmilk were dichotomized as < and > p50

peak was significantly earlier in infants with a breastmilk AUCg >p50. No differences were found in the timing of the salivary cortisone peaks.

When repeating the analyses while excluding mother-infant pairs who attended the POP-clinic (n=15), only one association became significant (Supplementary Table 5). None of the other associations changed.

DISCUSSION

In this study, we have shown that full-term infants at a group level have a biphasic diurnal GC rhythm at the age of 1 month with peaks in the morning as well as in the evening. Increased risk for maternal psychopathologic complaints (increased HADS-score or POP-clinic consultation), season at sampling and sex were not associated with the infants' cortisol and cortisone rhythm parameters. Maternal GCs in breastmilk might be associated with the infants' GC rhythm, since a more variable breastmilk GC rhythm appears to be associated with an earlier time of maximum in infants. However, the sample size of the study was small and results were not consistent between cortisol and cortisone parameters, and should therefore be interpreted with caution.

The most striking finding of our study is the double peak that was found in the infant GC rhythm at a group level. A double peak has been described before,[8,38,39] although those peaks were not related to a specific time of day. Several explanations are possible for the presence of this double peak rhythm. First, the double peak was seen at a group level. Since analyses at the individual level could not be performed, it is possible that the biphasic rhythm was caused by two or more separate groups of infants, with some peaking in the morning, while others had a peak occurring in the evening. However, visualizing the data per mother-infant pair revealed that several infants appeared to have a double peak, with examples shown in Figure 1C and 1D. Alternatively, a double peak could be a part of the development towards an adult-type GC rhythm. Several studies have shown that an adrenal rhythm, with a peak in the afternoon/evening, might be present in utero.[10,11] After birth, under the influence of exogenous factors,[15] an adult-type adrenal rhythm develops. The fetal GC peak in the evening could therefore slowly disappear, and a morning peak might take its place. During this development, a transitional period might exist, in which the remnants of the fetal evening peak and the beginnings of an adult-type morning peak are both present. However, to test this hypothesis, longitudinal data are necessary.

Nevertheless, as far as we are aware, this is the first study to show the presence of a double peak at the age of 1 month, although a biphasic rhythm has been reported at a younger age.[8] Ivars et al. (2015) have previously shown the presence of a significant GC rhythm at this age,[2] but did not report a double peak. Other studies have reported a

later establishment of a GC rhythm in infants.[1,7,39,40] These differences in outcomes could be due to heterogeneity in statistical methods. Price et al. (1983) [6] defined a circadian rhythm as a higher value in the morning than in the evening, with a steady decline throughout the day. Santiago et al. (1996) [7] and Antonini et al. (2000) [40] considered a circadian rhythm to be present when afternoon and evening values were 83.5% or less of the morning concentration, whereas Ivars et al. (2015) [2] used a ratio of <0.8 between morning and evening levels to determine the presence of a rhythm. De Weerth et al. (2003) [1] used hierarchical linear modeling. All of these methods are based on the premise that a rhythm is present when there is a linear decline in GC concentrations. However, as we have shown, at a group level a second peak is present in the evening. Since this peak is on average lower than the morning peak, it is possible that a circadian rhythm is considered to be present according to these other methods, while the second peak is overlooked.

The possible influencing factors considered in this study were not significantly associated with the salivary infant GC rhythm, although when analyzing only mother-infant pairs who did not attend the POP-clinic, some effects of season of sampling were found on the timing of the cortisol, but not cortisone, peak of the infants. However, several associations were found between breastmilk and salivary infant GC parameters. The cortisol maximum in breastmilk was associated with more salivary cortisol variability, a higher maximum and earlier time of maximum in infants according to linear regression analyses, and higher breastmilk cortisol variability and total exposure were associated with earlier times of salivary cortisol peaks as analyzed by SigmaPlot. More cortisone variability, total exposure and a higher maximum in breastmilk were associated with an earlier salivary cortisone peak in the infants. Additionally, breastmilk and infant GC concentrations were correlated between 6:00-9:00 (i.e., during the morning peak). Whether these findings are due to a true association is unclear, since findings were not consistent between cortisol and cortisone parameters. Cortisol concentrations especially are difficult to interpret in the saliva samples of infants, since 26% of the valid measurements were below the lower limit of detection (1 nmol/L). Cortisone levels were higher and did not reach the lower limit of detection, probably due to local conversion by 11β-HSD type 2.[30] Additionally, cortisone is has been found to be more reliable than cortisol, at least in saliva and hair.[31,32] The results which use cortisone parameters are therefore likely to be more trustworthy. However, inert breastmilk cortisone would have to be converted to active cortisol in the infant. We have previously speculated that the gut microbiota might play a role in this.[41] Additionally, 11β-HSD1 is expressed in the human intestine, liver and in the SCN.[42-44] Even if only the analyses performed with cortisone parameters are reliable, it would seem that a more variable cortisone rhythm with a high peak in breastmilk could bring the time of the morning peak forward in infants, although these associations were only found in regression analyses. In light of our previous hypothesis,

this could mean that GCs in breastmilk might aid in the transition from a fetal to an adult-type GC rhythm. The morning peak of GCs in breastmilk could have a role in this transition, since breastmilk and infant salivary GC levels were significantly correlated during this time-interval. The effects of breastmilk GCs could be caused by directly influencing GC concentrations in infant serum, although this is less likely due to the low absolute concentrations, or by acting as a signaling function in the infant's intestines (i.e., the "gut-brain axis hypothesis").[45,46] Alternatively, the associations might not be due to causality, but because another factor influences the maternal and infant HPA-axis in a similar fashion. Breastfeeding is associated with more responsive parenting [47] and increased maternal sensitivity [48] compared to formula feeding. Increased mother-infant synchrony caused by breastfeeding might therefore be a factor itself in influencing both maternal and infant GC rhythms.

This study has several strengths and limitations. First, our study's design enabled us to collect breastmilk and saliva samples at all hours of the day, with a total of 967 GC samples from 55 mother-infant pairs. Detailed analyses of infant and breastmilk rhythm could therefore be performed. Second, our analytical approach allowed for a detailed overview of infant GC rhythmicity, revealing a double peak. Third, maternal distress was measured at time of sampling and its associations with the infants' HPA-axis activity could therefore be taken into account. Our study also has its limitations. Due to collection errors, quite some infant samples did not contain enough saliva for laboratory analyses. This meant that several mother-infant pairs had to be excluded because no valid samples were available around the time of the expected (maternal) peak (i.e., 5:00-10:00) or because total sampling time was not sufficient (i.e., <8 hours). However, these exclusion criteria attempted to reduce the chances of a bias, since GC levels of the excluded mother-infants pairs were likely to have lower maximum concentrations as well as AUC's. Additionally, our sample size of 55 mother-infant pairs is quite limited, and the number of subjects that attended the POP-clinic (n=15, 27.3%) and/or had an increased HADS-score (n=10, 18.2%) was also small, although the incidence of increased psychological distress in this study was comparable with the prevalence in the general population.[49] The statistical power therefore might have been too low to detect certain associations. It also required us to pool the available data, and individual as well as adjusted analyses were therefore not possible. On the other hand, with the exception of one study,[2] our sample size was bigger than other studies assessing HPA-axis development in early-life in term infants.[1,3,6-8,39,40] Moreover, although we studied several possible influencing factors, we did not take all determinants into consideration. For instance, no data was collected about timing of daytime naps, while sleep has previously been associated with GC levels in infants.[1,8,50,51] In the previous studies, most of these associations were found in older infants than the ones included in this study, but an effect cannot be ruled out. However, daytime naps were associated with decreased cortisol levels im-

mediately after the nap,[51] and they are therefore unlikely to explain the biphasic rhythm found in this study. Sleeping through the night was found to be associated with a more pronounced circadian rhythm,[1] but it has previously also been shown that the establishment of a diurnal GC rhythm precedes a sleep-wake rhythm.[3] Additionally, the lack of a formula-fed control group limited our possibilities with regard to testing the associations between breastfeeding and GC rhythmicity in the infants, and the effect of the act of breastfeeding as a whole could therefore not be studied. Furthermore, the possibility of a selection bias cannot be excluded, because stressed mothers with infants who slept restlessly (indicative of a lack of rhythm) were probably less likely to participate, and it is therefore likely that the study population does not reflect the general or the POP-clinic population. However, we did not aim to find reference ranges, but designed this study to analyze effects of inter-individual variation in breastmilk rhythmicity. We did not collect data on mothers who were eligible for inclusion but opted out of participating, and a selection bias could consequently not be tested. Additionally, mothers who attended the POP-clinic might have had other reasons for not participating compared to mothers who did not attend the POP-clinic, which could have further skewed results. Lastly, a longitudinal study design would have enabled us a better understanding of HPA-axis development and which factors are of influence. However, we aimed to make this study as non-invasive as possible, and therefore decided to have mothers collect milk and saliva samples during one day only.

In conclusion, a biphasic GC rhythm appears to be present at a group level at the age of 1 month, with a peak in both the morning and the evening, which might be part of the developmental process towards an adult-type GC rhythm. Increased risk for maternal psychopathologic complaints (increased HADS-score or POP-clinic consultation), season at sampling and sex were not associated with infant GC rhythmicity in this study. However, breastmilk GC parameters might be associated with the infants' GC rhythm, which might be due to a causal signaling effect of breastmilk GCs, or because of an associative effect due to increased mother-infant synchrony. Although future studies should further elucidate HPA-axis development in early life, preferably with a longitudinal design and including a formula-fed control group, this exploratory study contributes to an increased understanding of this process, especially with regard to the role of breastmilk.

REFERENCES

1. de Weerth C, Zijl RH, Buitelaar JK. Development of cortisol circadian rhythm in infancy. Early Hum Dev 2003; 73:39-52
2. Ivars K, Nelson N, Theodorsson A, Theodorsson E, Strom JO, Morelius E. Development of Salivary Cortisol Circadian Rhythm and Reference Intervals in Full-Term Infants. PLoS One 2015; 10:e0129502
3. Joseph D, Chong NW, Shanks ME, Rosato E, Taub NA, Petersen SA, Symonds ME, Whitehouse WP, Wailoo M. Getting rhythm: how do babies do it? Arch Dis Child Fetal Neonatal Ed 2015; 100:F50-54
4. Kiess W, Meidert A, Dressendorfer RA, Schriever K, Kessler U, Konig A, Schwarz HP, Strasburger CJ. Salivary cortisol levels throughout childhood and adolescence: relation with age, pubertal stage, and weight. Pediatr Res 1995; 37:502-506
5. Lewis M, Ramsay DS. Developmental change in infants' responses to stress. Child Dev 1995; 66:657-670
6. Price DA, Close GC, Fielding BA. Age of appearance of circadian rhythm in salivary cortisol values in infancy. Arch Dis Child 1983; 58:454-456
7. Santiago LB, Jorge SM, Moreira AC. Longitudinal evaluation of the development of salivary cortisol circadian rhythm in infancy. Clin Endocrinol (Oxf) 1996; 44:157-161
8. Spangler G. The emergence of adrenocortical circadian function in newborns and infants and its relationship to sleep, feeding and maternal adrenocortical activity. Early Hum Dev 1991; 25:197-208
9. Custodio RJ, Junior CE, Milani SL, Simoes AL, de Castro M, Moreira AC. The emergence of the cortisol circadian rhythm in monozygotic and dizygotic twin infants: the twin-pair synchrony. Clin Endocrinol (Oxf) 2007; 66:192-197
10. Seron-Ferre M, Riffo R, Valenzuela GJ, Germain AM. Twenty-four-hour pattern of cortisol in the human fetus at term. Am J Obstet Gynecol 2001; 184:1278-1283
11. Patrick J, Challis J, Natale R, Richardson B. Circadian rhythms in maternal plasma cortisol, estrone, estradiol, and estriol at 34 to 35 weeks' gestation. Am J Obstet Gynecol 1979; 135:791-798
12. Iwata O, Okamura H, Saitsu H, Saikusa M, Kanda H, Eshima N, Iwata S, Maeno Y, Matsuishi T. Diurnal cortisol changes in newborn infants suggesting entrainment of peripheral circadian clock in utero and at birth. J Clin Endocrinol Metab 2013; 98:E25-32
13. Rivkees SA. Developing circadian rhythmicity in infants. Pediatrics 2003; 112:373-381
14. Seron-Ferre M, Torres C, Parraguez VH, Vergara M, Valladares L, Forcelledo ML, Constandil L, Valenzuela GJ. Perinatal neuroendocrine regulation. Development of the circadian time-keeping system. Mol Cell Endocrinol 2002; 186:169-173
15. McMillen IC, Kok JS, Adamson TM, Deayton JM, Nowak R. Development of circadian sleep-wake rhythms in preterm and full-term infants. Pediatr Res 1991; 29:381-384
16. Morin LP. The circadian visual system. Brain Res Brain Res Rev 1994; 19:102-127
17. Rivkees SA, Mayes L, Jacobs H, Gross I. Rest-activity patterns of premature infants are regulated by cycled lighting. Pediatrics 2004; 113:833-839
18. Martini J, Petzoldt J, Knappe S, Garthus-Niegel S, Asselmann E, Wittchen HU. Infant, maternal, and familial predictors and correlates of regulatory problems in early infancy: The differential role of infant temperament and maternal anxiety and depression. Early Hum Dev 2017; 115:23-31

19. Jarcho MR, Slavich GM, Tylova-Stein H, Wolkowitz OM, Burke HM. Dysregulated diurnal cortisol pattern is associated with glucocorticoid resistance in women with major depressive disorder. Biol Psychol 2013; 93:150-158

20. Vreeburg SA, Hoogendijk WJ, DeRijk RH, van Dyck R, Smit JH, Zitman FG, Penninx BW. Salivary cortisol levels and the 2-year course of depressive and anxiety disorders. Psychoneuroendocrinology 2013; 38:1494-1502

21. Thomas KA, Burr RL, Spieker S. Light and maternal influence in the entrainment of activity circadian rhythm in infants 4-12 weeks of age. Sleep Biol Rhythms 2016; 14:249-255

22. Cubero J, Narciso D, Terron P, Rial R, Esteban S, Rivero M, Parvez H, Rodriguez AB, Barriga C. Chrononutrition applied to formula milks to consolidate infants' sleep/wake cycle. Neuro Endocrinol Lett 2007; 28:360-366

23. Lavelli M, Poli M. Early mother-infant interaction during breast- and bottle-feeding. Infant Behavior & Development 1998; 21:667-683

24. Kuzela ALS, C.A.; Worobey, J. Breastfeeding and mother-infant interactions. Journal of Reproductive and Infant Psychology 1990; 8:185-194

25. Illnerova H, Buresova M, Presl J. Melatonin rhythm in human milk. J Clin Endocrinol Metab 1993; 77:838-841

26. van der Voorn B, de Waard M, van Goudoever JB, Rotteveel J, Heijboer AC, Finken MJ. Breast-Milk Cortisol and Cortisone Concentrations Follow the Diurnal Rhythm of Maternal Hypothalamus-Pituitary-Adrenal Axis Activity. J Nutr 2016; 146:2174-2179

27. Angelucci L, Patacchioli FR, Scaccianoce S, Di Sciullo A, Cardillo A, Maccari S. A model for later-life effects of perinatal drug exposure: maternal hormone mediation. Neurobehav Toxicol Teratol 1985; 7:511-517

28. Cao Y, Rao SD, Phillips TM, Umbach DM, Bernbaum JC, Archer JI, Rogan WJ. Are breast-fed infants more resilient? Feeding method and cortisol in infants. The Journal of pediatrics 2009; 154:452-454

29. Benjamin Neelon SE, Stroo M, Mayhew M, Maselko J, Hoyo C. Correlation between maternal and infant cortisol varies by breastfeeding status. Infant Behav Dev 2015; 40:252-258

30. Smith RE, Maguire JA, Stein-Oakley AN, Sasano H, Takahashi K, Fukushima K, Krozowski ZS. Localization of 11 beta-hydroxysteroid dehydrogenase type II in human epithelial tissues. J Clin Endocrinol Metab 1996; 81:3244-3248

31. Blair J, Adaway J, Keevil B, Ross R. Salivary cortisol and cortisone in the clinical setting. Curr Opin Endocrinol Diabetes Obes 2017; 24:161-168

32. Savas M, Wester VL, de Rijke YB, Rubinstein G, Zopp S, Dorst K, van den Berg SAA, Beuschlein F, Feelders RA, Reincke M, van Rossum EFC. Hair glucocorticoids as biomarker for endogenous Cushing's syndrome: validation in two independent cohorts. Neuroendocrinology 2019; 109(2):171-178

33. Herrmann C. International experiences with the Hospital Anxiety and Depression Scale--a review of validation data and clinical results. J Psychosom Res 1997; 42:17-41

34. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983; 67:361-370

35. van der Voorn B, Martens F, Peppelman NS, Rotteveel J, Blankenstein MA, Finken MJ, Heijboer AC. Determination of cortisol and cortisone in human mother's milk. Clin Chim Acta 2015; 444:154-155

36. Hollanders JJ, van der Voorn B, Rotteveel J, Finken MJJ. Is HPA axis reactivity in childhood gender-specific? A systematic review. Biol Sex Differ 2017; 8:23

37. Pruessner JC, Kirschbaum C, Meinlschmid G, Hellhammer DH. Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. Psychoneuroendocrinology 2003; 28:916-931

38. Francis SJ, Walker RF, Riad-Fahmy D, Hughes D, Murphy JF, Gray OP. Assessment of adrenocortical activity in term newborn infants using salivary cortisol determinations. J Pediatr 1987; 111:129-133

39. Vermes I, Dohanics J, Toth G, Pongracz J. Maturation of the circadian rhythm of the adrenocortical functions in human neonates and infants. Horm Res 1980; 12:237-244

40. Antonini SR, Jorge SM, Moreira AC. The emergence of salivary cortisol circadian rhythm and its relationship to sleep activity in preterm infants. Clin Endocrinol (Oxf) 2000; 52:423-426

41. Hollanders JJ, Heijboer AC, van der Voorn B, Rotteveel J, Finken MJJ. Nutritional programming by glucocorticoids in breast milk: Targets, mechanisms and possible implications. Best Pract Res Clin Endocrinol Metab 2017; 31:397-408

42. Bisschop PH, Dekker MJ, Osterthun W, Kwakkel J, Anink JJ, Boelen A, Unmehopa UA, Koper JW, Lamberts SW, Stewart PM, Swaab DF, Fliers E. Expression of 11beta-hydroxysteroid dehydrogenase type 1 in the human hypothalamus. J Neuroendocrinol 2013; 25:425-432

43. Stegk JP, Ebert B, Martin HJ, Maser E. Expression profiles of human 11beta-hydroxysteroid dehydrogenases type 1 and type 2 in inflammatory bowel diseases. Mol Cell Endocrinol 2009; 301:104-108

44. Tomlinson JW, Walker EA, Bujalska IJ, Draper N, Lavery GG, Cooper MS, Hewison M, Stewart PM. 11beta-hydroxysteroid dehydrogenase type 1: a tissue-specific regulator of glucocorticoid response. Endocr Rev 2004; 25:831-866

45. Cryan JF, Dinan TG. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. Nat Rev Neurosci 2012; 13:701-712

46. Foster JA, McVey Neufeld KA. Gut-brain axis: how the microbiome influences anxiety and depression. Trends Neurosci 2013; 36:305-312

47. Ventura AK. Associations between Breastfeeding and Maternal Responsiveness: A Systematic Review of the Literature. Adv Nutr 2017; 8:495-510

48. Kim P, Feldman R, Mayes LC, Eicher V, Thompson N, Leckman JF, Swain JE. Breastfeeding, brain activation to own infant cry, and maternal sensitivity. J Child Psychol Psychiatry 2011; 52:907-915

49. Andersson L, Sundstrom-Poromaa I, Bixo M, Wulff M, Bondestam K, aStrom M. Point prevalence of psychiatric disorders during the second trimester of pregnancy: a population-based study. Am J Obstet Gynecol 2003; 189:148-154

50. de Weerth CvG, P. A longitudinal study of basal cortisol in infants:Intra-individual variability, circadian rhythmand developmental trends. Infant Behav Dev 2002; 25:375-398

51. Larson MC, Gunnar MR, Hertsgaard L. The effects of morning naps, car trips, and maternal separation on adrenocortical activity in human infants. Child Dev 1991; 62:362-372

Supplementary Table 1: Fit parameters of single, double and triple peak equations as tested with Sigma-plot

		P value	adjusted R-squared	Residuals (mean squares)	Number of parameters	Additional notes
Cortisol	Single	1	NA	18,0329	3	
	Double	<0.0001	0,0652	16,1487	6	
	Triple	<0.0001	0,079	15,9099	9	1 non-significant parameter in model
Cortisone	Single	1	NA	409,5641	3	
	Double	<0.0001	0,0978	322,2971	6	
	Triple	<0.0001	0,096	322,9402	9	2 non-significant parameters in model

Supplementary Table 2: Associations between GC rhythm parameters in the infants' saliva and possible influencing factors for mother-infant pairs who did not attend the POP-clinic

			AUCi		AUCg		Maximum		Time of maximum	
		n	β (95%CI)	*P*	β (95%CI)	*P*	β (95%CI)	*P*	β (95%CI)	*P*
Cortisol	Increased HADS-score	36	1.3 (-0.5 to 3.3)	0.16	1.1 (-1.3 to 3.4)	0.37	7.9 (1.2 to 14.5)	0.02	-2.2 (8.4 to 3.7)	0.46
	Season at sampling	25	-0.4 (-2.0 to 1.2)	0.60	-0.6 (-2.6 to 1.3)	0.51	-2.0 (-7.9 to 3.9)	0.49	-4.7 (-9.4 to -0.04)	0.05
	Male sex	37	-0.9 (-2.2 to 0.2)	0.10	-0.9 (-2.3 to 0.6)	0.24	-5.4 (-9.6 to -1.3)	0.01	-0.6 (-4.4 to 3.1)	0.74
Cortisone	Increased HADS-score	36	4.4 (-3.3 to 12.0)	0.25	2.5 (-8.0 to 13.1)	0.63	8.8 (-14.4 to 31.9)	0.45	-1.8 (-7.4 to 3.9)	0.53
	Season at sampling	25	-6.2 (-12.0 to -0.5)	0.04	-7.8 (-15.8 to 0.3)	*0.06*	-14.1 (-32.5 to 4.4)	0.13	-0.8 (-5.4 to 3.8)	0.72
	Male sex	37	-1.2 (-6.2 to 3.7)	0.62	-0.9 (-7.8 to 5.9)	0.78	-4.1 (-19.0 to 10.7)	0.58	1.3 (-2.3 to 5.0)	0.46

Values represent β (95% CI) as analyzed with linear regression

Increased HADS-score: ≥8 on the HDS and/or HAS subscore

Season at sampling was divided into 4-month windows: 21/4 to 21/8 (summer) and 21/10 to 21/2 (winter)

Supplementary Table 3: Associations between GC rhythm parameters in the infants' saliva and breastmilk for mother-infant pairs who did not attend the POP-clinic

		Infants' saliva							
		AUCi		**AUCg**		**Maximum**		**Time of maximum**	
		β (95%CI)	P	β (95%CI)	P	β (95%CI)	P	β (95%CI)	P
Breastmilk — Cortisol	AUCi	0.1 (-0.1 to 0.4)	0.27	0.1 (-0.2 to 0.4)	0.40	0.5 (-0.5 to 1.4)	0.30	-0.4 (-1.2 to 0.4)	0.36
	AUCg	0.1 (-0.09 to 0.4)	0.21	0.1 (-0.2 to 0.4)	0.34	0.5 (-0.3 to 1.4)	0.23	-0.3 (-1.1 to 0.4)	0.34
	Maximum	0.05 (-0.02 to 0.11)	0.16	0.03 (-0.05 to 0.1)	0.42	0.1 (-0.1 to 0.4)	0.27	-0.1 (-0.3 to 0.07)	0.18
	Time of maximum	0.03 (-0.2 to 0.2)	0.80	0.2 (-0.1 to 0.4)	0.24	-0.1 (-0.9 to 0.7)	0.74	0.2 (-0.4 to 0.9)	0.50
Breastmilk — Cortisone	AUCi	0.5 (-0.09 to 1.1)	0.09	0.5 (-0.4 to 1.4)	0.26	1.6 (-0.2 to 3.5)	0.08	-0.4 (-0.9 to 0.04)	0.08
	AUCg	0.1 (-0.3 to 0.6)	0.57	0.1 (-0.5 to 0.8)	0.69	0.8 (-0.5 to 2.2)	0.24	-0.3 (-0.6 to -0.1)	0.04
	Maximum	0.1 (-0.1 to 0.4)	0.30	-0.04 (-0.4 to 0.3)	0.83	0.3 (-0.5 to 1.1)	0.42	-0.2 (-0.4 to -0.02)	0.03
	Time of maximum	0.2 (-1.0 to 1.3)	0.77	1.1 (-0.4 to 2.7)	0.14	1.3 (-2.1 to 4.7)	0.44	0.3 (-0.5 to 1.1)	0.47

Values represent β (95% CI) as analyzed with linear regression
AUCi: area under the curve increase, representing GC variability
AUCg: area under the curve ground, representing total GC exposure

Supplementary Table 4: Associations between breastmilk and infants' saliva GC concentrations per 3-hour time interval for mother-infant pairs who did not attend the POP-clinic

	Cortisol		**Cortisone**	
	β (95%CI)	P	β (95%CI)	P
0:00-3:00 (n=18)	-0.1 (-0.4 to 0.1)	0.25	-0.3 (-0.9 to 0.3)	0.33
3:00-6:00 (n=19)	0.0 (-0.4 to 0.4)	0.98	0.2 (-0.2 to 0.6)	0.30
6:00-9:00 (n=35/34)	0.2 (-0.07 to 0.4)	0.17	0.5 (-0.2 to 1.2)	0.17
9:00-12:00 (n=38/37)	0.0 (-0.4 to 0.3)	0.88	0.1 (-0.8 to 0.9)	0.88
12:00-15:00 (n=35/34)	0.2 (-0.04 to 0.4)	0.12	0.8 (-0.04 to 1.6)	0.06
15:00-18:00 (n=29)	-0.2 (-0.9 to 0.6)	0.65	0.2 (-0.7 to 1.2)	0.62
18:00-21:00 (n=29)	0.2 (-0.7 to 1.1)	0.59	-0.1 (-1.4 to 1.2)	0.86
21:00-24:00 (n=30)	-0.4 (-1.5 to 0.8)	0.51	-0.4 (-1.9 to 1.0)	0.56

Values represent β (95% CI) as analyzed with linear mixed models, while adjusting for intra-individual measurements

Supplementary Table 5: Mean differences (in hours) in time of peak for possible influencing factors for mother-infant pairs who did not attend the POP-clinic

	Cortisol		Cortisone	
	Peak 1	**Peak 2**	**Peak 1**	**Peak 2**
HADS-score	-1:42±2:36	-8:06±4:08	-0:18±9:42	-1:06±10:42
Season at sampling	-2:54±1:24[*]	-3:06±1:48	-3:30±6:06	-3:24±4:42
Sex	-0:06±2:42	0:30±3:57	-1:51±3:51	-0:27±3:54
AUCi breastmilk	-3:00±1:24[*]	Cannot be calculated	-1:00±7:09	-1:51±7:15
AUCg breastmilk	-2:51±1:09[*]	-6:12±3:12	-1:15±4:54	-0:21±4:51

Values represent mean differences ± SEM in hours as tested with t-tests
* *P* value <0.05, ** *P* value <0.01, *** *P* value <0.001
HADS-score was dichotomized as <8 or ≥8 on the anxiety and/or depression subscore
Season at sampling was divided into 4-month windows: 21/4 to 21/8 (summer) and 21/10 to 21/2 (winter)
AUCi and AUCg of breastmilk were dichotomized as < and > p50

EVA MOUTON 2019

7

No association between glucocorticoid diurnal rhythm in breastmilk and infant body composition at age 3 months

Jonneke Hollanders[*],
Lisette R. Dijkstra[*],
Bibian van der Voorn,
Stefanie M.P. Kouwenhoven,
Alyssa A. Toorop,
Johannes B. van Goudoever,
Joost Rotteveel,
Martijn J.J. Finken
[*] Authors contributed equally to this manuscript

ABSTRACT

Objective

Glucocorticoids (GCs) in breastmilk have previously been associated with infant body growth and body composition. However, the diurnal rhythm of breastmilk GCs was not taken into account, and we therefore aimed to assess the assocations between breast-milk GC rhythmicity at age 1 month and growth and body composition at age 3 months.

Methods

At one month postpartum, breastmilk GCs were collected over a 24-h period and ana-lyzed by LC-MS/MS. Body composition was measured using air-displacement plethys-mography at age 3 months. Length and weight were collected at age 1, 2 and 3 months.

Results

39 healthy mother-infant pairs were included. No associations were found between breastmilk GC rhythmicity (area-under-the-curve increase and ground, maximum and delta) and infant growth trajectories or body composition (fat and fat free mass index, fat%) at age 3 months.

Conclusions

This study did not find an association between breastmilk GC rhythmicity at 1 month and infant's growth or body composition at age 3 months. Therefore, this study suggests that previous observations linking breastmilk cortisol to changes in infant weight might be flawed by the lack of serial cortisol measurements and detailed information on body composition.

INTRODUCTION

Growing attention is focused on the etiology of obesity, and it has been hypothesized that part of its origin can be traced back to events occurring in early life (i.e., the Developmental Origins of Health and Disease [DOHaD] hypothesis).[1]

Given its effects on fat disposition and metabolism, the hypothalamus-pituitary-adrenal (HPA) axis has been implicated to play a role in the pathway leading to obesity.[2,3] Not only endogenous, but also maternal glucocorticoids (GCs) appear to be involved. Evidence from animal experiments indicates that increased transplacental supply of maternal GCs may be associated with a lower birth weight and cardiovascular correlates such as hypertension and hyperglycemia.[4] In humans, fetal exposure to excess maternal cortisol, e.g., due to maternal anxiety or depression, has been associated with a higher risk of childhood adiposity.[5]

After birth, small amounts of maternal GCs appear to be transferred to the developing infant through breastmilk. Maternal GCs in breastmilk have been shown to cross the intestinal barrier in animals,[6] and have been associated with growth and body composition. Hinde et al. (2015)[7] found that cortisol in breastmilk of rhesus macaques was positively associated with weight gain in offspring. In humans, Hahn-Holbrook et al. (2016)[8] showed that cortisol in breastmilk at age 3 months was inversely associated with body mass index (BMI) percentile gains in the first 2 years of life. Whether the findings from these studies are contradictory is unclear, since length was not taken into account by Hinde et al.[7] Moreover, the effect of GCs on growth might change between the ages of 3 months and 2 years.

Our group has previously shown that GCs in breastmilk follow maternal HPA-axis activity, with a peak in the morning and a nadir at night.[9] Although previous studies have found associations between cortisol in breastmilk and growth of offspring, none of them took GC rhythmicity into account. However, obesity has previously been associated with a flatter diurnal cortisol slope in adults,[10] and there is also some evidence that a blunted GC rhythm is associated with obesity in children.[11,12] Both Hinde et al. (2015)[7] and Hahn-Holbrook et al. (2016)[8] did not collect samples around peak GC levels, while Hahn-Holbrook et al. (2016) also had a wide time window during which samples could be collected (11:30-16:00).

We therefore aimed to assess the associations between breastmilk GC rhythmicity and infant growth and body composition. We measured cortisol and cortisone in breastmilk at age 1 month over a 24-hour period, measured body composition using air-displacement plethysmography at age 3 months, and collected length and weight data monthly up to that age. Due to associations found between a blunted endogenous GC rhythm and obesity in both children and adults,[10-12] we hypothesized that less GC variability in breastmilk could be associated with a higher fat mass in the infants.

METHODS

Population

Healthy mother-infant pairs were recruited at the maternity ward of the Amsterdam UMC, location VUMC (a tertiary hospital) in the Netherlands between March 2016 and July 2017. Subjects were eligible for inclusion when infants were born at term age (37-42 weeks) with a normal birth weight (-2 to +2 SDS), and when mothers had the intention to breastfeed for a minimum of three months. Exclusion criteria were: 1) major congenital anomalies, 2) multiple pregnancy, 3) pre-eclampsia or HELLP, 4) medication use other than "over the counter" drugs, 5) maternal alcohol consumption of >7 IU/week and/or 6) a maternal temperature of >38.5°C at the time of sampling. Approval of the Medical Ethics Committee of the VUMC was obtained (protocol number 2015.524), and written informed consent was obtained from all participating mothers.

Data collection

Peripartum

Shortly after inclusion, within the first days postpartum, mothers filled in a questionnaire pertaining to their pregnancy and birth, as well as maternal and infant anthropometric and demographic data.

One month postpartum

At 30 days postpartum (± 5 days), mothers collected a portion of breastmilk (1-2 mL) prior to each feeding moment, over a 24h period (i.e., five to eight times). Although only foremilk was collected through this method, previous research has shown that GC concentrations are similar in fore- and hindmilk.[13] Mothers could follow their own feeding schedule and were therefore asked to report the exact time of sampling. Milk was collected manually or with a breast pump; we requested that mothers used the same method for all samples. Milk was stored in the mother's freezer, and subsequently in the laboratory at -20°C for less than 3 months prior to analysis.

At the time of sampling, maternal distress was quantified with the Hospital Anxiety and Depression Scale (HADS).[14] This questionnaire contains 14 questions scored from 0-3, which assess self-reported levels of depression and anxiety symptoms. Seven questions concern depressive symptoms (HDS) and seven anxiety symptoms (HAS). A score of ≥8 on a subscale is indicative of clinically relevant depression and/or anxiety symptoms.

Three months postpartum

At 3 months postpartum (± 2 wks), body composition of the infants was assessed with the Pea Pod, an air-displacement plethysmography (ADP) system (COSMED USA, Inc., Concord, CA, USA)[15] It is based on a bi-compartmental model, which uses pressure and

volume changes in the chamber through which body density were determined. Age- and sex-specific fat and fat free mass density values were subsequently used to calculate fat mass (FM) and fat free mass (FFM). [15]

As part of the national standard care, weight and length at 1, 2 and 3 months of age were measured by the staff of the child health clinic and were obtained through a questionnaire. Weight was measured fully undressed on a balance scale with an accuracy of 1 gr. Length was measured in supine position to the nearest 0.1 cm. Additionally, all mothers were asked if their infants were still breastfed for >80% at the age of 3 months.

Laboratory analysis

Cortisol and cortisone concentrations in breastmilk were determined by isotope dilution liquid chromatography–tandem mass spectrometry (LC–MS/MS), as previously described.[16] In short, internal standards ($^{13}C_3$-labeled cortisol and $^{13}C_3$-labeled cortisone) were added to 200 µl of the samples. Then, breastmilk was washed 3 times with 2 mL hexane to remove lipids. Finally, samples were extracted and analyzed using Isolute plates (Biotage, Uppsala, Sweden) and analyzed by LC-MS/MS (Acquity with Quattro Premier XE, Milford MA, USA, Waters Corporation). The intra-assay coefficients of variation (CV%) were 4 and 5% for cortisol levels of 7 and 23 nmol/L, and 5% for cortisone levels of 8 and 33 nmol/L for LC-MS/MS measurements. The inter-assay CV% was <9% for both cortisol and cortisone. The Lower Limit of Quantitation (LLOQ) was 0.5 nmol/L for both cortisol and cortisone. All samples were measured in duplo.

Statistics

First, data of GC concentrations in breastmilk were converted into the following rhythm parameters, in order to provide a full overview of GC rhythmicity:
- The maximum GC concentration, as a proxy for peak concentrations
- The delta between maximum and minimum GC concentrations, as a measure of rhythm variability
- Area Under the Curve (AUC) ground (g) and increase (i), using the trapezoid rule.[17] Calculations were corrected for total sampling time, since this differed between mothers. AUCg is a measure of total GC exposure, while AUCi provides information on GC variability.

Mother-infant pairs were excluded from analyses when no valid GC data was available around the time of the expected morning peak (5:00-10:00) and/or when total sample collection was <8 hours.

Fat% was determined from FM and FFM values. Fat Mass Index (FMI) and Fat Free Mass Index (FFMI) were calculated by dividing FM and FFM values (in kg) respectively, by infant length squared (m^2), since fat mass and fat free mass are known to change with

length.[18] Length and weight data were converted to SDS.[19,20] Body mass index (BMI) was calculated for age 3 months only, and converted to SDS.[19]

Linear regressions were used to assess the associations between GC rhythm parameters at age 1 month and length SDS, weight SDS, BMI SDS, FMI and FFMI at age 3 months. First, unadjusted regression analyses were performed. Next, the following potential confounders were tested: sex, HADS-score (HAS and/or HDS ≥8), prepregnancy BMI, ethnicity (Caucasian vs. non-Caucasian), socio-economic status, birth weight SDS, gestational age, weight gain during pregnancy, parity (1 vs. >1), mode of delivery (vaginal vs. caesarian section), and %breastmilk at age 3 months (< or >80%). Due to our sample size, the three variables with the largest confounding effect (i.e., largest change in β of the independent variable) were used for the multiple linear regression analyses. Thus, weight gain during pregnancy, % breastmilk at age 3 months and ethnicity were included in the final model assessing the association between GC rhythm parameters and body composition outcomes. No effect modification was found for infant sex, and analyses were therefore not stratified.

Lastly, length and weight SDS growth trajectories between age 1 to 3 months were plotted against AUCi and AUCg values by using generalized estimating equations (GEEs), and 95% confidence intervals were calculated according to the method described by Figueiras et al. [21] AUCi and AUCg outcomes for cortisol and cortisone were categorized as ≤p25, p25-75 and ≥p75.

RESULTS

Population
Forty-four mother-infant pairs were included in the study. One mother-infant pair was lost to follow-up, three mother-infant pairs returned the growth questionnaires but did not consent to the Pea Pod measurement and one pair was excluded because no samples were collected between 5:00-10:00 and/or because total sampling time was <8 hours. Therefore, a total of 42 mother-infant pairs were included in the growth trajectory analyses, whereas 39 mother-infants pairs were included in the body composition analyses at age 3 months. Of the included mother-infant pairs, 59.5% were mother-son pairs. Table 1 shows the characteristics of the population. Supplementary Table 1 shows the cortisol and cortisone concentrations in breastmilk in 4-hour intervals.

Linear regression analyses
No associations were found between the GC rhythm parameters (AUCi, AUCg, maximum and delta) and body composition in the unadjusted analyses. Adjusting the analyses

Table 1: Characteristics of the study population (n=42)

Gestational age	wks	39.9±1.3
Birth weight	grams	3561±498
	SDS	0.2±1.0
Birth length*	cm	52.0±2.6
	SDS	1.0±1.6
Male sex		25 (59.5)
Primiparity		23 (54.8)
Caesarian section		21 (51.2)
HAS and/or HDS ≥8 at 1 month pp		6 (14.6)
Prepregnancy maternal BMI	kg/m^2	22.3±2.8
Weight gain during pregnancy	kg	13.1±3.2
Maternal age	Yrs	36.0±4.7
Non-Caucasian ethnicity		8 (20.0)
Socioeconomic status	SDS	0.6±1.2
>80% breastfed at age 3 months		35 (87.5)
Age at breastmilk sampling	days	30.8±2.6
Age at Pea Pod measurement**	days	90.5±7.0

Values represent mean±SD or n (%); pp= postpartum
HAS, Hospital Anxiety Score; HDS, Hospital Depression Score
* n=31
** n=39

for weight gain during pregnancy, % breastmilk at age 3 months, and ethnicity did not change the results (Table 2).

Growth trajectories

Figure 1 shows the growth trajectories for length and weight SDS according to breastmilk cortisone AUCi and AUCg outcomes. No differences were found between the categories ≤p25, p25-75 and ≥p75. Results for breastmilk cortisol AUCi and AUCg were similar [data not shown].

DISCUSSION

In this study, despite increased evidence for associations between blunted endogenous GC rhythms and obesity in both children and adults,[10-12] no associations were found between GC rhythmicity in breastmilk sampled at 1 month and infant body composition or growth at age 3 months. Therefore, our study could not confirm previous observations in animals and humans. Hinde et al. (2015)[7] measured cortisol in breastmilk of 108

Table 2: Adjusted associations between breastmilk GC rhythmicity at age 1 month and infant body composition at age 3 months (n=39)

		Length	Weight		BMI		FMI		FFMI		Fat %	
		B	95% CI	B	95% CI	B	95% CI	B	95% CI	B	95% CI	B	95% CI
Cortisol	Maximum	0.006 (-0.04 to 0.05)	0.022 (-0.02 to 0.07)	0.022 (-0.02 to 0.06)	0.003 (-0.04 to 0.05)	-0.006 (-0.04 to 0.03)	0.048 (-0.17 to 0.27)
	Delta	0.006 (-0.04 to 0.05)	0.024 (-0.02 to 0.07)	0.025 (-0.02 to 0.07)	0.003 (-0.04 to 0.05)	-0.004 (-0.04 to 0.03)	0.043 (-0.18 to 0.26)
	AUCi	0.025 (-0.15 to 0.20)	0.101 (-0.07 to 0.28)	0.1 (-0.06 to 0.26)	0.06 (-0.10 to 0.23)	-0.095 (-0.23 to 0.04)	0.53 (-0.34 to 1.39)
	AUCg	0.029 (-0.13 to 0.19)	0.06 (-0.11 to 0.23)	0.046 (-0.11 to 0.20)	0.06 (-0.10 to 0.22)	-0.107 (-0.24 to 0.02)	0.53 (-0.28 to 1.33)
Cortison	Maximum	-0.002 (-0.04 to 0.03)	0.01 (-0.03 to 0.05)	0.014 (-0.02 to 0.05)	-0.006 (-0.04 to 0.03)	-0.006 (-0.04 to 0.02)	-0.007 (-0.19 to 0.18)
	Delta	-0.002 (-0.04 to 0.04)	0.018 (-0.02 to 0.06)	0.024 (-0.01 to 0.06)	-0.003 (-0.04 to 0.03)	-0.001 (-0.03 to 0.03)	0.005 (-0.19 to 0.20)
	AUCi	-0.005 (-0.09 to 0.08)	0.042 (-0.05 to 0.13)	0.055 (-0.02 to 0.14)	0.002 (-0.08 to 0.09)	-0.008 (-0.08 to 0.06)	0.034 (-0.41 to 0.48)
	AUCg	-0.001 (-0.07 to 0.07)	0.003 (-0.07 to 0.07)	0.004 (-0.06 to 0.07)	-0.013 (-0.08 to 0.05)	-0.02 (-0.08 to 0.04)	-0.02 (-0.37 to 0.33)

Values represent β (95% CI) as analyzed with linear regression

Analyses were adjusted for weight gain during pregnancy, % breastmilk at age 3 months and ethnicity

AUCi or g, Area Under the Curve increase or ground; FMI, fat mass index; FFMI, fat free mass index
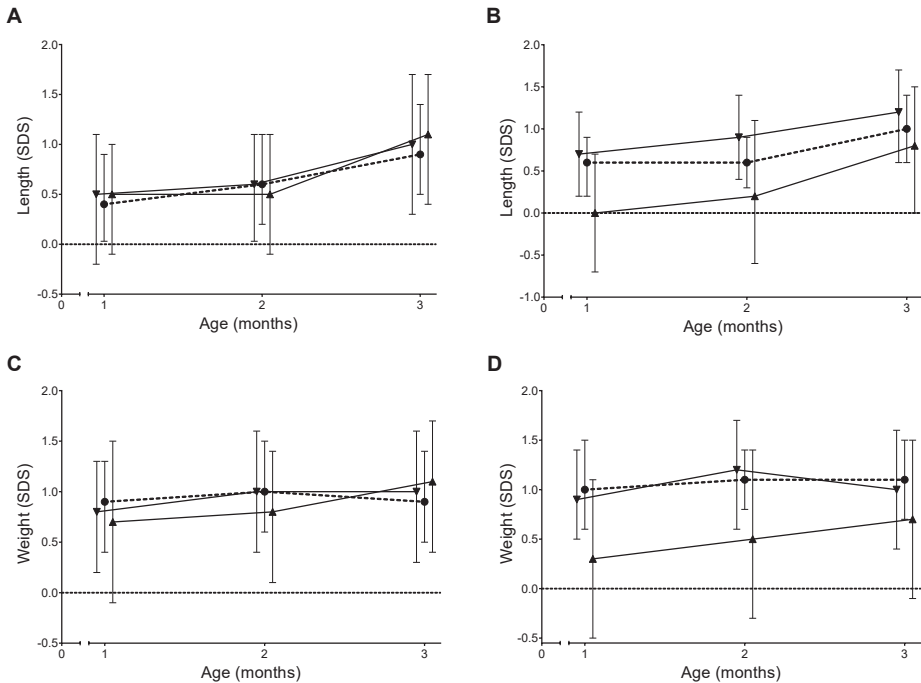
Figure 1: Growth trajectories between age 1 to 3 months for length and weight, according to breastmilk cortisone AUC outcomes (n=42). Results for breastmilk cortisol AUCi and AUCg were similar [data not shown]. A= length for AUCi, B= weight for AUCg, C= weight for AUCi, D= weight for AUCg; ▼ = AUCi or g <p25, ● = AUCi or g = p25-75, ▲ = AUCi or g >p75 ; AUCi or g, Area Under the Curve increase or ground

rhesus macaques at 1 month of age, and analyzed growth outcomes at 3.5 months of age. They found that higher cortisol concentrations were associated with greater weight gain over time. Hahn-Holbrook et al. (2016)[8] studied associations between breast-milk cortisol and BMI gains up until the age of 2 in 51 mother-infant pairs. They found that higher milk cortisol concentrations were associated with smaller BMI gains in offspring.

The different results between this study and previous studies could have several explanations. First, cortisol sampling in the previous studies did not take the diurnal rhythm of breastmilk GCs into account. Hinde et al.[7] sampled between 11:30-13:00, which did not capture peak GC concentration, since in Rhesus macaques, similar to humans, this occurs at around 8:00. Hahn-Holbrook et al.[8] collected a single breastmilk sample within a wide time window (11:30-16:00), which also did not capture peak concentrations. Analyses were corrected for time of collection, but it has previously been shown that correcting for time of sampling cannot account for all the variability observed in cortisol levels.[9,22] Second, in our study GC concentrations were determined by LC-MS/MS, which has been shown to be more sensitive and reliable than radioimmunoassay and chemiluminescent immunoassay,[23] which were used by Hinde et al. and Hahn-Holbrook et al. respectively.

Lastly, it has previously been shown that increases in fat mass specifically are associated with mid-childhood overweight and obesity[24]. Therefore, in this study, body composition was measured by ADP, which is able to differentiate between fat mass and fat free mass. In contrast, weight gain and changes in BMI were used as outcomes measured by Hinde et al.[7] and Hahn-Holbrook et al.,[8] respectively, both of which are less precise methods to determine body composition. Our more detailed methods when measuring GC concentrations in breastmilk as well as when determining body composition might therefore have led to more accurate conclusions.

Alternatively, the absence of associations might be due to the small sample size in this study, especially compared to Hinde et al.,[7] who included 108 mother-infant pairs, resulting in more power to detect small differences. However, this should be balanced against the use of of air-displacement plethysmography in this study, which is superior to weight gain for the assessment of body composition. Additionally, our follow-up until the age of 3 months was rather short. In contrast, follow-up took place up to 2 years of age in Hahn-Holbrook et al.'s study.[8] It is therefore possible that effects of GCs in breastmilk might only be noticeable at a later age. On the other hand, an increasing number of nutritional, life-style and family factors determine body composition with advancing age, and it is therefore progressively more difficult to determine to what extent breastmilk cortisol explains BMI gains.

This study has several strengths and limitations. This was the first study to assess the association between GC rhythmicity in breastmilk and body composition in the offspring. Body composition and GC rhythmicity were analyzed in detail, respectively by use of ADP and by measuring both cortisol and cortisone in breastmilk using samples that were collected over a 24-hour period. Cortisone concentrations have been shown to be more reliable than cortisol measurements, at least in saliva and hair.[25,26] This is possibly due to the local conversion of cortisol by 11β-hydroxysteroid dehydrogenase type 2, which leads to higher concentrations of cortisone.[27] However, this study also has its limitations. The sample size of this study was relatively small, and it is therefore possible that modest effects could not be detected. It was also not possible to correct for all potential confounders. However, many confounders were considered, and the three variables with the largest confounding effect were included in the final model, which did not change the results compared to unadjusted analyses. It is therefore unlikely that adjusting for more variables would have altered the results. Second, the follow-up in this study was relatively short, and it is therefore possible that breastmilk GC rhythmicity has an effect only noticeable at a later age. Additionally, a selection bias cannot be ruled out, since we did not collect data on mothers who were eligible for inclusion but decided against participation and since we included mother-infant pairs at a (tertiary) hospital. The study population might therefore not reflect the general population; for example, 51% of the mothers have birth via Caesarian section, compared to approximately 17%

in the general population.[28] Lastly, the interplay between GCs and infant body composition is complex, and could be moderated by, for example, exposure to GCs and other conditions in utero, the number of feeds per day, and the extra-uterine environment of the infants, including synchrony in mother-infant interactions as well as stressful events. Unfortunately, we were not able to take these factors into account.

CONCLUSIONS

This study did not find an association between breastmilk GC rhythmicity at 1 month and growth trajectories as well as body composition of the offspring at age 3 months. Therefore, this study suggests that previous observations linking breastmilk cortisol to changes in infant weight might be flawed by the lack of serial cortisol measurements and detailed information on body composition.

REFERENCES

1. Barker DJ, Osmond C. Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. Lancet 1986; 1:1077-1081

2. Finken MJ, van der Voorn B, Heijboer AC, de Waard M, van Goudoever JB, Rotteveel J. Glucocorticoid Programming in Very Preterm Birth. Horm Res Paediatr 2016; 85:221-231

3. Rosmond R, Bjorntorp P. The hypothalamic-pituitary-adrenal axis activity as a predictor of cardiovascular disease, type 2 diabetes and stroke. J Intern Med 2000; 247:188-197

4. Drake AJ, Tang JI, Nyirenda MJ. Mechanisms underlying the role of glucocorticoids in the early life programming of adult disease. Clin Sci (Lond) 2007; 113:219-232

5. Entringer S. Impact of stress and stress physiology during pregnancy on child metabolic function and obesity risk. Curr Opin Clin Nutr Metab Care 2013; 16:320-327

6. Angelucci L, Patacchioli FR, Scaccianoce S, Di Sciullo A, Cardillo A, Maccari S. A model for later-life effects of perinatal drug exposure: maternal hormone mediation. Neurobehav Toxicol Teratol 1985; 7:511-517

7. Hinde K, Skibiel AL, Foster AB, Rosso LD, Mendoza SP, Capitanio JP. Cortisol in mother's milk across lactation reflects maternal life history and predicts infant temperament. Behavioral Ecology 2015; 26:269-281

8. Hahn-Holbrook J, Le TB, Chung A, Davis EP, Glynn LM. Cortisol in human milk predicts child BMI. Obesity (Silver Spring) 2016; 24:2471-2474

9. van der Voorn B, de Waard M, van Goudoever JB, Rotteveel J, Heijboer AC, Finken MJ. Breast-Milk Cortisol and Cortisone Concentrations Follow the Diurnal Rhythm of Maternal Hypothalamus-Pituitary-Adrenal Axis Activity. J Nutr 2016; 146:2174-2179

10. Adam EK, Quinn ME, Tavernier R, McQuillan MT, Dahlke KA, Gilbert KE. Diurnal cortisol slopes and mental and physical health outcomes: A systematic review and meta-analysis. Psychoneuroendocrinology 2017; 83:25-41

11. Ruttle PL, Javaras KN, Klein MH, Armstrong JM, Burk LR, Essex MJ. Concurrent and longitudinal associations between diurnal cortisol and body mass index across adolescence. J Adolesc Health 2013; 52:731-737

12. Wirix AJ, Finken MJ, von Rosenstiel-Jadoul IA, Heijboer AC, Nauta J, Groothoff JW, Chinapaw MJ, Kist-van Holthe JE. Is There an Association Between Cortisol and Hypertension in Overweight or Obese Children? J Clin Res Pediatr Endocrinol 2017; 9:344-349

13. Patacchioli F, Cigliana G. Maternal plasma and milk free cortisol during the first 3 days of breast-feeding following spontaneous delivery or elective cesarean section. Gynecolog Obstet Invest 1992; 34:159-163

14. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983; 67:361-370

15. Ma G, Yao M, Liu Y, Lin A, Zou H, Urlando A, Wong WW, Nommsen-Rivers L, Dewey KG. Validation of a new pediatric air-displacement plethysmograph for assessing body composition in infants. Am J Clin Nutr 2004; 79:653-660

16. van der Voorn B, Martens F, Peppelman NS, Rotteveel J, Blankenstein MA, Finken MJ, Heijboer AC. Determination of cortisol and cortisone in human mother's milk. Clin Chim Acta 2015; 444:154-155

17. Pruessner JC, Kirschbaum C, Meinlschmid G, Hellhammer DH. Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. Psychoneuroendocrinology 2003; 28:916-931

18. Kyle UG, Schutz Y, Dupertuis YM, Pichard C. Body composition interpretation. Contributions of the fat-free mass index and the body fat mass index. Nutrition 2003; 19:597-604

19. TNO. De Vijfde Landelijke Groeistudie. 2010;

20. Schonbeck Y, Talma H, van Dommelen P, Bakker B, Buitendijk SE, HiraSing RA, van Buuren S. The world's tallest nation has stopped growing taller: the height of Dutch children from 1955 to 2009. Pediatr Res 2013; 73:371-377

21. Figueiras A, Domenech-Massons JM, Cadarso C. Regression models: calculating the confidence interval of effects in the presence of interactions. Stat Med 1998; 17:2099-2105

22. de Weerth C, Zijl RH, Buitelaar JK. Development of cortisol circadian rhythm in infancy. Early Hum Dev 2003; 73:39-52

23. Ackermans MT, Endert E. LC-MS/MS in endocrinology: what is the profit of the last 5 years? Bio-analysis 2014; 6:43-57

24. Koontz MB, Gunzler DD, Presley L, Catalano PM. Longitudinal changes in infant body composition: association with childhood obesity. Pediatr Obes 2014; 9:e141-144

25. Blair J, Adaway J, Keevil B, Ross R. Salivary cortisol and cortisone in the clinical setting. Curr Opin Endocrinol Diabetes Obes 2017; 24:161-168

26. Savas M, Wester VL, de Rijke YB, Rubinstein G, Zopp S, Dorst K, van den Berg SAA, Beuschlein F, Feelders RA, Reincke M, van Rossum EFC. Hair glucocorticoids as biomarker for endogenous Cushing's syndrome: validation in two independent cohorts. Neuroendocrinology 2019;

27. Smith RE, Maguire JA, Stein-Oakley AN, Sasano H, Takahashi K, Fukushima K, Krozowski ZS. Localization of 11 beta-hydroxysteroid dehydrogenase type II in human epithelial tissues. J Clin Endocrinol Metab 1996; 81:3244-3248

28. Macfarlane AJ, Blondel B, Mohangoo AD, Cuttini M, Nijhuis J, Novak Z, Olafsdottir HS, Zeitlin J, Euro-Peristat Scientific C. Wide differences in mode of delivery within Europe: risk-stratified analyses of aggregated routine data from the Euro-Peristat study. BJOG 2016; 123:559-568

Supplementary Table 1: Cortisol and cortisone concentrations in breastmilk in 4-hour intervals.

	Cortisol (nmol/L)	Cortisone (nmol/L)
0:00-4:00	4.1±5.5	15.1±11.8
4:00-8:00	11.6±8.7	29.7±12.8
8:00-12:00	8.2±6.5	27.9±8.6
12:00-16:00	4.4±2.9	21.3±6.8
16:00-20:00	2.1±1.4	13.0±6.1
20:00-24:00	2.1±3.9	10.4±9.2

Values represent mean±SD

EVA MOUTON 2019

8
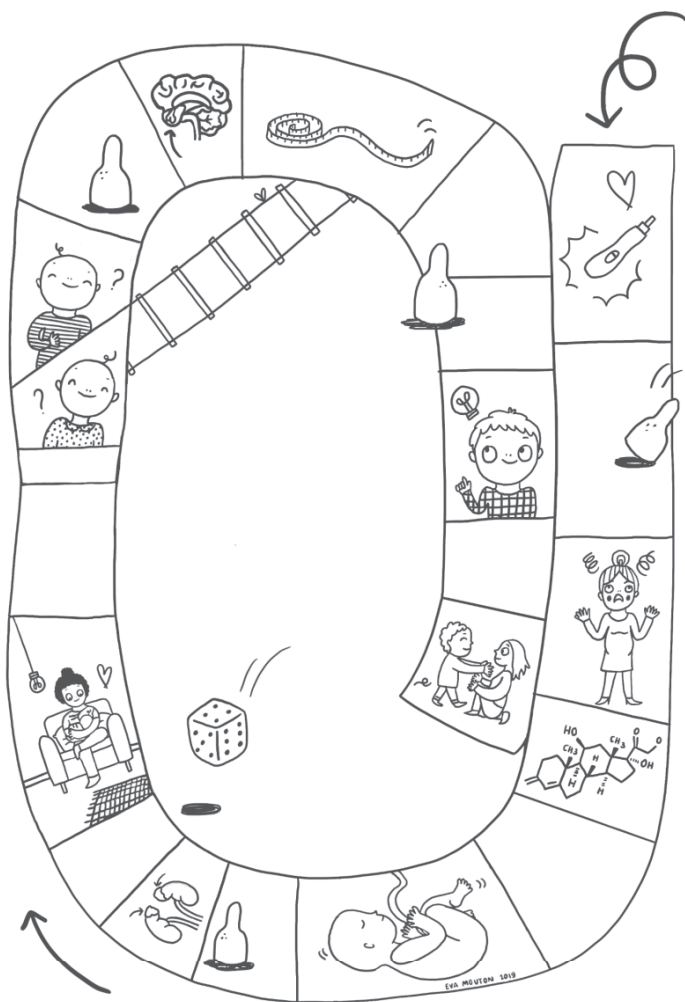
Diurnal rhythmicity in breast-milk glucocorticoids and infant behavior and sleep at age three months

Alyssa A. Toorop*,
Bibian van der Voorn*,
Jonneke J. Hollanders,
Lisette R. Dijkstra,
Koert M. Dolman,
Annemieke C. Heijboer,
Joost Rotteveel,
Adriaan Honig,
Martijn J.J. Finken

* Authors contributed equally to this manuscript

ABSTRACT

Purpose

In previous studies, associations between breast-milk cortisol levels obtained on one occasion and infant neurodevelopment were demonstrated. However, more recent evidence indicates that breast-milk cortisol and cortisone concentrations follow a diurnal rhythm, and therefore these levels fluctuate throughout the day. We studied associations between breast-milk glucocorticoid (GC) rhythmicity and infant behavior and sleep.

Methods

We included 59 mothers, and their infants, of whom 17 had consulted an expert center during pregnancy for an increased risk of psychological distress. At 1 month postpartum, breast milk was sampled (on average 6 times) over a 24h period for assessment of cortisol and cortisone using LC-MS/MS, and experienced maternal distress was assessed using the Hospital Anxiety and Depression Scale questionnaire. At 3 months postpartum, infant behavior was assessed using the Infant Behavior Questionnaire, and infant sleep was quantified by questionnaire. Associations between breast-milk GC rhythmicity (maximum, delta, and Area Under the Curve [AUC]) and infant behavior and sleep were tested with linear regression analyses.

Results

No consistent associations between breast-milk GC rhythmicity or exposure and infant behavior or sleep were found.

Conclusions

Breast-milk GC rhythmicity or exposure at 1 month postpartum was not associated with infant behavior or sleep at the age of 3 months. Findings from previous studies linking breast-milk cortisol to infant neurodevelopment might be biased by the lack of GC measurements across the full diurnal cycle, and should therefore be interpreted with caution.

INTRODUCTION

Approximately 15% of pregnant women in Western countries are diagnosed with psychiatric conditions.[1] Depressive and anxiety disorders are associated with alterations in hypothalamic-pituitary-adrenal (HPA-) axis activity, such as a lower morning peak or less diurnal variability in cortisol level.[2,3] Maternal glucocorticoids (GCs) that cross the placenta may influence the fetal HPA-axis, possibly through alterations in the expression of GC and mineralocorticoid receptors in the developing hippocampus and amygdala.[4] In humans, fetal exposure to maternal depression or anxiety symptoms was associated with a more fearful temperament and disorganized sleep in infancy, along with a flattened cortisol rhythm.[5-7] The results of these studies suggest that fetal GC exposure may alter neurodevelopment and HPA-axis settings in later life.[8-10]

The development of an adult-type diurnal cortisol rhythm, characterized by cortisol concentrations that are higher in the morning than in the evening, is thought to start at approximately 1 month of age and continues to develop during the first year of life.[11] It has been hypothesized that the development of HPA-axis rhythmicity may serve as a modulator for the development of behavioral rhythms, such as the sleep-wake cycle.[12,13] Previous research has demonstrated that a diurnal GC rhythm develops before sleep rhythmicity is established around the age of 2-4 months.[14,15] Multiple factors may be involved in the development of HPA-axis rhythmicity, such as environmental time cues (e.g., daylight) and maternal care.[15] In addition, it has been proposed that non-nutritive bioactive compounds in breast milk might be involved in the development of sleep regulation in infants.[16]

Animal studies have shown that GCs in breast milk are able to cross the intestinal wall and to enter the circulation in offspring.[17,18] Among Rhesus monkeys, offspring exposed to higher levels of breast-milk cortisol were found to exhibit a more nervous, less confident behavior and impulsivity, albeit with few gender-specific differences.[19-21] In rats, exposure to physiological ranges of ingested GCs was associated with reduction of fearfulness and stress-induced corticosterone secretion throughout the lifespan.[22] In breastfed human infants, exposure to maternal cortisol has also been associated with behavior.[23-25] Two studies showed that higher breast-milk cortisol was associated with negative affectivity among girls, but not among boys.[24,26] Another study showed that higher plasma cortisol, which has strong correlation with breast-milk cortisol,[27] was associated with increased infant fear behavior [23]. However, none of these studies collected samples multiple times during the day or around the morning peak of cortisol secretion, in spite of evidence indicating that breast-milk GCs follow the diurnal rhythm of maternal HPA-axis activity.[27] Some of these studies statistically adjusted for inter-individual differences in collection time, which assumes that the cortisol slope barely differs between subjects. However, it has been demonstrated that *post-hoc* statistical

correction for time of sampling may not be able to provide an adequate representation of an individual's HPA-axis dynamics.[27,28]

The aim of this study was to investigate associations between exposure to breast-milk GCs over a 24-hr period at 1 month postpartum, and infant behavior and sleep at 3 months postpartum. In this study, we oversampled mothers at risk of psychological distress during and after pregnancy in an attempt to capture a wide range of maternal HPA-axis activity, since depression and anxiety have previously been associated with GC rhythmicity.[2,3]

METHODS

Participants

From March 2016 to July 2017, mothers were approached within the first days after delivery at the maternity wards of the Amsterdam University Medical Center, location VUmc (Group 1, n=42), and the OLVG hospital (Group 2, n=17), The Netherlands. Mothers included at the OLVG hospital had an increased risk of psychological distress and therefore consulted the Psychiatric Obstetric Pediatric (POP) outpatient clinic during pregnancy. Breastfeeding mothers of infants born after full-term gestation (37-42 weeks of pregnancy) with a birth weight appropriate for gestational age (i.e., between -2 and 2 SD score) were eligible for inclusion. Exclusion criteria were preeclampsia/HELLP, multiple pregnancy, consumption of >7 IU of alcohol per week, fever >38.5 °C at the time of sampling, and major congenital anomalies. Additionally, mothers who used drugs other than 'over the counter' drugs were excluded, with the exception of Selective Antidepressants (SADs) use for mothers included at the OLVG. Approval of the Medical Ethics Committee of the Amsterdam University Medical Center, location VUmc was obtained (*protocol number 2015.524),* and written informed consent was obtained from all participating mothers.

Data collection

Infant and maternal characteristics
During the first days postpartum, maternal and infant characteristics were obtained by questionnaire (Table 1). At the time of milk sample collection (1 month postpartum), mothers were asked to fill in the Hospital Anxiety and Depression Scale (HADS) for assessment of maternal psychological distress experienced during the past 2 weeks.[29] The HADS contains 14 items, including seven items for depressive symptoms (Hospital Depression Subscale [HDS]) and seven items for anxiety symptoms (Hospital Anxiety Subscale [HAS]). Items are scored as 0-3, and a score ≥8 on either subscale indicates clinically relevant depression and/or anxiety symptoms. Accordingly, we defined increased maternal stress as HDS score and/or HAS score ≥8.

Table 1: Maternal and infant characteristics of participants[1]

Maternal characteristics	Group 1[2] (n=42)	Group 2[3] (n=17)
Maternal age, yrs	33.6 ± 4.7	31.6 ± 4.7
Maternal BMI, kg/m[2]	22.3 ± 2.8	22.9 ± 2.2
Social Economic Status (SES)[4]	0.6 ± 1.2	0.4 ± 1.3
Caucasian ethnicity	34 (81)	15 (88)
Primiparity	23 (55)	7 (41)
Selective antidepressant use	0	12 (71)*
HAS/HDS score≥8 1 mo. pp (n=58)[5]	6 (15)	6 (35)
Neonatal characteristics		
Male gender	25 (60)	11 (65)
Birth weight (grams)	3389 ± 39	3561 ± 498
Gestational age (weeks)	39.1 ± 1.1	39.9 ± 1.3*
Vaginal birth (n=58)[5]	20 (49)	13 (77)
≥80% breast milk 3 mo. pp (n=58)[5]	37 (90)	14 (82)

[1] Values are presented as means ± SD or frequencies (%).
[2] Mothers included at the maternity ward of the Amsterdam UMC.
[3] Mothers included at the Psychiatry Obstetric Pediatric (POP) expert center, OLVG hospital
[4] Z-score based on average income, % low income, % low-skilled and % unemployed civilians per postal code area, based on data from the Dutch Social Cultural Project office [2014, The Netherlands]).
[5] Three participating mothers in group 1 did not provide these data
*p<.05

Breast-milk sample collection

At 30±5 days postpartum, 1-2mL of breast milk was collected before every feed over a 24h-period, either manually or with a breast pump. Mothers were requested to report the exact time of the sample collection, since they were breastfeeding their child on demand. Mothers were asked to abstain from alcohol at least one day before and during the sample collection. Following collection, samples were stored in plastic tubes at -20°C until they were thawed for analysis.

Infant Behavior and Sleep

At 3 months (±2 wks) postpartum, mothers were asked to fill in the Infant Behavior Questionnaire (IBQ) and a questionnaire for the quantification of infant sleep. The IBQ is a validated instrument for the assessment of temperament in infants aged 3 months to 1 year.[30,31] The original IBQ contains 94 items on six scales of temperament dimensions (distress to limitations, approach to novel stimuli, soothability, duration of orienting, smiling and laughter, and activity) that show considerable stability over time.[32] The IBQ assesses behavior on a 7-point Likert scale, with answers ranging from 'never' to 'always', or 'does not apply'. To minimize recall bias, the answers pertain to the infant's behavior over the past 1-2 weeks. The mean scoring represents the outcome for each dimension

separately. The sleep questionnaire (see Supplementary File 1) included the total hours of night-time and day-time sleep, the number of daytime naps, and the number of nights with more than 6 hours of consecutive sleep during one week.

Determination of cortisol and cortisone levels in milk

An isotope dilution liquid chromatography-tandem mass spectrometry (LC-MS/MS) method was used to assess cortisol and cortisone concentrations in milk, as described previously.[33] In short, milk samples were washed with hexane after adding internal standards to the samples ($^{13}C_3$ labeled cortisol and cortisone). Samples were extracted using Isolute plates (Biotage, Uppsala, Sweden) and analyzed by LC-MS/MS (Acquity with Quattro Premier XE, Milford MA, USA, Waters Corporation). For cortisol, the intra-assay coefficient of variation (CV) was 4 to 5%, and for cortisone it was 5% at different levels. For both cortisol and cortisone, the inter-assay CV was <9%, and the Lower Limit of Quantitation was 0.5 nmol/L.

Data analyses

Breast-milk GC parameters were recorded over a period of 24 hours, and were defined as: maximum (i.e., the maximum measured concentration), delta (Δ, i.e., the difference between the maximum and minimum measured concentrations), and Area Under the Curve increase (AUCi) and ground (AUCg) per hour collection. AUCi/h was used as an index for GC rhythmicity. AUCg/h was used to reflect total breast-milk GC exposure. Both AUCi/h and AUCg/h were calculated using the trapezoid rule.[34] Participants who did not provide morning samples (between 05:00-10:00 a.m.) or with <8 hours of total sample collection time were excluded from the analyses.

   Maternal and infant characteristics were compared between mothers who had no increased risk of psychological distress (group 1) and those who had (group 2), using independent samples T-tests and Chi-square tests (Table 1). Of all characteristics, only maternal antidepressant use and gestational age were significantly different between the groups. Although one-third of the mothers monitored at the POP outpatient clinic reported increased psychological distress, no statistically significant differences were found between group 1 and group 2 with regard to HADS-score as well as breast-milk GC parameters. Therefore, all mother-infant pairs were analyzed as one group, while considering gestational age and maternal use of antidepressants as potential confounders. Subsequently, associations between breast-milk GC rhythmicity or total GC exposure, and IBQ scores or infant sleep were tested using linear or logistic regression, as appropriate. Second, we performed multivariate analyses correcting for a set of potential confounders based on previous literature (infant gender, socio-economic status and maternal stress) or statistical impact (maternal use of antidepressants and gestational age).[23,24,26] Results were presented as beta or Odds Ratio (OR) [95% confident interval (CI)]. A p-value <.05 was considered statistically significant.

RESULTS

Figure 1 shows the stepwise inclusion procedure for the study. A total of 303 mothers were approached, of whom 110 gave written informed consent. Of these, 59 completed the study. Main reasons for drop-out were switching to formula feeding or withdrawal of consent. Characteristics of the participating mothers and their infants are presented in Table 1. Breast-milk GC parameters, IBQ scores and sleep outcomes are shown in Table 2. Mothers collected on average 6 (Range: 4 to 8) milk samples over a 24-hour period.

Breast-milk glucocorticoids and infant behavior

Tables 3 and 4 show the multivariate associations between breast-milk GC parameters and infant behavior. An association between total breast-milk cortisol exposure and more infant soothability was found (β= 0.15 [0.02 to 0.29], p<.05). Other breast-milk GC parameters were not associated with infant behavior.
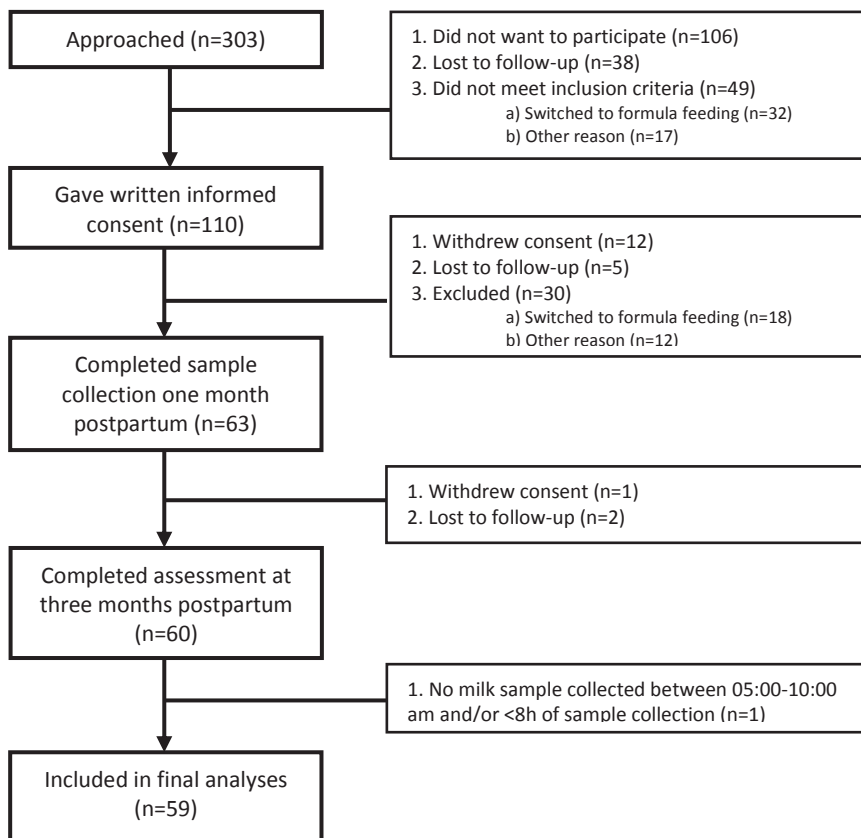
Figure 1: Flowchart of the inclusion of mother-infant pairs

Table 2: Measurements by group[1]

	Group 1[2] (n=42)	Group 2[3] (n=17)	Total (n=59)
Breast-milk GC levels			
Cortisol maximum (nmol/L) *Average time of maximum*	15.8 ±8.9 [9.2 to 20.8] *8:15h*	13.6 ±8.2 [8.8 to 16.7] *6:45h*	15.2 ±8.7 [9.2 to 19.2] *7:45h*
Δcortisol (nmol/L)	14.7 ±9.0 [8.0 to 19.0]	12.5 ±8.3 [8.3 to 16.0]	14.2 ±8.8 [8.1 to 18.4]
AUCi/h of cortisol in 24h	4.1 ±2.3 [2.6 to 5.0]	3.0 ±1.5 [1.9 to 3.8]	3.7 ±2.2 [2.2 to 4.7]
AUCg/h of cortisol in 24h	5.2 ± 2.4 [3.6 to 6.2] *	3.7 ± 1.5 [2.5 to 4.7]	4.8 ± 2.3 [3.3 to 5.8]
Cortisone maximum (nmol/L) *Average time of maximum*	36.0±10.6 [27.7 to 42.7] *08:00h*	33.8 ±9.0 [26.6 to 40.2] *8:30h*	35.3 ±10.1 [27.4 to 42.3] *8:15h*
Δcortisone (nmol/L)	28.9 ±10.1 [23.5 to 35.6]	27.8 ±8.7 [22.8 to 32.2]	28.6 ±9.7 [23.3 to 34.2]
AUCi/h of cortisone in 24h	12.2 ±4.4 [10.0 to 14.9]	10.8 ±4.4 [8.0 to 14.5]	11.8±4.4 [9.4 to 14.7]
AUCg/h of cortisone in 24h	19.2 ± 5.7 [16.0 to 21.8]	16.8 ± 4.5 [13.2 to 19.8]	18.5 ± 5.4 [14.8 to 21.1]
IBQ domain [4]			
Activity	3.2 ± 0.9	3.0 ± 0.7	3.1 ± 0.9
Distress to limitations	3.3 ± 0.9	3.2 ± 0.9	3.3 ± 0.9
Approach to novel stimuli	2.0 ± 1.0	2.1 ± 0.7	2.0 ±0.9
Duration of orienting	3.6 ± 1.2	4.0 ± 0.9	3.7 ± 1.2
Smiling and laughter	4.3 ± 1.0	4.5 ± 1.0	4.4 ± 1.0
Soothability	5.0 ± 1.2	4.9 ± 1.1	5.0 ± 1.1
Infant sleep parameters			
Hours of night-time sleep	8.2 ± 2.3	8.6 ± 2.3	8.3 ± 2.3
Hours of daytime sleep	4.3 ± 1.9	4.5 ±1.6	4.4 ± 1.8
Number of naps during daytime (n=58)	3.3 ± 0.8	3.3 ± 1.1	1.5 ± 1.0
Number of nights per week with >6h sleep	4.1 ± 3.0	4.4 ± 2.7	4.2 ± 2.9

[1] Values are presented as means ± SD and range. Maximum= maximum value during 24h period. Δ= Difference between maximum and minimum value during 24h period. AUCi/h= Area under the curve for increase per hour collection. AUCg/h= Area under the curve with respect to ground per hour collection. Groups were compared using the independent samples T-test.

[2] Mothers included at the maternity ward of the Amsterdam UMC.

[3] Mothers included at the Psychiatry Obstetric Pediatric (POP) expert center, OLVG hospital.

[4] IBQ= Infant Behavior Questionnaire.

*p<.05

Breast-milk glucocorticoids and infant sleep

Tables 3 and 4 show the multivariate associations between breast-milk GC parameters and infant sleep. A positive association between breast-milk delta cortisone and infant sleep at night-time was found (β= 0.07 [0.01 to 0.20], p<.05). Other breast-milk GC parameters were not associated with infant sleep.

Table 3: Associations between breast-milk cortisol parameters and infant behavior and sleep[1]

Infant behavior outcomes	Maximum cortisol (nmol/L)	Δcortisol (nmol/L)	Cortisol AUCi/h (in 24h)	Cortisol AUCg/h (in 24h)
Distress to limitations	-0.02 [-0.05 to 0.01]	-0.02 [-0.04 to 0.01]	-0.02 [-0.13 to 0.09]	-0.03 [-0.13 to 0.08]
Approach to novel stimuli	-0.02 [-0.05 to 0.01]	-0.02 [-0.05 to 0.01]	-0.05 [-0.16 to 0.06]	-0.03 [-0.13 to 0.07]
Soothability	0.02 [-0.02 to 0.04]	0.02 [-0.02 to 0.05]	0.13 [-0.02 to 0.27]	0.15 [0.02 to 0.28] *
Smiling and laughter	0.01 [-0.02 to 0.04]	0.01 [-0.02 to 0.04]	0.005 [-0.11 to 0.12]	0.02 [-0.10 to 0.13]
Duration of orienting	-0.004 [-0.04 to 0.03]	-0.01 [-0.05 to 0.03]	-0.04 [-0.19 to 0.11]	0.03 [-0.11 to 0.17]
Activity	0.01 [-0.02 to 0.04]	0.01 [-0.02 to 0.03]	0.02 [-0.09 to 0.13]	0.05 [-0.05 to 0.15]
Infant sleep outcomes				
Hours of night-time sleep	0.04 [-0.03 to 0.11]	0.05 [-0.02 to 0.12]	0.07 [-0.22 to 0.35]	-0.03 [-0.29 to 0.23]
Hours of daytime sleep	0.01 [-0.05 to 0.07]	0.01 [-0.05 to 0.07]	0.01 [-0.23 to 0.25]	0.01 [-0.23 to 0.22]
Number of nights per week with >6h sleep[2] (0-4 days=0, 5-7 days=1)	0.98 [0.92 to 1.04]	0.98 [0.92 to 1.05]	0.92 [0.71 to 1.20]	0.87 [0.67 to 1.12]
Number of naps during daytime[2] (<3 naps=0, ≥3 naps=1)	0.93 [0.86 to 1.01]	0.94 [0.88 to 1.02]	0.87 [0.66 to 1.15]	0.77 [0.57 to 1.04]

[1]Values are presented as betas or ORs [95% CI] and adjusted for gender, socio-economic status, elevated HADS score at one mo, pp (HAS/HDS score ≥8), maternal use of antidepressants and gestational age.

[2]OR.

Maximum= maximum value during 24h period. Δ= Difference between maximum and minimum value during 24h period. AUCi/h= Area under the curve for increase per hour collection. AUCg/h= Area under the curve with respect to ground per hour collection. *p<0.05.

Table 4: Associations between breast-milk cortisone parameters and infant behavior and sleep[1]

Infant behavior outcomes	Maximum cortisone (nmol/L)	Δcortisone (nmol/L)	Cortisone AUCi/h (in 24h)	Cortisone AUCg/h (in 24h)
Distress to limitations	-0.01 [-0.03 to 0.02]	-0.01 [-0.04 to 0.01]	-0.03 [-0.09 to 0.03]	-0.01 [-0.05 to 0.04]
Approach to novel stimuli	-0.02 [-0.04 to 0.01]	-0.02 [-0.05 to 0.003]	-0.03 [-0.09 to 0.03]	-0.004 [-0.05 to 0.04]
Soothability	0.003 [-0.03 to 0.04]	0.001 [-0.04 to 0.04]	0.02 [-0.06 to 0.10]	0.02 [-0.04 to 0.08]
Smiling and laughter	-0.01 [-0.04 to 0.02]	-0.01 [-0.04 to 0.02]	-0.04 [-0.10 to 0.02]	-0.03 [-0.08 to 0.02]
Duration of orienting	0.002 [-0.03 to 0.04]	-0.01 [-0.04 to 0.03]	-0.04 [-0.12 to 0.03]	-0.001 [-0.06 to 0.06]
Activity	0.001 [-0.02 to 0.03]	-0.006 [-0.03 to 0.02]	-0.02 [-0.08 to 0.03]	0.01 [-0.04 to 0.05]
Infant sleep outcomes				
Hours of night-time sleep	0.06 [-0.001 to 0.12]	0.07 [0.006 to 0.14] *	0.11 [-0.04 to 0.25]	0.06 [-0.06 to 0.18]
Hours of daytime sleep	-0.01 [-0.06 to 0.05]	0.001 [-0.06 to 0.06]	0.001 [-0.13 to 0.13]	-0.03 [-0.13 to 0.07]
Number of nights per week with >6h sleep[2] (0-4 days=0, 5-7 days=1)	0.98 [0.92 to 1.04]	0.99 [0.94 to 1.06]	0.94 [0.82 to 1.08]	0.89 [0.79 to 1.0]
Number of naps during daytime[2] (<3 naps=0, ≥3 naps=1)	0.94 [0.88 to 1.00]	0.97 [0.91 to 1.03]	1.00 [0.87 to 1.15]	0.88 [0.78 to 1.01]

[1] Values are presented as betas or ORs [95% CI] adjusted for gender, socio-economic status, elevated HADS score at one mo. pp (HAS/HDS score ≥8), maternal use of antidepressants and gestational age.

[2] OR. Maximum= maximum value during 24h period.

Δ= Difference between maximum and minimum value during 24h period. AUCi/h= Area under the curve for increase per hour collection. AUCg/h= Area under the curve with respect to ground per hour collection.

*p<0.05.

DISCUSSION

In this study, with few exceptions, no associations were found between breast-milk GC rhythmicity and total exposure at 1 month postpartum and infant behavior or sleep at age 3 months. Therefore, our study could not confirm previous observations in animals and humans.[19,20,24,25]

The results of our study differed from those of previous studies. Importantly, cortisol sampling in the previous studies did not take the diurnal rhythm of breastmilk GCs into account. Although some of these studies corrected analyses for time of collection,[23,24,26] it has previously been demonstrated that correcting for sampling time cannot account for variability in cortisol levels over time.[27,35] There is no doubt that sampling fluctuation has a major impact on the interpretation of HPA-axis dynamics, which plausibly leads to false conclusions when only cross-sectional breast-milk GC levels are studied; e.g., an outcome might be associated with the height of the cortisol level, whereas it actually reflects the time of sampling.[28]

There is some evidence suggesting that breast-milk GCs influence infant neurodevelopment in a sex-specific manner. Among Rhesus monkeys, the associations between higher milk cortisol levels and more nervous, less confident temperament in offspring differed between males and females in such a way that male offspring appeared to be more sensitive to cortisol increments over time and female offspring to the absolute cortisol concentration.[20] In humans, milk cortisol was positively associated with negative affectivity among girls, but not among boys.[24,26] This might be attributed to differences in the developmental timing of GC sensitivity between males and females. Indeed, studies in rodents showed that forebrain and hippocampal GC receptor expression patterns developed in a sex-specific manner.[22,36,37] Due to the small sample size in our study, we were unable to perform sex-specific analyses.

This study has several strengths. It is the first to assess the association between GC diurnal rhythmicity in breast milk and infant behavior and sleep. Moreover, measurement of cortisone along with cortisol in breast milk carries the advantage of having a more precise estimate of GC exposure. Epithelial tissues have been demonstrated to harbor 11β HSD type 2, which converts cortisol into cortisone upon entrapment.[38] Consequently, cortisone may be a more accurate marker of the circulating cortisol level than cortisol itself, at least in saliva and hair.[38-40] This is corroborated by observations demonstrating that cortisone is less likely to have concentrations below the lower limit of quantification at the nadir.[33] Moreover, the presence of 11β-reductase activity in infants implies that cortisol can be regenerated from cortisone. Breast-milk cortisone may thus become a part of the biologically available GC pool.[35,41] Another strength is the use of mass spectrometry for the measurements of cortisol and cortisone. LC-MS/MS is superior to the immunoassays that were used by previous studies in terms of specificity.[42]

However, this study also has its limitations. First, the sample size of our study, although comparable to previous studies in this field,[23-26] might have been too small to detect subtle differences or to stratify for sex. However, this limitation must be balanced against having multiple measurements across the diurnal cycle. Second, we cannot exclude the possibility that distressed mothers were less likely to participate. Non-participants did not sign informed consent, and non-response analyses could therefore not be performed. This might offer an explanation for the observation that experienced maternal stress did not differ between the groups. Third, in view of the limited number of participants, it was not possible to correct for all potential confounders, such as parental temper and other environmental factors that might interfere with infant behavior or sleep. Fourth, infant temperament and sleep were self-reported by the mothers, while parenting behavior was not assessed. Even though behavior is still thought to be best reported by the infant's primary caregiver, distressed mothers may rate their infant's behavior as more difficult,[43-45] To account for this phenomenon, all outcomes were corrected for elevated HADS scores.[46,47]

In conclusion, in our study breast-milk GC rhythmicity at 1 month postpartum was not associated with infant behavior or sleep at 3 months postpartum. Therefore, this study suggests that findings from previous studies linking breast-milk cortisol to infant neurodevelopment might be biased by the lack of GC measurements across the full diurnal cycle or the choice for less reliable immunoassay measurements.

REFERENCES

1.   Andersson L, Sundstrom-Poromaa I, Bixo M, Wulff M, Bondestam K, aStrom M. Point prevalence of psychiatric disorders during the second trimester of pregnancy: a population-based study. Am J Obstet Gynecol 2003; 189:148-154

2.   Jarcho MR, Slavich GM, Tylova-Stein H, Wolkowitz OM, Burke HM. Dysregulated diurnal cortisol pattern is associated with glucocorticoid resistance in women with major depressive disorder. Biol Psychol 2013; 93:150-158

3.   Vreeburg SA, Hoogendijk WJ, DeRijk RH, van Dyck R, Smit JH, Zitman FG, Penninx BW. Salivary cortisol levels and the 2-year course of depressive and anxiety disorders. Psychoneuroendocrinology 2013; 38:1494-1502

4.   Seckl JR. Prenatal glucocorticoids and long-term programming. Eur J Endocrinol 2004; 151 Suppl 3:U49-62

5.   Davis EP, Glynn LM, Schetter CD, Hobel C, Chicz-Demet A, Sandman CA. Prenatal exposure to maternal depression and cortisol influences infant temperament. J Am Acad Child Adolesc Psychiatry 2007; 46:737-746

6.   Field T, Diego M, Hernandez-Reif M, Figueiredo B, Schanberg S, Kuhn C. Sleep disturbances in depressed pregnant women and their newborns. Infant Behav Dev 2007; 30:127-133

7.   Luijk MP, Saridjan N, Tharner A, van Ijzendoorn MH, Bakermans-Kranenburg MJ, Jaddoe VW, Hofman A, Verhulst FC, Tiemeier H. Attachment, depression, and cortisol: Deviant patterns in insecure-resistant and disorganized infants. Dev Psychobiol 2010; 52:441-452

8.   Glover V, O'Connor TG, O'Donnell K. Prenatal stress and the programming of the HPA axis. Neurosci Biobehav Rev 2010; 35:17-22

9.   Zijlmans MA, Riksen-Walraven JM, de Weerth C. Associations between maternal prenatal cortisol concentrations and child outcomes: A systematic review. Neurosci Biobehav Rev 2015; 53:1-24

10.   Brunton PJ. Programming the brain and behaviour by early-life stress: a focus on neuroactive steroids. J Neuroendocrinol 2015; 27:468-480

11.   Ivars K, Nelson N, Theodorsson A, Theodorsson E, Strom JO, Morelius E. Development of Salivary Cortisol Circadian Rhythm and Reference Intervals in Full-Term Infants. PLoS One 2015; 10:e0129502

12.   de Weerth C, Zijl RH, Buitelaar JK. Development of cortisol circadian rhythm in infancy. Early Hum Dev 2003; 73:39-52

13.   Oster H, Challet E, Ott V, Arvat E, de Kloet ER, Dijk DJ, Lightman S, Vgontzas A, Van Cauter E. The Functional and Clinical Significance of the 24-Hour Rhythm of Circulating Glucocorticoids. Endocr Rev 2017; 38:3-45

14.   Joseph D, Chong NW, Shanks ME, Rosato E, Taub NA, Petersen SA, Symonds ME, Whitehouse WP, Wailoo M. Getting rhythm: how do babies do it? Arch Dis Child Fetal Neonatal Ed 2015; 100:F50-54

15.   Nishihara K, Horiuchi S, Eto H, Uchida S. The development of infants' circadian rest-activity rhythm and mothers' rhythm. Physiol Behav 2002; 77:91-98

16.   Arslanoglu S, Bertino E, Nicocia M, Moro GE. WAPM Working Group on Nutrition: potential chronobiotic role of human milk in sleep regulation. J Perinat Med 2012; 40:1-8

17.   Pacha J. Development of intestinal transport function in mammals. Physiol Rev 2000; 80:1633-1667

18.   Yeh KY, Yeh M, Holt PR. Induction of intestinal differentiation by systemic and not by luminal corticosterone in adrenalectomized rat pups. Endocrinology 1989; 124:1898-1904

19. Dettmer AM, Murphy AM, Guitarra D, Slonecker E, Suomi SJ, Rosenberg KL, Novak MA, Meyer JS, Hinde K. Cortisol in Neonatal Mother's Milk Predicts Later Infant Social and Cognitive Functioning in Rhesus Monkeys. Child Dev 2017;

20. Hinde K, Skibiel AL, Foster AB, Del Rosso L, Mendoza SP, Capitanio JP. Cortisol in mother's milk across lactation reflects maternal life history and predicts infant temperament. Behav Ecol 2015; 26:269-281

21. Sullivan EC, Hinde K, Mendoza SP, Capitanio JP. Cortisol concentrations in the milk of rhesus monkey mothers are associated with confident temperament in sons, but not daughters. Dev Psychobiol 2011; 53:96-104

22. Catalani A, Casolini P, Scaccianoce S, Patacchioli FR, Spinozzi P, Angelucci L. Maternal corticosterone during lactation permanently affects brain corticosteroid receptors, stress response and behaviour in rat progeny. Neuroscience 2000; 100:319-325

23. Glynn LM, Davis EP, Schetter CD, Chicz-Demet A, Hobel CJ, Sandman CA. Postnatal maternal cortisol levels predict temperament in healthy breastfed infants. Early Hum Dev 2007; 83:675-681

24. Grey KR, Davis EP, Sandman CA, Glynn LM. Human milk cortisol is associated with infant temperament. Psychoneuroendocrinology 2013; 38:1178-1185

25. Hart S, Boylan LM, Border B, Carroll SR, McGunegle D, Lampe RM. Breast milk levels of cortisol and Secretory Immunoglobulin A (SIgA) differ with maternal mood and infant neuro-behavioral functioning. Infant Behavior & Development 2004; 27:101-106

26. Nolvi S, Uusitupa HM, Bridgett DJ, Pesonen H, Aatsinki AK, Kataja EL, Korja R, Karlsson H, Karlsson L. Human milk cortisol concentration predicts experimentally induced infant fear reactivity: moderation by infant sex. Dev Sci 2018; 21:e12625

27. van der Voorn B, de Waard M, van Goudoever JB, Rotteveel J, Heijboer AC, Finken MJ. Breast-Milk Cortisol and Cortisone Concentrations Follow the Diurnal Rhythm of Maternal Hypothalamus-Pituitary-Adrenal Axis Activity. J Nutr 2016; 146:2174-2179

28. Stalder T, Kirschbaum C, Kudielka BM, Adam EK, Pruessner JC, Wust S, Dockray S, Smyth N, Evans P, Hellhammer DH, Miller R, Wetherell MA, Lupien SJ, Clow A. Assessment of the cortisol awakening response: Expert consensus guidelines. Psychoneuroendocrinology 2016; 63:414-432

29. Herrmann C. International experiences with the Hospital Anxiety and Depression Scale--a review of validation data and clinical results. J Psychosom Res 1997; 42:17-41

30. Parade SH, Leerkes EM. The reliability and validity of the Infant Behavior Questionnaire-Revised. Infant Behav Dev 2008; 31:637-646

31. Worobey J, Blajda VM. Temperament Ratings at 2 Weeks, 2 Months, and 1 Year - Differential Stability of Activity and Emotionality. Dev Psychol 1989; 25:257-263

32. Rothbart MK. Measurement of Temperament in Infancy. Child Development 1981; 52:569-578

33. van der Voorn B, Martens F, Peppelman NS, Rotteveel J, Blankenstein MA, Finken MJ, Heijboer AC. Determination of cortisol and cortisone in human mother's milk. Clin Chim Acta 2015; 444:154-155

34. Pruessner JC, Kirschbaum C, Meinlschmid G, Hellhammer DH. Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. Psychoneuroendocrinology 2003; 28:916-931

35. Murphy BE. Ontogeny of cortisol-cortisone interconversion in human tissues: a role for cortisone in human fetal development. J Steroid Biochem 1981; 14:811-817

36. Catalani A, Casolini P, Cigliana G, Scaccianoce S, Consoli C, Cinque C, Zuena AR, Angelucci L. Maternal corticosterone influences behavior, stress response and corticosteroid receptors in the female rat. Pharmacol Biochem Behav 2002; 73:105-114

37. Slotkin TA, Seidler FJ, Wood CR, Lau C. Development of glucocorticoid receptor regulation in the rat forebrain: implications for adverse effects of glucocorticoids in preterm infants. Brain Res Bull 2008; 76:531-535

38. Smith RE, Maguire JA, Stein-Oakley AN, Sasano H, Takahashi K, Fukushima K, Krozowski ZS. Localization of 11 beta-hydroxysteroid dehydrogenase type II in human epithelial tissues. J Clin Endocrinol Metab 1996; 81:3244-3248

39. Blair J, Adaway J, Keevil B, Ross R. Salivary cortisol and cortisone in the clinical setting. Curr Opin Endocrinol Diabetes Obes 2017; 24:161-168

40. Savas M, Wester VL, de Rijke YB, Rubinstein G, Zopp S, Dorst K, van den Berg SAA, Beuschlein F, Feelders RA, Reincke M, van Rossum EFC. Hair glucocorticoids as biomarker for endogenous Cushing's syndrome: validation in two independent cohorts. Neuroendocrinology 2019;

41. Watterberg KL. Adrenocortical function and dysfunction in the fetus and neonate. Semin Neonatol 2004; 9:13-21

42. Ackermans MT, Endert E. LC-MS/MS in endocrinology: what is the profit of the last 5 years? Bioanalysis 2014; 6:43-57

43. Hane AA, Fox NA, Polak-Toste C, Ghera MM, Guner BM. Contextual basis of maternal perceptions of infant temperament. Dev Psychol 2006; 42:1077-1088

44. Reck C, Muller M, Tietz A, Mohler E. Infant distress to novelty is associated with maternal anxiety disorder and especially with maternal avoidance behavior. J Anxiety Disord 2013; 27:404-412

45. Tikotzky L, Chambers AS, Gaylor E, Manber R. Maternal sleep and depressive symptoms: links with infant Negative Affectivity. Infant Behav Dev 2010; 33:605-612

46. Gartstein MA, Rothbart MK. Studying infant temperament via the Revised Infant Behavior Questionnaire. Infant Behavior & Development 2003; 26:64-86

47. Marysko M, Finke P, Wiebel A, Resch F, Moehler E. Can Mothers Predict Childhood Behavioural Inhibition in Early Infancy? Child Adol Ment H-Uk 2010; 15:91-96

SUPPLEMENTARY FILE 1

Sleep questionnaire

Amount of sleep

These questions relate to how much sleep you and your child have had in the past week. We will also ask some short questions about the quality of your own sleep. The questions have to be answered differently than the previous questions of this questionnaire. This time the questions are open. Your answers can therefore be a rough estimate.

In the past week...

1. On average, how many hours did your child sleep at night?   ….. hour(s)

2. On average, how often did your child wake at night?   ….. time(s)

3. How often did your child sleep more than 6 hours at night?   ….. time(s)

4. On average, how many naps did your child take during daytime?   ….. nap(s)

5. On average, how many hours of sleep did your child get during daytime   ….. hour(s)

Part 2

Glucocorticoid regulation and sex
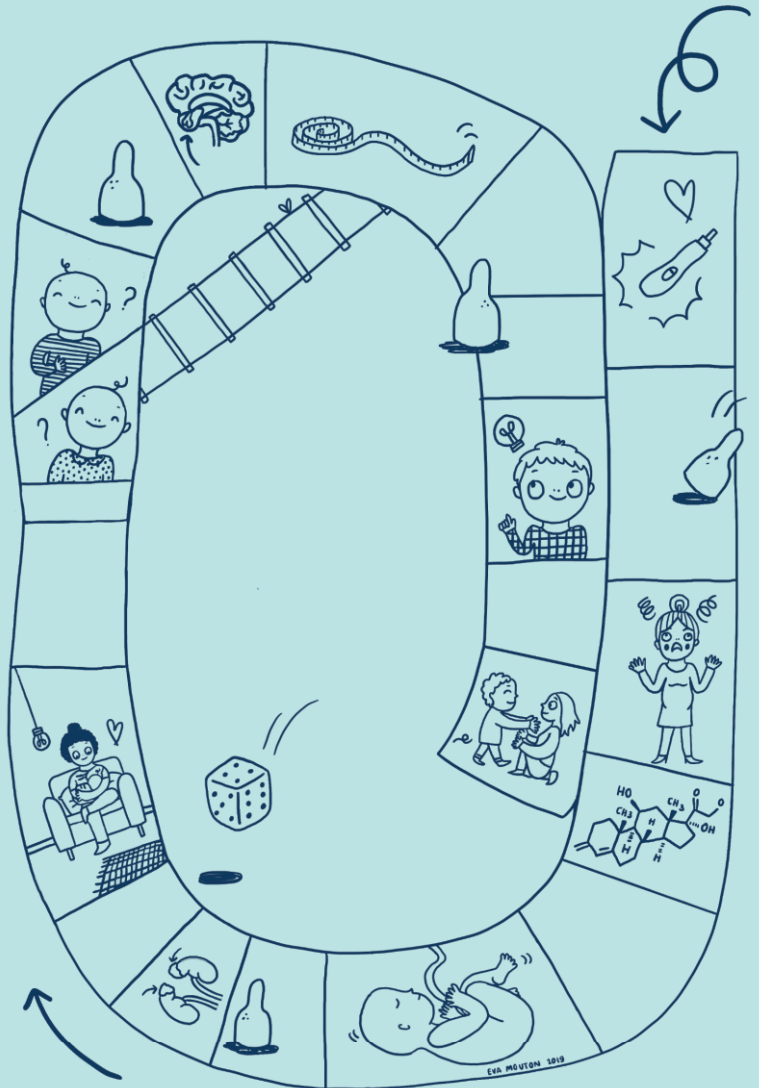
9

Gender-specific differences in hypothalamus–pituitary–adrenal axis activity during childhood: a systematic review and meta-analysis

Bibian van der Voorn,
Jonneke J. Hollanders,
Johannes C.F. Ket,
Joost Rotteveel,
Martijn J.J. Finken

ABSTRACT

Background

Gender-specific differences in HPA axis activity have been postulated to emerge during puberty. We conducted a systematic review and meta-analysis to test the hypothesis that gender-specific differences in HPA axis activity are already present in childhood.

Methods

From inception to January 2016, PubMed and Embase.com were searched for studies that assessed non-stimulated cortisol in serum or saliva, or cortisol in 24h-urine in healthy males and females aged ≤18yr. Studies were reported conform the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement. Standardized mean differences (95%CIs) were calculated and analyzed using fixed-effect meta-analysis stratified for age: <8yr (prepubertal) and 8-18yr (peri-/postpubertal). For comparison, we ran the same analyses using random-effects models.

Results

Two independent assessors selected 413 out of 6,158 records (7%) for full-text screening, of which 79 articles were included. Of these, 58 (with data on 16,551 subjects) were included in the meta-analysis. Gender differences in cortisol metabolism differed per age group. Boys aged <8yr had 0.18 (0.06 to 0.30) nmol/L higher serum and 0.21 (0.05 to 0.37) nmol/L higher salivary cortisol levels, while between 8-18yr, boys had 0.34 (0.28 to 0.40) nmol/L lower serum and 0.42 (0.38; 0.47) nmol/L lower salivary cortisol levels. In 24h-urine, cortisol was consistently higher in boys, being 0.34 (0.05 to 0.64) and 0.32 (0.17 to 0.47) μg/24h higher in the <8yr and 8-18yr groups, respectively. However, gender-differences in serum cortisol <8 yr and between 8-18 yr were absent when using random-effects models.

Conclusions

Gender differences in cortisol metabolism are already present in childhood, with higher salivary cortisol in boys aged <8yr compared to girls. This pattern was reversed after age 8 yr. In contrast, the gender-specific difference in cortisol production as assessed through 24h-urine did not change with age. Although differences were small, and analyses of gender differences in serum cortisol were inconclusive, they might contribute to gender-specific origins of health and disease.

BACKGROUND

The hypothalamus-pituitary-adrenal (HPA) and hypothalamus-pituitary-gonadal (HPG) axes are closely connected. Animal studies demonstrated that CRH inhibits the HPG axis at all levels, while testosterone inhibits the HPA axis at the hypothalamic level. Additionally, estrogens stimulate the HPA axis at both the hypothalamic and adrenal levels. Moreover, CRH levels were dependent on the phase of the menstrual cycle, with the highest concentrations occurring during the follicular phase.[1,2]

Human studies suggested that estrogens decrease the hepatic A-ring reduction of cortisol, albeit not in the short term[3] and increase the production of CBG, thereby affecting the bioavailability of cortisol.[1,4,5] The latter being enhanced by the use of oral contraceptives. Furthermore, HPA axis responses to acute psychological stress were different depending on the phase of the menstrual cycle.[2,4]

Due to an increase in sex steroid concentrations, gender differences in HPA axis activity have been postulated to emerge during puberty.[6,7] However, more recent evidence suggests that gender differences in HPA axis activity are already present early in life.[1,8,9] Putative mediators of these prepubertal gender differences are the postnatal reproductive hormone surge, also known as mini-puberty,[10] and sex-specific effects of styles in parental care, such as psychosocial stress reactivity to maternal over-controlling behavior.[11] However, physiological gender differences in cortisol concentrations during childhood have not been studied yet.

Therefore, the question was raised whether gender differences in unstimulated HPA axis activity emerge during puberty or whether they are already present earlier in life. Accordingly, we conducted a systematic review and meta-analysis with the hypothesis that gender-specific differences in unstimulated HPA axis activity are present in early life and are subsequently influenced by puberty.

METHODS

Search strategy

From inception up to 14 January 2016, PubMed and Embase.com were searched (by BvdV and JCFK) for studies that reported non-stimulated cortisol in serum or saliva, or cortisol in 24h-urine for healthy boys and girls aged ≤18 yr separately. Appendix 1 presents the full search strategy, which was based on the following index terms or free-text words: 'cortisol' or 'glucocorticoid', and 'sex difference' or 'sexual characteristics', and 'child' or 'adolescent'. Studies in children with (psycho)pathology, on synthetic glucocorticoids, or with risk for abnormal HPA axis activity (e.g., a history of maltreatment) were excluded. An English language restriction was applied for abstracts of published

articles. No restrictions for year of publication or study design, apart from reviews and case reports, were applied. The review protocol was based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement.

Data collection

Two independent assessors (BvdV and JJH) screened 6,158 titles and abstracts without consideration of outcomes. Studies were not assessed blindly. Disagreement between assessors was discussed until consensus was reached. When gender differences were analyzed without reporting on cortisol levels for boys and girls separately or when data were only presented in graphs, authors were requested for additional quantitative data. Data were stratified into two age groups: <8 yr (prepubertal) and between 8–18 yr (peri-/postpubertal). Ideally, stratification would have been based on pubertal staging according to Tanner. Unfortunately, only a minority of the included studies reported on the subjects' Tanner stages. Because pubertal onset before age 8 years is considered to be pathologic,[12] we chose 8 yr as cut off for stratification. When articles reported on serial cortisol measurements, we included only data on the youngest assessment age. When cortisol levels were reported prepubertally as well as peri-/postpubertally within the same individual, we included one sampling moment for each stratified group. When articles reported on the same study population, we included the article with the lowest bias risk. When articles reported on dynamic tests of HPA axis activity, we only included baseline cortisol. We only included the control subjects of case-control studies. If known, we excluded female subjects on oral contraceptives. When gender differences were described but not quantified, the articles were included in the descriptive analysis rather than the meta-analysis.

Meta-analysis

When necessary, we converted serum and salivary cortisol levels into nmol/L, and 24h-urine cortisol levels into μg/24h. When means ± SDs were not reported, the SD was calculated based on the following assumptions: the 95% CI is 3.92 SDs wide (2 × 1.96); the inter-quartile range is 1.35 SDs wide; the range is 4 SDs wide; the SD is the SE multiplied by the square root of the sample size.[13] To assess parametricity, we assumed that a normal distribution extends no more than 2 SDs from the mean,[14] i.e., when normally distributed, the mean minus 2 SDs should be >0 nmol/L. Data analyses were performed using Review Manager (RevMan) version 5.3.5, 2014. For each study, the standardized mean gender difference (95% CI) in cortisol concentration was calculated by combining the SD with the sample size. Subsequently, fixed-effect meta-analyses were performed first, which assumes that the effect estimate of the group differences was fixed across studies. Second, the results of these analyses were compared with random-effects meta-analysis, which weigh studies of variable sample sizes more equally. We reported any

source of bias from each included article conform the PRISMA statement and assessed selection, performance, detection and other biases. (Figure 1, Appendix 2) Bias was assessed as low, unclear or high. A sensitivity analysis was done by excluding studies that had ≥ 1 high bias risks. Heterogeneity of the data was assessed by the I^2 statistic, with significance defined as $I^2 >50\%$. Publication bias was assessed through funnel plots.
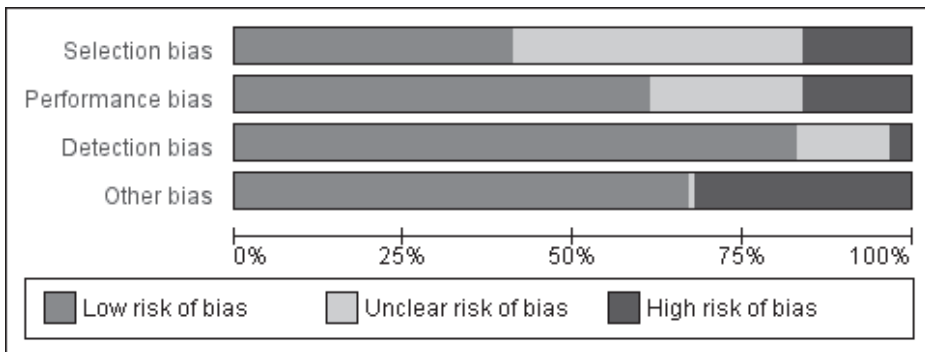
Figure 1: Risk of bias graph presenting a summary of the judgements of the assessors concerning risk of bias across all studies included in the meta-analysis. Bias risk is presented as percentage of total studies (n = 58).

RESULTS

Figure 2 shows the flowchart of the descriptive analysis and meta-analysis. Of the 6,158 titles and abstracts, 414 (7%) were eligible for full-text screening, from which 79 articles (19%) were included. Thirty-one authors of articles with insufficient quantitative data were contacted, of whom 12 responded: six provided the necessary quantitative data, five did not have access to the raw data anymore and one was not willing to participate. Two articles reported the cortisol production rate assessed through 24h-serum sampling, which hampered inclusion in the meta-analysis. The authors of 27 articles that only provided gender-specific data in figures were contacted, but could not be reached. Subsequently, these articles were excluded. Finally, 21 articles were included only in the descriptive analysis, and 58 articles (with data on 16,551 subjects) had sufficient data for inclusion in the meta-analysis.

Description of included studies

Studies were conducted in Europe (n = 36), North-America (n = 37), Asia (n = 3), South-America (n = 2) or Africa (n = 1), and were published between 1973 and 2016. Sample sizes ranged from 11 to 2,824 subjects, with seven studies having a sample size >500 subjects. Study designs were as follows: randomized placebo-controlled (n = 2), pro-
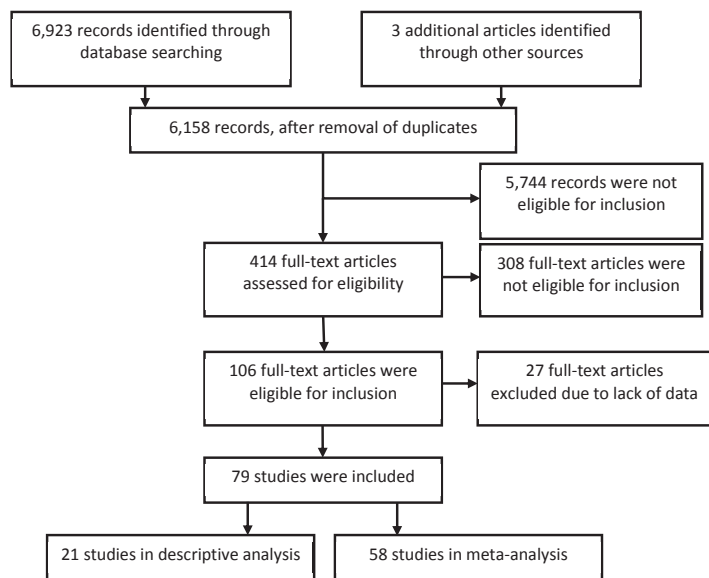
Figure 2: This flowchart presents the different phases of the systematic review and meta-analysis, conform the PRISMA-statement. (www.prisma-statement.org)

spective observational (n = 29), non-randomized intervention, i.e., stress tests (n = 15), cross-sectional (n = 16), longitudinal (n = 11) and case-control (n = 6). All studies that assessed serum or salivary cortisol used immunoassays, except for one that used high-performance liquid chromatography (HPLC). Studies that assessed 24h-urine cortisol used immunoassays (n = 4), gas chromatography–mass spectrometry (n = 3), HPLC (n = 1), and liquid chromatography-UV detection (n = 1). Twenty-two studies (28%) did not collect morning samples, of which 11 did not report the time of collection and 11 described specifically that samples were collected in the afternoon. Online Supplementary File 1 presents the data extracted from the articles included in the meta-analysis. Three out of 21 studies (14%) included in the descriptive analysis had no high bias risk (Table 1), while 16 out of 58 studies (28%) included in the meta-analysis had no high bias risk (Figure 1).

Gender-specific differences

Descriptive analysis

Table 1 summarizes the data on the 21 studies included in the descriptive analysis. The majority (90%) of these studies reported no significant gender differences in cortisol levels. Before age 8 yr, one study [15] found significantly lower salivary cortisol levels for boys at awakening. Between ages 8-18 yr, one study [8] found significantly lower morning salivary cortisol levels in boys.

Table 1: Summary of studies included in the descriptive analysis.

Group	First author (year)	N (%girls)	Age (yr)	Sample protocol	Assay	Result	Bias*
Saliva <8 yr	Klug (2000)[35]	119 (46%)	0	3 point day curve	Immunoassay	No gender differences	2
	Eiden (2015)[36]	257 (?)	0.75	Laboratory Temperament Assessment	Immunoassay	No gender differences	3
	Plusquellec (2011)[37]	466 (?)	1.6 ± 0.1	Morning sample	Immunoassay	No gender differences	2
	Spinrad (2009)[38]	84 (49%)	4.5	Preschool Laboratory Assessment	Immunoassay	No gender differences	2
	Hatzinger (2007)[15]	102 (42%)	4.9 ± 0.4	CAR	Immunoassay	Cortisol levels were lower in boys at awakening (p<0.1)	1
Saliva 8–18 yr	Safarzadeh (2005)[39]	100 (58%)	6 – 14	Morning sample	Immunoassay	No gender differences	1
	Isaksson (2015)[40]	68 (50%)	9	Morning sample	Immunoassay	No gender differences	2
	Kjölhede (2014)[41]	231 (50%)	9.5 ± 1.5	Morning sample	Immunoassay	No gender differences	1
	Vaillancourt (2008)[8]	154 (52%)	12.3 ± 0.8	Six samples standardized across time and day	Immunoassay	On Saturday morning boys had significantly lower morning levels. On Monday and Thursday no gender differences were found.	1
	Gunnar (2009)[42]	82 (49%)	9 – 15	TSST	Immunoassay	No gender differences	1
Serum <8 yr	Fadalti (1999)[43]	72 (49%)	0 – 2	Morning sample	Immunoassay	No gender differences	0
	Ballerini (2010)[44]	319 (45%)	0 – 5	Surplus serum	Immunoassay	No gender differences	2
	Parker (1978)[45]	106 (43%)	2 – 12	Morning sample	Immunoassay	No gender differences	2
Serum 8–18 yr	Kulasingam (2010)[46]	419 (?)	0 – 15	Surplus serum	Immunoassay	No gender differences	3
	Soldin (2005)[47]	376 (?)	0 – 18	Surplus serum	Immunoassay	No gender differences	1
	Karbasy (2015)[48]	711 (?)	0 – 19	?	Immunoassay	No gender differences	1
	Fadalti (1999)[43]	82 (49%)	6 – 18	Morning sample	Immunoassay	No gender differences	0
	Barra (2015)[49]	120 (45%)	12.4 ± 3	Morning sample	Immunoassay	No gender differences	1
	Chalew (1997)[50]	15 (73%)	12.7 ± 2.2	24h-blood withdrawal	Immunoassay	No gender differences	1
	Linder (1990)[51]	82 (58%)	8 – 17	24h-blood withdrawal	HPL	No gender differences .	0
Urine <8 yr	-						
Urine 8–18 yr	Dorn (1996)[52]	20 (55%)	15.2 ± 1.1	24h-urine sample	Immunoassay	No gender differences	1

* number of high risks of bias out of 4 bias categories (selection, performance, detection and other biases)

A

Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
0 yr (Davis 1995)	30.1	36.4	18	29.5	32.8	18	6.1%	0.02 [-0.64, 0.67]
3 yr (Gunnar 2010)	4.03	4.58	37	3.89	3.01	54	14.9%	0.04 [-0.38, 0.46]
3.3 yrs (De Bruijn 2009)	2.82	1.86	19	2.45	1.82	27	7.5%	0.20 [-0.39, 0.79]
4 yr (Pérez-Edgar 2008)	11.9	4.7	51	10.8	5	60	18.6%	0.22 [-0.15, 0.60]
4.1yr (Mills 2008)	3.03	2.76	115	2.76	2.76	85	33.1%	0.10 [-0.18, 0.38]
4.3 yr (Tout 1998)	5.79	2.48	38	4.41	1.38	37	12.0%	0.68 [0.21, 1.14]
4yr (Gunnar 2010)	2.92	1.35	11	4.47	4.88	13	3.9%	-0.40 [-1.22, 0.41]
7.3yr (Törnhage 2002)	10.5	5.2	19	5.2	1.88	11	4.0%	1.19 [0.38, 2.00]
Total (95% CI)			**308**			**305**	**100.0%**	**0.21 [0.05, 0.37]**

Heterogeneity: Chi² = 13.34, df = 7 (P = 0.06); I² = 48%
Test for overall effect: Z = 2.53 (P = 0.01)

B

Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
8yr (Jones 2006)	9.58	1.51	68	10.59	1.49	72	1.6%	-0.67 [-1.01, -0.33]
8 yr (Azurmendi 2016)	8	4.14	49	6.35	3.59	41	1.0%	0.42 [-0.00, 0.84]
8yr (Martikainen 2013)	6.8	1.7	126	8	1.7	132	2.9%	-0.70 [-0.96, -0.45]
8.4 yr (Michels 2012)	12.66	6.12	183	13.84	8.96	202	4.6%	-0.15 [-0.35, 0.05]
9 yr (Alghadir 2009) active	0.13	0.026	12	0.16	0.017	12	0.3%	-1.25 [-2.00, -0.50]
9yr (Alghadir 2009) sedentary	0.18	0.014	12	0.2	0.028	16	0.3%	-0.84 [-1.62, -0.05]
9.2 yr (Cicchetti 2001)	8.64	5.16	124	9.52	5.21	80	2.3%	-0.17 [-0.45, 0.11]
9.4 yr (Turan 2015)	5.2	3.2	79	5.6	10.8	74	1.8%	-0.05 [-0.37, 0.27]
9.9 yr (Osika 2007)	5.34	2.1	34	5.5	3.33	50	1.0%	-0.05 [-0.49, 0.38]
10yr (Törnhage 2002)	11	4.08	28	8.9	13.03	31	0.7%	0.21 [-0.30, 0.72]
10.4±0.4yr (Cieslak 2003)	3	0.5	29	3	0.3	32	0.7%	0.00 [-0.50, 0.50]
11.1±0.55yr (Dietrich 2013)	11.2	4.7	818	11.9	4.73	786	19.0%	-0.15 [-0.25, -0.05]
11.9±0.6yr (Chen 2014)	4.69	5.52	213	5.52	5.24	212	5.0%	-0.15 [-0.34, 0.04]
11.9 yr (Portnoy 2015)	4.06	3.72	224	5.79	22.18	214	5.2%	-0.11 [-0.30, 0.08]
12 yr (Tzortzi 2009)	9.05	3.63	11	10.24	4.86	10	0.2%	-0.27 [-1.13, 0.59]
12.7 ± 2.9yr (Allen 2009)	0.17	0.18	119	0.13	0.09	116	2.8%	0.28 [0.02, 0.54]
12.7 (Yu 2009)	6.22	0.92	10	4.83	2.04	10	0.2%	0.84 [-0.08, 1.76]
13.4 yr (Mrug 2016)	3.43	2.83	40	2.31	2.5	39	0.9%	0.42 [-0.03, 0.86]
14.0±0.5yr (Belva 2013)	24.6	10.6	115	29.2	13.2	108	2.6%	-0.38 [-0.65, -0.12]
14.9yr (Alghadir 2009) active	0.22	0.015	20	0.24	0.034	30	0.5%	-0.70 [-1.29, -0.12]
14.9 yr (Alghadir 2009) sedentary	0.16	0.018	25	0.17	0.014	10	0.3%	-0.57 [-1.32, 0.17]
14 -16yr (Fransson 2014)	6.6	2	43	10.9	2	89	0.9%	-2.14 [-2.59, -1.69]
15.3 ± 1.1yr (Covelli 2012)	27	11	49	25	11.2	57	1.2%	0.18 [-0.20, 0.56]
15.5 yr (West 2010)	10.64	1.02	1418	11.67	1.02	1406	29.8%	-1.01 [-1.09, -0.93]
16 ±1yr (Georgopoulos 2011)	9.1	8.3	35	11.6	8.3	29	0.7%	-0.30 [-0.79, 0.20]
16.1 ±1.0yr (Daughters 2013)	3.22	2.13	59	2.3	1.38	73	1.5%	0.52 [0.17, 0.87]
16.6 ± 0,5yr (Reynolds 2013)	25.93	14.1	573	29.5	19	362	10.5%	-0.22 [-0.35, -0.09]
17.9 yr (Minckley 2012)	10.47	6.29	14	8.42	2.84	14	0.3%	0.41 [-0.34, 1.16]
18 yr (Colomina 1997)	13.7	7.7	37	13.4	7.6	36	0.9%	0.04 [-0.42, 0.50]
Total (95% CI)			**4580**			**4343**	**100.0%**	**-0.42 [-0.47, -0.38]**

Heterogeneity: Chi² = 481.05, df = 28 (P < 0.00001); I² = 94%
Test for overall effect: Z = 19.37 (P < 0.00001)

C

Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
0 yr (Elmlinger 2002)	253	95.7	42	271	97.4	44	7.7%	-0.18 [-0.61, 0.24]
0 yr (Forest 1978)	312	168	15	448	187	3	0.9%	-0.76 [-2.03, 0.51]
0 yr (Tennes 1973)	435.9	313.6	23	389.8	227.6	17	3.5%	0.16 [-0.47, 0.79]
0 yrs (Garagorri 2008)	249.41	99.32	80	280.31	124.16	58	12.0%	-0.28 [-0.62, 0.06]
0-2 (Bailey 2013)	159.8	135.3	261	190.2	186.2	208	41.6%	-0.19 [-0.37, -0.01]
0.4 yr (Lashansky 1991)	342	146	13	353	196	8	1.8%	-0.06 [-0.94, 0.82]
2-8 yr (Bailey 2013)	134.4	75.6	110	129.9	71.5	86	17.4%	0.06 [-0.22, 0.34]
3 yr (Lashansky 1991)	337	160	14	281	113	8	1.8%	0.37 [-0.51, 1.25]
4-10 yr (Tsvetkova 1977)	518.69	193.1	9	510.42	157.3	8	1.5%	0.04 [-0.91, 1.00]
6.8 yr (Soriano-Rodriguez 2010)	441.4	13.8	106	408.3	13.8	117	11.6%	2.39 [2.05, 2.74]
Total (95% CI)			**673**			**557**	**100.0%**	**0.17 [0.05, 0.29]**

Heterogeneity: Chi² = 186.39, df = 9 (P < 0.00001); I² = 95%
Test for overall effect: Z = 2.79 (P = 0.005)

D

Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI	Std. Mean Difference IV, Fixed, 95% CI
8 yr (Apter 1979)	240	80.3	7	226.2	109.5	7	0.4%	0.13 [-0.91, 1.18]	
8yr (Lashansky 1991)	254	86	8	229	94	7	0.4%	0.26 [-0.76, 1.28]	
8yr (Ong 2004)	210	58	423	215	69	347	22.4%	-0.08 [-0.22, 0.06]	
9 yr (Ilias 2009)	112.97	49.14	5	153.26	131.05	6	0.3%	-0.36 [-1.56, 0.84]	
6-15 yr (Ross 1986)	74.5	35.9	23	80	44.1	24	1.4%	-0.13 [-0.71, 0.44]	
10yr (Stroud 2011)	164.4	5.9	39	146.5	5.99	30	0.9%	2.98 [2.28, 3.68]	
11-14 (Tsvetkova 1977)	424.89	171.1	7	408.3	160	7	0.4%	0.09 [-0.95, 1.14]	
11y (Elmlinger 2002)	341	113.5	22	364	117.3	23	1.3%	-0.20 [-0.78, 0.39]	
12.5 yr (Susman 1991)	364.7	149.5	56	344.3	148.7	52	3.2%	0.14 [-0.24, 0.51]	
13 yr (Hackney 2003)	397.3	126.9	62	466.3	162.8	63	3.6%	-0.47 [-0.82, -0.11]	
13yr (Lashansky 1991)	223	74	10	229	94	11	0.6%	-0.07 [-0.92, 0.79]	
14.6 yr (Ghaziuddin 2003)	1,757.5	1,191.9	12	1,914.7	1,658.2	9	0.6%	-0.11 [-0.97, 0.76]	
14.7 yr (Huybrechts 2014)	299.5	1.6	297	305.6	1.7	426	7.8%	-3.67 [-3.91, -3.43]	
15 yr (Syme 2008)	546.3	140.7	77	714.6	306.24	82	4.4%	-0.70 [-1.02, -0.38]	
15.5yr (Lashansky 1991)	262	80	13	279	77	8	0.6%	-0.21 [-1.09, 0.68]	
16.6 ± 0,5yr (Reynolds 2013)	551	154	653	590.98	202.79	401	29.1%	-0.23 [-0.35, -0.10]	
17 yr (Stupnicki 1995)	347	130	14	397	125	15	0.8%	-0.38 [-1.12, 0.35]	
8-18yr (Bailey 2013)	218.3	98.9	370	215.7	110.7	369	21.7%	0.02 [-0.12, 0.17]	
Total (95% CI)			**2098**			**1887**	**100.0%**	**-0.39 [-0.46, -0.32]**	

Heterogeneity: Chi² = 876.34, df = 17 (P < 0.00001); I² = 98%
Test for overall effect: Z = 11.36 (P < 0.00001)

-2 -1 0 1 2
Female Male

E

Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI	Std. Mean Difference IV, Fixed, 95% CI
3yr (Lundberg 1981)	3.35	1.12	4	2.46	0.44	9	5.2%	1.19 [-0.11, 2.49]	
4.5yr (Lundberg 1983)	8.5	1.8	8	11.9	5.5	3	4.3%	-1.02 [-2.45, 0.41]	
3-4 yr (Wudy 2007)	20.5	6.63	25	21.2	6.49	25	28.6%	-0.11 [-0.66, 0.45]	
5-6 yr (Wudy 2007)	30.5	10.2	25	25.4	7.52	25	27.4%	0.56 [-0.01, 1.13]	
5-10 yr (Nakamura 1984)	14.8	3.6	7	9.98	3.2	8	6.6%	1.34 [0.18, 2.50]	
7-8 yr (Wudy 2007)	36.9	12.9	25	31.9	10.9	25	28.0%	0.41 [-0.15, 0.97]	
Total (95% CI)			**94**			**95**	**100.0%**	**0.34 [0.05, 0.64]**	

Heterogeneity: Chi² = 11.11, df = 5 (P = 0.05); I² = 55%
Test for overall effect: Z = 2.28 (P = 0.02)

-2 -1 0 1 2
Female Male

F

Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI	Std. Mean Difference IV, Fixed, 95% CI
8yr (Honour 2007)	2,826.86	774.81	244	2,557.71	655.56	188	57.7%	0.37 [0.18, 0.56]	
9-10 yr (Wudy 2007)	35.5	9.26	25	34.9	12.5	25	6.9%	0.05 [-0.50, 0.61]	
4-15yr (Canalis 1982)	15.2	6.3	14	13	5.6	15	3.9%	0.36 [-0.38, 1.09]	
11-12 yr (Wudy 2007)	54.1	13.8	25	43.6	12.2	25	6.4%	0.79 [0.22, 1.37]	
13 ±3yr (Vaindirlis 2000)	94	32	12	69	33	10	2.8%	0.74 [-0.13, 1.61]	
13-14 yr (Wudy 2007)	53.2	21.4	25	49.5	15.1	25	6.9%	0.20 [-0.36, 0.75]	
14-19yr (Nakamura 1984)	40.9	23	6	33.3	3.1	7	1.7%	0.45 [-0.66, 1.56]	
15-16 yr (Wudy 2007)	68.5	21	25	65.1	22	25	6.9%	0.16 [-0.40, 0.71]	
17-18 yr (Wudy 2007)	75	25.3	25	81.1	32.3	25	6.9%	-0.21 [-0.76, 0.35]	
Total (95% CI)			**401**			**345**	**100.0%**	**0.32 [0.17, 0.47]**	

Heterogeneity: Chi² = 8.67, df = 8 (P = 0.37); I² = 8%
Test for overall effect: Z = 4.31 (P < 0.0001)

-2 -1 0 1 2
Female Male
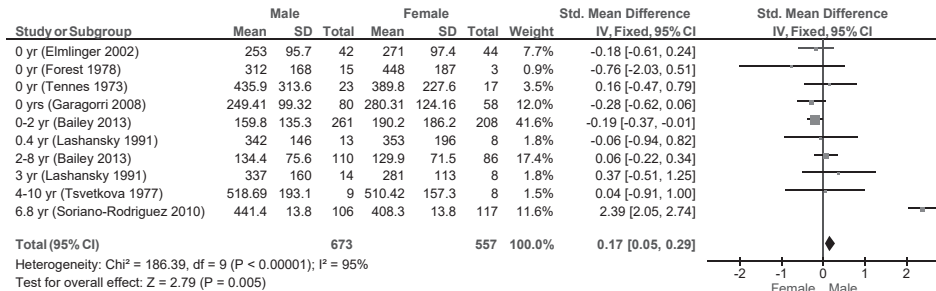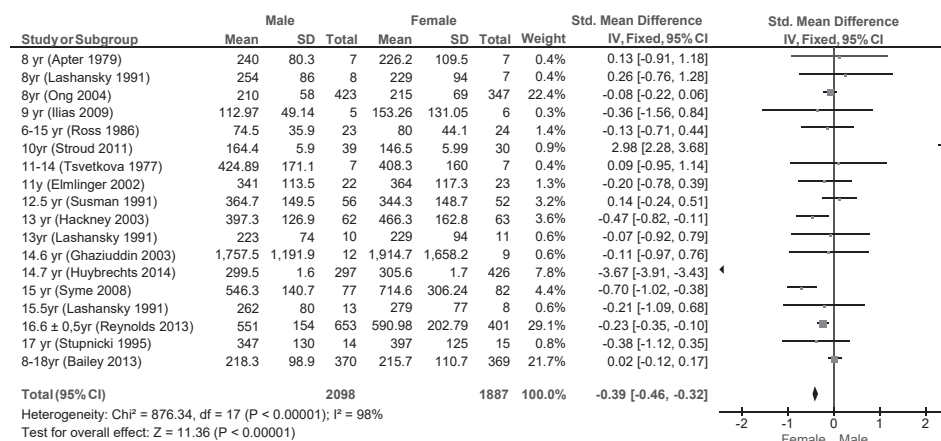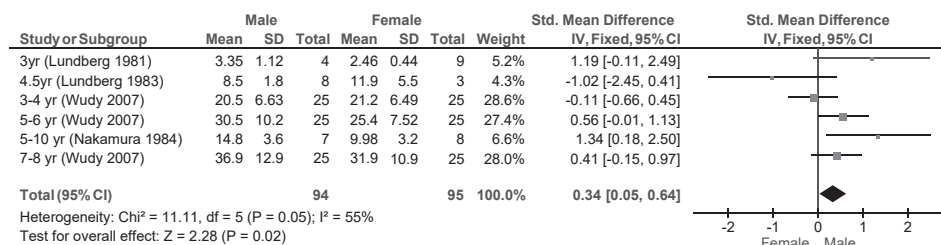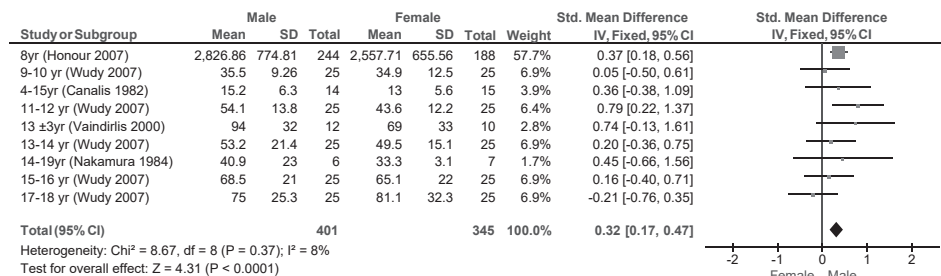
Figure 3: Forest plots of gender differences per subgroup (Fixed effect analyses)
A Salivary cortisol (nmol/L) <8 yr of age **B** Salivary cortisol (nmol/L) 8–18 yr of age **C** Serum cortisol (nmol/L) <8 yr of age **D** Serum cortisol (nmol/L) 8–18 yr of age **E** 24h-urine cortisol (μg/24h) <8 yr of age **F** 24h-urine cortisol (μg/24h) 8–18 yr of age

Meta-analysis

Nine articles (16%) did not report mean and SD values, which were therefore calculated. Figure 3 shows the results of the fixed-effect meta-analysis. Compared to girls, boys <8 yr had 0.21 (0.05 to 0.37) nmol/L ($P = 0.01$, $I^2 = 48\%$) higher salivary and 0.18 (0.06 to 0.30) nmol/L ($P < 0.01$, $I^2 = 94\%$) higher serum cortisol levels. Between ages 8-18 yr, boys had 0.42 (0.38 to 0.47) nmol/L ($P < 0.01$, $I^2 = 94\%$) lower salivary and 0.34 (0.28 to 0.40) nmol/L ($P < 0.01$, $I^2 = 97\%$) lower serum cortisol levels. In contrast, free cortisol in 24h-urine was 0.34 (0.05 to 0.64) µg/24h ($P = 0.02$, $I^2 = 55\%$) higher in boys aged <8 yr and 0.32 (0.17 to 0.47) µg/24h ($P < 0.01$, $I^2 = 8\%$) higher in boys between ages 8-18 yr. The sensitivity analyses did not significantly change the results, although it decreased the heterogeneity: boys <8 yr had 0.40 (0.11 to 0.69) nmol/L ($P < 0.01$, $I^2 = 55\%$) higher salivary, 0.45 (0.30 to 0.61) nmol/L ($P < 0.01$, $I^2 = 94\%$) higher serum and 0.28 (-0.04 to 0.61) µg/24h ($P = 0.08$, $I^2 = 33\%$) higher 24h-urine cortisol; boys 8–18 yr had 0.20 (0.13 to 0.26) nmol/L ($P < 0.01$, $I^2 = 47\%$) lower salivary, 0.10 (0.02 to 0.18) nmol/L ($P = 0.01$, $I^2 = 33\%$) lower serum and 0.24 (0.02 to 0.47) µg/24h ($P = 0.04$, $I^2 = 24\%$) higher 24h-urine cortisol.

Appendix 3 shows the results of the random-effects meta-analyses. When analyzed by the random-effects method, the effect estimates of serum cortisol <8 yr and between 8-18 yr became non-significant ($P = 0.46$ and $P = 0.62$, respectively). This also applied to salivary cortisol <8 yr ($P = 0.06$) and urinary cortisol <8yr ($P = 0.12$), although trends in the same direction were observed.

Funnel plots showed no evidence of publication bias. (Appendix 4)

DISCUSSION

The results from this meta-analysis suggest that gender-specific differences in HPA axis activity are already present early in life. They also support previous observations which show that cortisol metabolism diverges between genders at pubertal age. Before age 8 yr, cortisol in both serum and saliva was higher in boys compared to girls, at least in fixed-effect meta-analysis. These patterns were reversed after age 8 yr. In contrast, gender differences in 24h-urine cortisol remained consistent with age, with higher cortisol levels in urine for boys before and after age 8 yr.

Total serum cortisol and free salivary cortisol reflect the balance between cortisol production and degradation, i.e., the bioavailability. Our meta-analysis suggests that puberty induces gender-specific changes in the bioavailability of cortisol, as reflected by similar changes in both total serum and free salivary cortisol levels, at least in fixed-effect models. Even though associations were absent for total serum cortisol in random-effects models, the change in free salivary cortisol could not be explained by an estrogen-induced increase in the production of CBG.[4] Moreover, the gender differ-

ence in cortisol in 24hr-urine (i.e., non-metabolized, free cortisol, representing cortisol production rate) remained consistent with age. Consequently, sex-hormone dependent effects on the hepatic metabolism of cortisol are more likely to explain our observations. Cortisol is metabolized reversibly by 11βHSD type 2, and irreversibly by α- and β-ring reductases, and CYP3A. Animal studies showed a lower bioavailability of glucocorticoids in females due to decreased 11βHSD type 1 [16-18] and relatively increased 11βHSD type 2 activity,[18] as compared to males. In addition, previous observations in humans suggest that estrogens could alter hepatic cortisol metabolism through increased CYP3A activity,[19,20] and decreased A-ring reduction.[3,21] In contrast, sex-specificity in the activities of 11βHSD isozymes is debated in humans.[3,21,22] Since analyses of gender-specific differences in total serum cortisol were inconclusive in random-effects models (Appendix 3) and only one of the included studied had assessed CBG levels next to cortisol, we cannot exclude a gender-specific influence of CBG [4] on the serum cortisol level.

The HPA axis set point can be modified through an altered balance between mineralocorticoid and glucocorticoid receptor expression.[23] Animal studies have suggested that patterns in receptor expression develop in a gender-specific manner from birth onwards.[24] In humans, behavioral patterns that impact a child's stress vulnerability have been associated with gender-specific changes in cortisol levels from age 1.5 yr onwards.[11,25] Therefore, even in our sample of normal children, gender-specific effects of stress exposure could be an explanation for our results.[9]

Even subtle disturbances in HPA axis activity have been associated with cardiovascular disease and its risk factors.[26-28] Cardiovascular disease susceptibility is gender-specific,[7,29] which has been suggested to be due to gender differences in HPA axis activity, stress vulnerability and responsivity.[4,30-32] Early in life, developmental plasticity offers the child the capacity to change his HPA axis set point based on stress experiences.[9,33] This ability offers opportunities to withstand early-life challenges, but it has also been suggested to affect disease risk later in life. Accordingly, although the gender differences found in our study were small, these patterns might contribute to gender-specific origins of health and disease.[9]

The major strength of this study is our systematic approach and the effort to contact all authors of eligible publications, enabling us to include the data on 16,551 healthy children. Moreover, articles with a lack of quantitative data were included in our descriptive analysis with the aim to be as complete as possible. The large sample size enabled us to perform a sensitivity analysis, which decreased the heterogeneity between studies. Furthermore, we accounted for this heterogeneity by calculating standardized mean differences, based on the intervention effects relative to the variability observed.[13] Additionally, we chose fixed-effect meta-analysis, because the studies with a large sample size were most likely conducted with greater methodological accuracy.[13] Fixed-effect meta-analysis has the advantage of increasing the impact of large studies on the effect

estimate. For comparison, results of random-effects meta-analyses, which put more weight on studies with small sample sizes, were also included. (Appendix 3).

A limitation of this study is that only a subset of studies (16%) considered gender differences as the primary outcome. In addition, in 22 studies (28%) samples were not collected specifically during mornings. Both could have led to a selection or performance bias, which we accounted for in our sensitivity analysis. Furthermore, 21 articles with data on 3,985 subjects could not be included in the meta-analysis due to lack of gender-stratified quantitative data, while most of these articles reported no significant gender differences. However, funnel plots of the articles included in the meta-analysis were not suggestive of publication bias. Instead, the plots seem to indicate that most articles reported on the nonexistence of gender differences, which might be a result of the common idea that gender differences are nonexistent at this early age. Nonetheless, our meta-analysis shows that significant gender differences are already present early in life. Another limitation is that almost all studies that reported on salivary or serum cortisol used immunoassays. Due to its superior specificity, liquid chromatography-tandem mass spectrometry is the method of choice for steroid hormone analysis.[34] Furthermore, we stratified studies based on the mean age or age range of the study group. Since study samples differed in age range, we have probably included some subjects < 8 yr of age in the 8-18 yr groups, and vice versa. An overview of the age ranges of studies included in the meta-analysis is presented in Appendix 5. Moreover, only a minority of the included studies assessed Tanner pubertal staging. Therefore, we were unable to address the question at which maturational stage the direction of the gender-specific dimorphism in cortisol changes.

CONCLUSIONS

In conclusion, gender differences in HPA axis activity are present early in life, with higher salivary cortisol concentrations in boys. A gender-specific evolution of cortisol metabolism is suggested to be induced by puberty, resulting in lower bioavailability of cortisol in boys. Although results from random-effects analyses were inconclusive for serum cortisol, the gender difference in cortisol production seems to be consistent between genders with age. Future research should take gender differences in HPA axis activity into account, regardless of age. Whether gender differences in stress-induced cortisol levels also exist is unknown and remains to be explored.

REFERENCES

1. Panagiotakopoulos L, Neigh GN. Development of the HPA axis: Where and when do sex differences manifest? Frontiers in Neuroendocrinology 2014; 35:285-302

2. Bourke CH, Harrell CS, Neigh GN. Stress-induced sex differences: Adaptations mediated by the glucocorticoid receptor. Hormones and Behavior 2012; 62:210-218

3. Finken MJ, Andrews RC, Andrew R, Walker BR. Cortisol metabolism in healthy young adults: sexual dimorphism in activities of A-ring reductases, but not 11beta-hydroxysteroid dehydrogenases. J Clin Endocrinol Metab 1999; 84:3316-3321

4. Kajantie E, Phillips DIW. The effects of sex and hormonal status on the physiological response to acute psychosocial stress. Psychoneuroendocrinology 2006; 31:151-178

5. Perogamvros I, Ray DW, Trainer PJ. Regulation of cortisol bioavailability--effects on hormone measurement and action. Nat Rev Endocrinol 2012; 8:717-727

6. Wudy SA, Hartmann MF, Remer T. Sexual dimorphism in cortisol secretion starts after age 10 in healthy children: urinary cortisol metabolite excretion rates during growth. Am J Physiol Endocrinol Metab 2007; 293:E970-E976

7. McCormick CM, Mathews IZ. HPA function in adolescence: Role of sex hormones in its regulation and the enduring consequences of exposure to stressors. Pharmacology Biochemistry and Behavior 2007; 86:220-233

8. Vaillancourt T, Duku E, Decatanzaro D, Macmillan H, Muir C, Schmidt LA. Variation in hypothalamic-pituitary-adrenal axis activity among bullied and non-bullied children. Aggress Behav 2008; 34:294-305

9. Hanson MA, Gluckman PD. Early developmental conditioning of later health and disease: physiology or pathophysiology? Physiol Rev 2014; 94:1027-1076

10. Raivio T, Toppari J, Kaleva M, Virtanen H, Haavisto AM, Dunkel L, Janne OA. Serum androgen bioactivity in cryptorchid and noncryptorchid boys during the postnatal reproductive hormone surge. J Clin Endocrinol Metab 2003; 88:2597-2599

11. Martinez-Torteya C, Muzik M, McGinnis EW, Rosenblum KL, Bocknek EL, Beeghly M, DeCator D, Abelson JL. Longitudinal examination of infant baseline and reactivity cortisol from ages 7 to 16 months. Dev Psychobiol 2015; 57:356-364

12. Dattani MTT, V.; Hindmarsh, P.C. Evaluation of disordered puberty. In: Brook's Clinical Pediatric Endcrinology. Sixth ed.

13. Higgins JPT GSe. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration 2011 Available from www.handbook.cochrane.org;

14. Altman DG, Bland JM. Detecting skewness from summary information. BMJ 1996; 313:1200

15. Hatzinger M, Brand S, Perren S, von Wyl A, von Klitzing K, Holsboer-Trachsler E. Hypothalamic-pituitary-adrenocortical (HPA) activity in kindergarten children: importance of gender and associations with behavioral/emotional difficulties. J Psychiatr Res 2007; 41:861-870

16. Jamieson PM, Chapman KE, Seckl JR. Tissue- and temporal-specific regulation of 11beta-hydroxysteroid dehydrogenase type 1 by glucocorticoids in vivo. J Steroid Biochem Mol Biol 1999; 68:245-250

17. Albiston AL, Smith RE, Krozowski ZS. Sex- and tissue- specific regulation of 11 beta-hydroxysteroid dehydrogenase mRNA. Mol Cell Endocrinol 1995; 109:183-188

18. Loizzo S, Vella S, Loizzo A, Fortuna A, Di Biase A, Salvati S, Frajese GV, Agrapart V, Ramirez Morales R, Spampinato S, Campana G, Capasso A, Galietta G, Guarino I, Carta S, Carru C, Zinellu A, Ghir-

landa G, Seghieri G, Renzi P, Franconi F. Sexual dimorphic evolution of metabolic programming in non-genetic non-alimentary mild metabolic syndrome model in mice depends on feed-back mechanisms integrity for pro-opiomelanocortin-derived endogenous substances. Peptides 2010; 31:1598-1605

19. Wolbold R, Klein K, Burk O, Nussler AK, Neuhaus P, Eichelbaum M, Schwab M, Zanger UM. Sex is a major determinant of CYP3A4 expression in human liver. Hepatology 2003; 38:978-988

20. Hunt CM, Westerkam WR, Stave GM. Effect of age and gender on the activity of human hepatic CYP3A. Biochem Pharmacol 1992; 44:275-283

21. Toogood AA, Taylor NF, Shalet SM, Monson JP. Sexual dimorphism of cortisol metabolism is maintained in elderly subjects and is not oestrogen dependent. Clin Endocrinol (Oxf) 2000; 52:61-66

22. Weaver JU, Taylor NF, Monson JP, Wood PJ, Kelly WF. Sexual dimorphism in 11 beta hydroxysteroid dehydrogenase activity and its relation to fat distribution and insulin sensitivity; a study in hypopituitary subjects. Clin Endocrinol (Oxf) 1998; 49:13-20

23. De Kloet ER, Vreugdenhil E, Oitzl MS, Joels M. Brain corticosteroid receptor balance in health and disease. Endocr Rev 1998; 19:269-301

24. Slotkin TA, Seidler FJ, Wood CR, Lau C. Development of glucocorticoid receptor regulation in the rat forebrain: implications for adverse effects of glucocorticoids in preterm infants. Brain Res Bull 2008; 76:531-535

25. Tout K, de Haan M, Campbell EK, Gunnar MR. Social behavior correlates of cortisol activity in child care: gender differences and time-of-day effects. Child Dev 1998; 69:1247-1262

26. Chanson P, Salenave S. Metabolic syndrome in Cushing's syndrome. Neuroendocrinology 2010; 92 Suppl 1:96-101

27. Rosmond R, Bjorntorp P. The hypothalamic-pituitary-adrenal axis activity as a predictor of cardiovascular disease, type 2 diabetes and stroke. J Intern Med 2000; 247:188-197

28. Watt GC, Harrap SB, Foy CJ, Holton DW, Edwards HV, Davidson HR, Connor JM, Lever AF, Fraser R. Abnormalities of glucocorticoid metabolism and the renin-angiotensin system: a four-corners approach to the identification of genetic determinants of blood pressure. J Hypertens 1992; 10:473-482

29. Hassan-Smith ZK, Morgan SA, Sherlock M, Hughes B, Taylor AE, Lavery GG, Tomlinson JW, Stewart PM. Gender-Specific Differences in Skeletal Muscle 11beta-HSD1 Expression Across Healthy Aging. J Clin Endocrinol Metab 2015; 100:2673-2681

30. Kudielka BM, Buske-Kirschbaum A, Hellhammer DH, Kirschbaum C. HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: impact of age and gender. Psychoneuroendocrinology 2004; 29:83-98

31. Otte C, Hart S, Neylan TC, Marmar CR, Yaffe K, Mohr DC. A meta-analysis of cortisol response to challenge in human aging: importance of gender. Psychoneuroendocrinology 2005; 30:80-91

32. Stoney CM, Davis MC, Matthews KA. Sex differences in physiological responses to stress and in coronary heart disease: a causal link? Psychophysiology 1987; 24:127-131

33. Bosch NM, Riese H, Reijneveld SA, Bakker MP, Verhulst FC, Ormel J, Oldehinkel AJ. Timing matters: long term effects of adversities from prenatal period up to adolescence on adolescents' cortisol stress response. The TRAILS study. Psychoneuroendocrinology 2012; 37:1439-1447

34. Ackermans MT, Endert E. LC-MS/MS in endocrinology: what is the profit of the last 5 years? Bioanalysis 2014; 6:43-57

Included in the descriptive analysis

35.   Klug I, Dressendorfer R, Strasburger C, Kuhl GP, Reiter HL, Reich A, Muller G, Meyer K, Kratzsch J, Kiess W. Cortisol and 17-hydroxyprogesterone levels in saliva of healthy neonates: normative data and relation to body mass index, arterial cord blood ph and time of sampling after birth. Biol Neonate 2000; 78:22-26

36.   Eiden RD, Veira Y, Granger DA. Prenatal cocaine exposure and infant cortisol reactivity. Child Dev 2009; 80:528-543

37.   Plusquellec P, Ouellet-Morin I, Feng B, Perusse D, Tremblay RE, Lupien SJ, Boivin M. Salivary cortisol levels are associated with resource control in a competitive situation in 19 month-old boys. Horm Behav 2011; 60:159-164

38.   Spinrad TL, Eisenberg N, Granger DA, Eggum ND, Sallquist J, Haugen RG, Kupfer A, Hofer C. Individual differences in preschoolers' salivary cortisol and alpha-amylase reactivity: relations to temperament and maladjustment. Horm Behav 2009; 56:133-139

39.   Safarzadeh E, Mostafavi F, Ashtiani MTH. Determination of salivary cortisol in healthy children and adolescents. Acta Medica Iranica 2005; 43:32-36

40.   Isaksson J, Nilsson KW, Lindblad F. The Pressure-Activation-Stress scale in relation to ADHD and cortisol. Eur Child Adolesc Psychiatry 2015; 24:153-161

41.   Kjolhede EA, Gustafsson PE, Gustafsson PA, Nelson N. Overweight and obese children have lower cortisol levels than normal weight children. Acta Paediatr 2014; 103:295-299

42.   Gunnar MR, Wewerka S, Frenn K, Long JD, Griggs C. Developmental changes in hypothalamus-pituitary-adrenal activity over the transition to adolescence: normative changes and associations with puberty. Dev Psychopathol 2009; 21:69-85

43.   Fadalti M, Petraglia F, Luisi S, Bernardi F, Casarosa E, Ferrari E, Luisi M, Saggese G, Genazzani AR, Bernasconi S. Changes of serum allopregnanolone levels in the first 2 years of life and during pubertal development. Pediatr Res 1999; 46:323-327

44.   Ballerini MG, Chiesa A, Scaglia P, Gruneiro-Papendieck L, Heinrich JJ, Ropelato MG. 17alpha-hydroxyprogesterone and cortisol serum levels in neonates and young children: influence of age, gestational age, gender and methodological procedures. J Pediatr Endocrinol Metab 2010; 23:121-132

45.   Parker LN, Sack J, Fisher DA, Odell WD. The adrenarche: prolactin, gonadotropins, adrenal androgens, and cortisol. Journal of Clinical Endocrinology and Metabolism 1978; 46:396-401

46.   Kulasingam V, Jung BP, Blasutig IM, Baradaran S, Chan MK, Aytekin M, Colantonio DA, Adelia K. Pediatric reference intervals for 28 biochemistry analytes on the Roche cobas(registered trademark) 6000 analyzer. Clinical Biochemistry 2010; 43:775

47.   Soldin OP, Hoffman EG, Waring MA, Soldin SJ. Pediatric reference intervals for FSH, LH, estradiol, T3, free T3, cortisol, and growth hormone on the DPC IMMULITE 1000. Clin Chim Acta 2005; 355:205-210

48.   Karbasy K, Lin DC, Stoianov A, Chan MK, Bevilacqua V, Chen Y, Adeli K. Pediatric reference value distributions and covariate-stratified reference intervals for 29 endocrine and special chemistry biomarkers on the Beckman Coulter Immunoassay Systems: a CALIPER study of healthy community children. Clin Chem Lab Med 2016; 54:643-657

49.   Barra CB, Silva IN, Rodrigues TM, Santos JL, Colosimo EA. Morning serum Basal cortisol levels are affected by age and pubertal maturation in school-aged children and adolescents. Horm Res Paediatr 2015; 83:55-61

50. Chalew SA, Nagel H, Burt D, Edwards CR. The integrated concentration of cortisone is reduced in obese children. J Pediatr Endocrinol Metab 1997; 10:287-290

51. Linder BL, Esteban NV, Yergey AL, Winterer JC, Loriaux DL, Cassorla F. Cortisol production rate in childhood and adolescence. J Pediatr 1990; 117:892-896

52. Dorn LD, Burgess ES, Susman EJ, von Eye A, DeBellis MD, Gold PW, Chrousos GP. Response to oCRH in depressed and nondepressed adolescents: does gender make a difference? J Am Acad Child Adolesc Psychiatry 1996; 35:764-773

Included in the meta-analysis

Alghadir AH, Gabr SA, Al-Eisa E. Effects of Physical Activity on Trace Elements and Depression Related Biomarkers in Children and Adolescents. Biological Trace Element Research 2015:1-8

Allen LB, Lu Q, Tsao JCI, Worthman CM, Zeltzer LK. Sex differences in the association between cortisol concentrations and laboratory pain responses in healthy children. Gend Med 2009; 6 Suppl 2:193-207

Apter D, Pakarinen A, Hammond GL, Vihko R. Adrenocortical function in puberty. serum ACTH, cortisol and dehydroepiandrosterone in girls and boys. Acta Paediatr Scand 1979; 68:599-604

Azurmendi A, Pascual-Sagastizabal E, Vergara AI, Munoz JM, Braza P, Carreras R, Braza F, Sanchez-Martin JR. Developmental trajectories of aggressive behavior in children from ages 8 to 10: The role of sex and hormones. Am J Hum Biol 2016; 28:90-97

Bailey D, Colantonio D, Kyriakopoulou L, Cohen AH, Chan MK, Armbruster D, Adeli K. Marked biological variance in endocrine and biochemical markers in childhood: establishment of pediatric reference intervals using healthy community children from the CALIPER cohort. Clin Chem 2013; 59:1393-1405

Belva F, Painter RC, Schiettecatte J, Bonduelle M, Roelants M, Roseboom TJ, Tournaye H, De Schepper J. Gender-specific alterations in salivary cortisol levels in pubertal intracytoplasmic sperm injection offspring. Horm Res Paediatr 2013; 80:350-355

Canalis E, Reardon GE, Caldarella AM. A more specific, liquid-chromatographic method for free cortisol in urine. Clinical Chemistry 1982; 28:2418-2420

Chen FR, Raine A, Rudo-Hutt AS, Glenn AL, Soyfer L, Granger DA. Harsh discipline and behavior problems: The moderating effects of cortisol and alpha-amylase. Biol Psychol 2015; 104:19-27

Cicchetti D, Rogosch FA. The impact of child maltreatment and psychopathology on neuroendocrine functioning. Dev Psychopathol 2001; 13:783-804

Cieslak TJ, Frost G, Klentrou P. Effects of physical activity, body fat, and salivary cortisol on mucosal immunity in children. J Appl Physiol (1985) 2003; 95:2315-2320

Colomina MT, Canals J, Carbajo G, Domingo JL, Domenech E. Salivary cortisol in a young population: Relationship with psychopathological disorders. Research Communications in Biological Psychology and Psychiatry 1997; 22:1-10

Covelli MM, Wood CE, Yarandi HN. Biologic measures as epidemiological indicators of risk for the development of hypertension in an African American adolescent population. J Cardiovasc Nurs 2012; 27:476-484

Daughters SB, Gorka SM, Matusiewicz A, Anderson K. Gender specific effect of psychological stress and cortisol reactivity on adolescent risk taking. J Abnorm Child Psychol 2013; 41:749-758

Davis M, Emory E. Sex differences in neonatal stress reactivity. Child Dev 1995; 66:14-27

de Bruijn ATCE, van Bakel HJA, Wijnen H, Pop VJM, van Baar AL. Prenatal maternal emotional complaints are associated with cortisol responses in toddler and preschool aged girls. Dev Psychobiol 2009; 51:553-563

Dietrich A, Ormel J, Buitelaar JK, Verhulst FC, Hoekstra PJ, Hartman CA. Cortisol in the morning and dimensions of anxiety, depression, and aggression in children from a general population and clinic-referred cohort: An integrated analysis. The TRAILS study. Psychoneuroendocrinology 2013; 38:1281-1298

Elmlinger MW, Kuhnel W, Ranke MB. Reference ranges for serum concentrations of lutropin (LH), follitropin (FSH), estradiol (E2), prolactin, progesterone, sex hormone-binding globulin (SHBG), dehydroepiandrosterone sulfate (DHEAS), cortisol and ferritin in neonates, children and young adults. Clin Chem Lab Med 2002; 40:1151-1160

Forest MG. Age-related response of plasma testosterone, delta 4-androstenedione, and cortisol to adrenocorticotropin in infants, children, and adults. J Clin Endocrinol Metab 1978; 47:931-937

Fransson E, Folkesson L, Bergstrom M, Ostberg V, Lindfors P. Exploring salivary cortisol and recurrent pain in mid-adolescents living in two homes. BMC Psychol 2014; 2:46

Garagorri JM, Rodriguez G, Lario-Elboj AJ, Olivares JL, Lario-Munoz A, Orden I. Reference levels for 17-hydroxyprogesterone, 11-desoxycortisol, cortisol, testosterone, dehydroepiandrosterone sulfate and androstenedione in infants from birth to six months of age. Eur J Pediatr 2008; 167:647-653

Georgopoulos NA, Rottstein L, Tsekouras A, Theodoropoulou A, Koukkou E, Mylonas P, Polykarpou G, Lampropoulou E, Iconomou G, Leglise M, Vagenakis AG, Markou KB. Abolished circadian rhythm of salivary cortisol in elite artistic gymnasts. Steroids 2011; 76:353-357

Ghaziuddin N, Welch K, Greden J. Central serotonergic effects of m-chlorophenylpiperazine (mCPP) among normal control adolescents. Neuropsychopharmacology 2003; 28:133-139

Gunnar MR, Kryzer E, Van Ryzin MJ, Phillips DA. The rise in cortisol in family day care: associations with aspects of care quality, child behavior, and child sex. Child Dev 2010; 81:851-869

Hackney AC, McMurray RG, Judelson DA, Harrell JS. Relationship between caloric intake, body composition, and physical activity to leptin, thyroid hormones, and cortisol in adolescents. Jpn J Physiol 2003; 53:475-479

Honour JW, Jones R, Leary S, Golding J, Ong KK, Dunger DB. Relationships of urinary adrenal steroids at age 8 years with birth weight, postnatal growth, blood pressure, and glucose metabolism. J Clin Endocrinol Metab 2007; 92:4340-4345

Huybrechts I, De Vriendt T, Breidenassel C, Rogiers J, Vanaelst B, Cuenca-Garcia M, Moreno LA, Gonzalez-Gross M, Roccaldo R, Kafatos A, Clays E, Bueno G, Beghin L, Sjostrom M, Manios Y, Molnar D, Pisa PT, De Henauw S. Mechanisms of stress, energy homeostasis and insulin resistance in European adolescents--the HELENA study. Nutr Metab Cardiovasc Dis 2014; 24:1082-1089

Ilias I, Ghizzoni L, Mastorakos G. Orderliness of cortisol, growth hormone, and leptin secretion in short-normal pre-pubertal boys and girls. Med Sci Monit 2009; 15:CR242-CR247

Jones A, Godfrey KM, Wood P, Osmond C, Goulden P, Phillips DIW. Fetal growth and the adrenocortical response to psychological stress. J Clin Endocrinol Metab 2006; 91:1868-1871

Lashansky G, Saenger P, Fishman K, Gautier T, Mayes D, Berg G, Di Martino-Nardi J, Reiter E. Normative data for adrenal steroidogenesis in a healthy pediatric population: age- and sex-related changes after adrenocorticotropin stimulation. J Clin Endocrinol Metab 1991; 73:674-686

Lundberg U, de Chateau P, Winberg J, Frankenhaeuser M. Catecholamine and cortisol excretion patterns in three-year-old children and their parents. J Human Stress 1981; 7:3-11

Lundberg U. Sex differences in behaviour pattern and catecholamine and cortisol excretion in 3-6 year old day-care children. Biol Psychol 1983; 16:109-117

Martikainen S, Pesonen AK, Lahti J, Heinonen K, Feldt K, Pyhala R, Tammelin T, Kajantie E, Eriksson JG, Strandberg TE, Raikkonen K. Higher levels of physical activity are associated with lower hypothalamic-pituitary-adrenocortical axis reactivity to psychosocial stress in children. J Clin Endocrinol Metab 2013; 98:E619-E627

Michels N, Sioen I, Huybrechts I, Bammann K, Vanaelst B, De Vriendt T, Iacoviello L, Konstabel K, Ahrens W, De Henauw S. Negative life events, emotions and psychological difficulties as determinants of salivary cortisol in Belgian primary school children. Psychoneuroendocrinology 2012; 37:1506-1515

Mills RSL, Imm GP, Walling BR, Weiler HA. Cortisol reactivity and regulation associated with shame responding in early childhood. Dev Psychol 2008; 44:1369-1380

Minkley N, Kirchner WH. Influence of test tasks with different cognitive demands on salivary cortisol concentrations in school students. Int J Psychophysiol 2012; 86:245-250

Mrug S, Tyson A, Turan B, Granger DA. Sleep problems predict cortisol reactivity to stress in urban adolescents. Physiology and Behavior 2016; 155:95-101

Nakamura J, Yakata M. Age- and sex-related differences in urinary cortisol level. Clin Chim Acta 1984; 137:77-80

Ong KK, Potau N, Petry CJ, Jones R, Ness AR, Honour JW, De Zegher F, Ibanez L, Dunger DB. Opposing influences of prenatal and postnatal weight gain on adrenarche in normal boys and girls. J Clin Endocrinol Metab 2004; 89:2647-2651

Osika W, Friberg P, Wahrborg P. A new short self-rating questionnaire to assess stress in children. Int J Behav Med 2007; 14:108-117

Perez-Edgar K, Schmidt LA, Henderson HA, Schulkin J, Fox NA. Salivary cortisol levels and infant temperament shape developmental trajectories in boys at risk for behavioral maladjustment. Psychoneuroendocrinology 2008; 33:916-925

Portnoy J, Raine A, Glenn AL, Chen FR, Choy O, Granger DA. Digit ratio (2D:4D) moderates the relationship between cortisol reactivity and self-reported externalizing behavior in young adolescent males. Biological Psychology 2015; 112:94-106

Reynolds RM, Hii HL, Pennell CE, McKeague IW, de Kloet ER, Lye S, Stanley FJ, Mattes E, Foster JK. Analysis of baseline hypothalamic-pituitary-adrenal activity in late adolescence reveals gender specific sensitivity of the stress axis. Psychoneuroendocrinology 2013; 38:1271-1280

Ross JL, Schulte HM, Gallucci WT, Cutler GBJ, Loriaux DL, Chrousos GP. Ovine corticotropin-releasing hormone stimulation test in normal children. J Clin Endocrinol Metab 1986; 62:390-392

Soriano-Rodriguez P, Osiniri I, Grau-Cabrera P, Riera-Perez E, Prats-Puig A, Carbonell-Alferez M, Schneider S, Mora-Maruny C, De Zegher F, Ibanez L, Bassols J, Lopez-Bermejo A. Physiological concentrations of serum cortisol are related to vascular risk markers in prepubertal children. Pediatr Res 2010; 68:452-455

Stroud LR, Papandonatos GD, Williamson DE, Dahl RE. Sex differences in cortisol response to corticotropin releasing hormone challenge over puberty: Pittsburgh Pediatric Neurobehavioral Studies. Psychoneuroendocrinology 2011; 36:1226-1238

Stupnicki R, Obminski Z, Klusiewicz A, Viru A. Pre-exercise serum cortisol concentration and responses to laboratory exercise. European Journal of Applied Physiology and Occupational Physiology 1995; 71:439-443

Susman EJ, Dorn LD, Chrousos GP. Negative affect and hormone levels in young adolescents: Concurrent and predictive perspectives. J Youth Adolesc 1991; 20:167-190

Syme C, Abrahamowicz M, Leonard GT, Perron M, Pitiot A, Qiu X, Richer L, Totman J, Veillette S, Xiao Y, Gaudet D, Paus T, Pausova Z. Intra-abdominal adiposity and individual components of the

metabolic syndrome in adolescence: sex differences and underlying mechanisms. Arch Pediatr Adolesc Med 2008; 162:453-461

Tennes K, Carter D. Plasma cortisol levels and behavioral states in early infancy. Psychosom Med 1973; 35:121-128

Tornhage CJ. Reference values for morning salivary cortisol concentrations in healthy school-aged children. J Pediatr Endocrinol Metab 2002; 15:197-204

Tout K, de Haan M, Campbell EK, Gunnar MR. Social behavior correlates of cortisol activity in child care: gender differences and time-of-day effects. Child Dev 1998; 69:1247-1262

Tsvetkova V. Adrenocortical function after stimulation with synthetic ACTH. Curr Med Res Opin 1977; 4:635-639

Turan B, Tackett JL, Lechtreck MT, Browning WR. Coordination of the cortisol and testosterone responses: A dual axis approach to understanding the response to social status threats. Psychoneuroendocrinology 2015; 62:59-68

Tzortzi C, Proff P, Redlich M, Aframian DJ, Palmon A, Golan I, Muessig D, Wichelhaus A, Baumert U. Cortisol daily rhythm in saliva of healthy school children. Int Dent J 2009; 59:12-18

Vaindirlis I, Peppa-Patrikiou M, Dracopoulou M, Manoli I, Voutetakis A, Dacou-Voutetakis C. 'White coat hypertension' in adolescents: Increased values of urinary cortisol and endothelin. Journal of Pediatrics 2000; 136:359-364

West P, Sweeting H, Young R, Kelly S. The relative importance of family socioeconomic status and school-based peer hierarchies for morning cortisol in youth: an exporatory study. Soc Sci Med 2010; 70:1246-1253

Wudy SA, Hartmann MF, Remer T. Sexual dimorphism in cortisol secretion starts after age 10 in healthy children: urinary cortisol metabolite excretion rates during growth. Am J Physiol Endocrinol Metab 2007; 293:E970-E976

Yu YZ, Shi JX. Relationship between levels of testosterone and cortisol in saliva and aggressive behaviors of adolescents. Biomed Environ Sci 2009; 22:44-49

APPENDIX 1

Appendix 1A: Search strategy for PubMed (14 January 2016)

Search	Query	Records (n)
#1	"Hydrocortisone"[Mesh] OR "Glucocorticoids"[Mesh] OR "11-beta-Hydroxysteroid Dehydrogenases"[Mesh] OR "Tetrahydrocortisone"[Mesh] OR "Tetrahydrocortisol"[Mesh] OR cortisol*[tiab] OR hydrocortison*[tiab] OR epicortisol*[tiab] OR cortifair*[tiab] OR cortril*[tiab] OR glucocorticoid*[tiab] OR beta hydroxysteroid dehydrogenase*[tiab] OR 11 oxoreductase*[tiab] OR 11 oxidoreductase*[tiab] OR 11 hydroxysteroid dehydrogenase*[tiab] OR 11b hydroxysteroid dehydrogenase*[tiab] OR 11 reductase*[tiab] OR 11beta hydroxysteroid dehydrogenase*[tiab] OR tetrahydrocortiso*[tiab] OR "hsd11b2"[tiab] OR "11bhsd2"[tiab] OR "11betahsd2"[tiab] OR 11beta hsd*[tiab] OR hydroxycortisol*[tiab] OR "Circadian Rhythm"[Mesh] OR "twenty four hour"[tiab] OR circadian*[tiab] OR diurnal*[tiab] OR nyctohemeral*[tiab]	250,920
#2	child*[tw] OR schoolchild*[tw] OR infan*[tw] OR adolescen*[tw] OR pediatri*[tw] OR paediatr*[tw] OR neonat*[tw] OR boy[tw] OR boys[tw] OR boyhood[tw] OR girl[tw] OR girls[tw] OR girlhood[tw] OR youth[tw] OR youths[tw] OR baby[tw] OR babies[tw] OR toddler*[tw] OR "Mental Disorders Diagnosed in Childhood"[MeSH] OR teen[tw] OR teens[tw] OR teenager*[tw] OR newborn*[tw] OR postneonat*[tw] OR postnat*[tw] OR perinat*[tw] OR puberty[tw] OR preschool*[tw] OR suckling*[tw] OR picu[tw] OR nicu[tw] OR "Arthritis, Juvenile"[Mesh] OR "Myoclonic Epilepsy, Juvenile"[Mesh] OR "Leukemia, Myelomonocytic, Juvenile"[Mesh] OR "Xanthogranuloma, Juvenile"[Mesh] OR "Juvenile Delinquency"[Mesh] OR "Corneal Dystrophy, Juvenile Epithelial of Meesmann"[Mesh]	3,664,351
#3	"Sex Characteristics"[Mesh] OR "Sex Factors"[Mesh] OR sex characteristic*[tiab] OR sex difference*[tiab] OR sex dimorphism*[tiab] OR sexual dimorphism*[tiab] OR sexual difference*[tiab] OR sexual characteristic*[tiab] OR sex factor*[tiab] OR sexual factor*[tiab] OR sexual dimorphi*[tiab] OR sex influenc*[tiab] OR sexual influenc*[tiab] OR gender*[tiab]	451,894
#4	(#1 AND #2 AND #3)	2,643

Abbreviations: Mesh = Medical subject headings; tiab = words in title OR abstract; tw = words in title, abstract, MeSH and other content related fields

Appendix 1B: Search strategy for Embase.com (14 January 2016)

Search	Query	Records (n)
#1	'hydrocortisone'/exp OR 'glucocorticoid'/de OR '11beta hydroxysteroid dehydrogenase'/exp OR 'tetrahydrocortisone'/exp OR 'tetrahydrocortisol'/exp OR cortisol*:ab,ti OR hydrocortison*:ab,ti OR epicortisol*:ab,ti OR cortifair*:ab,ti OR cortril*:ab,ti OR glucocorticoid*:ab,ti OR ('beta hydroxysteroid' NEAR/3 dehydrogenase*):ab,ti OR (11 NEXT/1 oxoreductase*):ab,ti OR (11 NEXT/1 oxidoreductase*):ab,ti OR ('11 hydroxysteroid' NEXT/1 dehydrogenase*):ab,ti OR ('11b hydroxysteroid' NEXT/1 dehydrogenase*):ab,ti OR (11 NEXT/1 reductase*):ab,ti OR ('11beta hydroxysteroid' NEXT/1 dehydrogenase*):ab,ti OR 'hsd11b2':ab,ti OR tetrahydrocortiso*:ab,ti OR 'hsd11b2':ab,ti OR '11bhsd2':ab,ti OR '11betahsd2':ab,ti OR (11beta NEXT/1 hsd*):ab,ti OR hydroxycortisol*:ab,ti OR ('trier social stress' NEXT/1 test*):ab,ti OR tsst:ab,ti OR (stress NEAR/3 hormone*):ab,ti OR (stress NEAR/3 marker*):ab,ti	**235,221**
#2	adolescen*:ab,ti OR 'adolescence'/exp OR 'adolescent coping orientation for problem experiences'/exp OR 'adolescent development'/exp OR 'adolescent disease'/exp OR 'adolescent health'/exp OR 'adolescent parent'/exp OR 'adolescent pregnancy'/exp OR 'adolescent smoking'/exp OR 'adolescent'/exp OR 'adolescent-family inventory of life events and changes'/exp OR babies:ab,ti OR baby:ab,ti OR 'birth weight'/exp OR boy:ab,ti OR boyhood:ab,ti OR boys:ab,ti OR 'brazelton neonatal behavioral assessment scale'/exp OR 'child abuse'/exp OR 'child advocacy'/exp OR 'child behavior checklist'/exp OR 'child behavior'/exp OR 'child care'/exp OR 'child death'/exp OR 'child health care'/exp OR 'child health'/exp OR 'child nutrition'/exp OR 'child parent relation'/exp OR 'child psychology'/exp OR 'child restraint system'/exp OR 'child safety'/exp OR 'child welfare'/exp OR child*:ab,ti OR 'child'/exp OR 'childhood disease'/exp OR 'childhood mortality'/exp OR 'childhood'/exp OR girl:ab,ti OR girlhood:ab,ti OR girls:ab,ti OR 'high risk infant'/exp OR infan*:ab,ti OR 'infant disease'/exp OR 'infant mortality'/exp OR 'infant nutrition'/exp OR 'infant welfare'/exp OR 'infanticide'/exp OR 'infantile diarrhea'/exp OR 'infantile hypotonia'/exp OR 'juvenile delinquency'/exp OR neonat*:ab,ti OR 'neonatal weight loss'/exp OR 'newborn disease'/exp OR 'newborn morbidity'/exp OR 'newborn period'/exp OR newborn*:ab,ti OR 'newborn'/exp OR nicu:ab,ti OR 'only child'/exp OR paediatr*:ab,ti OR pediatri*:de,ab,ti OR 'pediatric advanced life support'/exp OR 'pediatric anesthesia'/exp OR 'pediatric cardiology'/exp OR 'pediatric hospital'/exp OR 'pediatric intensive care nursing'/exp OR 'pediatric nurse practitioner'/exp OR 'pediatric nursing'/exp OR 'pediatric rehabilitation'/exp OR 'pediatric surgery'/exp OR 'newborn hypoxia'/exp OR 'pediatric ward'/exp OR 'pediatrics'/exp OR perinat*:ab,ti OR 'perinatal development'/exp OR 'perinatal period'/exp OR 'persistent hyperinsulinemic hypoglycemia of infancy'/exp OR picu:ab,ti OR postnat*:ab,ti OR 'postnatal care'/exp OR 'postnatal development'/exp OR 'postnatal growth'/exp OR postneonat*:ab,ti OR preschool*:ab,ti OR puberty:ab,ti OR 'runaway behavior'/exp OR 'school child':ab,ti OR schoolchild*:ab,ti OR 'severe myoclonic epilepsy in infancy'/exp OR suckling*:ab,ti OR teen:ab,ti OR teenager*:ab,ti OR teens:ab,ti OR toddler*:ab,ti OR 'transient hypogammaglobulinemia of infancy'/exp OR youth:ab,ti OR youths:ab,ti	**4,477,134**
#3	'sex difference'/exp OR 'sex ratio'/exp OR ('boy'/exp AND 'girl'/exp) OR (sex NEAR/3 characteristic*):ab,ti OR (sex NEAR/3 difference*):ab,ti OR (sex NEAR/3 dimorphism*):ab,ti OR (sexual NEAR/3 dimorphism*):ab,ti OR (sexual NEAR/3 difference*):ab,ti OR (sexual NEAR/3 characteristic*):ab,ti OR (sex NEAR/3 factor*):ab,ti OR (sexual NEAR/3 factor*):ab,ti OR (sexual NEAR/3 dimorphi*):ab,ti OR (sex NEAR/3 influenc*):ab,ti OR (sexual NEAR/3 influenc*):ab,ti OR gender*:ab,ti OR (boy*:ab,ti AND girl*:ab,ti) OR sex:ab,ti	**1,014,014**
#4	(#1 AND #2 AND #3)	**4,280**

/exp = EMtree keyword with explosion; /de = EMtree keyword without explosion; :ab,ti = words in title or abstract; NEXT/x = words in that order next to each other, x places apart; NEAR/x = words near to each other, x places apart

APPENDIX 2

Risk of bias of studies included in the meta-analysis (See for argumentation Online Supplementary File 2)

Risk of selection bias included: participants' age range, and sex-specific differences in participation or baseline characteristics. Risk of performance bias included: time of sample collection, protocol transparency, and sex-specific differences in protocol compliance. Risk of detection bias included: sex-specific differences in assay methods. Non-parametric distribution of the data was recorded as a risk of other biases. Bias could be assessed as low (i.e., unlikely to alter the results), unclear (i.e., raises doubt about results) or high (i.e., weakens confidence in results). Colored squares indicate: ■ = Low risk ■ = Unclear risk, ■ = High risk of bias.

	Selection bias	Performance bias	Detection bias	Other bias
Bailey 2013	+	-	+	+
Elmlinger 2002	?	?	+	+
Forest 1978	?	?	+	+
Garagorri 2008	+	+	+	+
Lashansky 1991	?	+	?	+
Soriano-Rodriguez 2010	+	?	+	+
Tennes 1973	?	+	-	-
Tsvetkova 1977	?	+	+	+

A. Serum <8 yr

	Selection bias	Performance bias	Detection bias	Other bias
Apter 1979	?	?	+	-
Bailey 2013	+	-	+	+
Elmlinger 2002	?	?	+	+
Ghaziuddin 2003	+	+	+	-
Hackney 200	-	+	+	+
Huybrechts 2014	-	+	+	+
Ilias 2009	-	+	+	-
Lashansky 1991	?	+	?	+
Ong 2004	+	+	+	+
Reynolds 2013	?	+	+	+
Ross 1986	-	-	+	-
Stroud 2011	+	-	+	+
Stupnicki 1995	-	-	+	+
Susman 1991	?	+	+	+
Syme 2008	-	?	?	+
Tsvetkova 1977	?	+	+	+

B. Serum 8–18 yr

	Selection bias	Performance bias	Detection bias	Other bias
Davis 1995	?	-	+	+
De Bruijn 2009	+	-	+	-
Gunnar 2010	+	+	?	-
Mills 2008	?	-	+	-
PÈrez-Edgar 2008	+	+	+	+
Tˆrnhage 2002	?	+	+	-
Tout 1998	+	+	+	+

C. Saliva <8 yr

	Selection bias	Performance bias	Detection bias	Other bias
Alghadir 2009	+	+	+	-
Allen 2009	+	-	+	+
Azurmendi 2016	?	+	-	-
Belva 2013	+	+	+	+
Chen 2014	?	+	+	+
Cicchetti 2001	-	+	+	-
Cieslak 2003	+	-	+	+
Colomina 1997	-	+	+	-
Covelli 2012	-	+	+	-
Daughters 2013	+	-	+	+
Dietrich 2013	?	?	+	+
Fransson 2014	+	+	+	-
Georgopoulos 2011	-	?	-	-
Jones 2006	?	+	+	+
Martikainen 2013	-	+	+	+
Michels 2012	+	+	+	+
Minckley 2012	?	+	+	?
Mrug 2016	-	+	+	-
Osika 2007	+	+	+	-
Portnoy	+	-	+	-
Reynolds 2013	?	+	+	+
Törnhage 2002	?	+	+	-
Turan 2015	+	-	?	-
Tzortzi 2009	+	+	+	+
West 2010	-	+	+	-
Yu 2009	?	-	+	+

D. Saliva 8–18 yr

	Selection bias	Performance bias	Detection bias	Other bias
Lundberg 1981	?	+	+	-
Lundberg 1983	?	+	?	-
Nakamura 1984	-	?	+	+
Wudy 2007	+	+	+	+

E. Urine <8 yr

	Selection bias	Performance bias	Detection bias	Other bias
Canalis 1982	?	?	?	+
Honour 2007	+	+	+	-
Nakamura 1984	-	?	+	+
Vaindirlis 2000	?	+	?	+
Wudy 2007	+	+	+	+

F. Urine 8–18 yr

APPENDIX 3

Forest plots of gender differences per subgroup (Random effect analyses)

Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
0 yr (Davis 1995)	30.1	36.4	18	29.5	32.8	18	9.2%	0.02 [-0.64, 0.67]
3 yr (Gunnar 2010)	4.03	4.58	37	3.89	3.01	54	15.4%	0.04 [-0.38, 0.46]
3.3 yrs (De Bruijn 2009)	2.82	1.86	19	2.45	1.82	27	10.6%	0.20 [-0.39, 0.79]
4 yr (Pérez-Edgar 2008)	11.9	4.7	51	10.8	5	60	16.9%	0.22 [-0.15, 0.60]
4.1yr (Mills 2008)	3.03	2.76	115	2.76	2.76	85	20.6%	0.10 [-0.18, 0.38]
4.3 yr (Tout 1998)	5.79	2.48	38	4.41	1.38	37	13.8%	0.68 [0.21, 1.14]
4yr (Gunnar 2010)	2.92	1.35	11	4.47	4.88	13	6.7%	-0.40 [-1.22, 0.41]
7.3yr (Törnhage 2002)	10.5	5.2	19	5.2	1.88	11	6.8%	1.19 [0.38, 2.00]
Total (95% CI)			**308**			**305**	**100.0%**	**0.23 [-0.01, 0.48]**

Heterogeneity: Tau² = 0.05; Chi² = 13.34, df = 7 (P = 0.06); I² = 48%
Test for overall effect: Z = 1.90 (P = 0.06)

A. Serum <8 yr

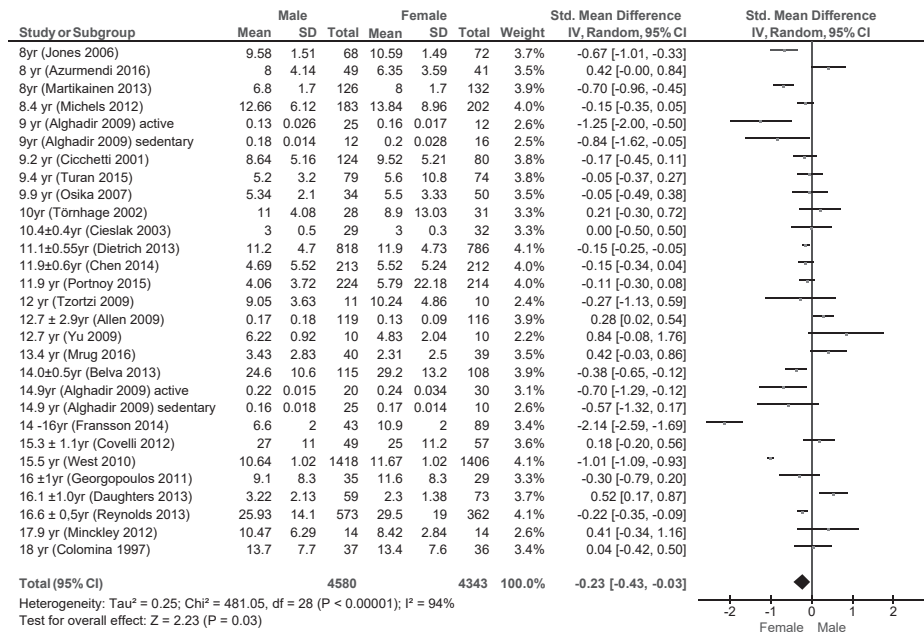| Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
8yr (Jones 2006)	9.58	1.51	68	10.59	1.49	72	3.7%	-0.67 [-1.01, -0.33]
8 yr (Azurmendi 2016)	8	4.14	49	6.35	3.59	41	3.5%	0.42 [-0.00, 0.84]
8yr (Martikainen 2013)	6.8	1.7	126	8	1.7	132	3.9%	-0.70 [-0.96, -0.45]
8.4 yr (Michels 2012)	12.66	6.12	183	13.84	8.96	202	4.0%	-0.15 [-0.35, 0.05]
9 yr (Alghadir 2009) active	0.13	0.026	25	0.16	0.017	12	2.6%	-1.25 [-2.00, -0.50]
9yr (Alghadir 2009) sedentary	0.18	0.014	12	0.2	0.028	16	2.5%	-0.84 [-1.62, -0.05]
9.2 yr (Cicchetti 2001)	8.64	5.16	124	9.52	5.21	80	3.8%	-0.17 [-0.45, 0.11]
9.4 yr (Turan 2015)	5.2	3.2	79	5.6	10.8	74	3.8%	-0.05 [-0.37, 0.27]
9.9 yr (Osika 2007)	5.34	2.1	34	5.5	3.33	50	3.5%	-0.05 [-0.49, 0.38]
10yr (Törnhage 2002)	11	4.08	28	8.9	13.03	31	3.3%	0.21 [-0.30, 0.72]
10.4±0.4yr (Cieslak 2003)	3	0.5	29	3	0.3	32	3.3%	0.00 [-0.50, 0.50]
11.1±0.55yr (Dietrich 2013)	11.2	4.7	818	11.9	4.73	786	4.1%	-0.15 [-0.25, -0.05]
11.9±0.6yr (Chen 2014)	4.69	5.52	213	5.52	5.24	212	4.0%	-0.15 [-0.34, 0.04]
11.9 yr (Portnoy 2015)	4.06	3.72	224	5.79	22.18	214	4.0%	-0.11 [-0.30, 0.08]
12 yr (Tzortzi 2009)	9.05	3.63	11	10.24	4.86	10	2.4%	-0.27 [-1.13, 0.59]
12.7 ± 2.9yr (Allen 2009)	0.17	0.18	119	0.13	0.09	116	3.9%	0.28 [0.02, 0.54]
12.7 yr (Yu 2009)	6.22	0.92	10	4.83	2.04	10	2.2%	0.84 [-0.08, 1.76]
13.4 yr (Mrug 2016)	3.43	2.83	40	2.31	2.5	39	3.5%	0.42 [-0.03, 0.86]
14.0±0.5yr (Belva 2013)	24.6	10.6	115	29.2	13.2	108	3.9%	-0.38 [-0.65, -0.12]
14.9yr (Alghadir 2009) active	0.22	0.015	20	0.24	0.034	30	3.1%	-0.70 [-1.29, -0.12]
14.9 yr (Alghadir 2009) sedentary	0.16	0.018	25	0.17	0.014	10	2.6%	-0.57 [-1.32, 0.17]
14 -16yr (Fransson 2014)	6.6	2	43	10.9	2	89	3.4%	-2.14 [-2.59, -1.69]
15.3 ± 1.1yr (Covelli 2012)	27	11	49	25	11.2	57	3.6%	0.18 [-0.20, 0.56]
15.5 yr (West 2010)	10.64	1.02	1418	11.67	1.02	1406	4.1%	-1.01 [-1.09, -0.93]
16 ±1yr (Georgopoulos 2011)	9.1	8.3	35	11.6	8.3	29	3.3%	-0.30 [-0.79, 0.20]
16.1 ±1.0yr (Daughters 2013)	3.22	2.13	59	2.3	1.38	73	3.7%	0.52 [0.17, 0.87]
16.6 ± 0,5yr (Reynolds 2013)	25.93	14.1	573	29.5	19	362	4.1%	-0.22 [-0.35, -0.09]
17.9 yr (Minckley 2012)	10.47	6.29	14	8.42	2.84	14	2.6%	0.41 [-0.34, 1.16]
18 yr (Colomina 1997)	13.7	7.7	37	13.4	7.6	36	3.4%	0.04 [-0.42, 0.50]
Total (95% CI)			**4580**			**4343**	**100.0%**	**-0.23 [-0.43, -0.03]**

Heterogeneity: Tau² = 0.25; Chi² = 481.05, df = 28 (P < 0.00001); I² = 94%
Test for overall effect: Z = 2.23 (P = 0.03)

B. Serum 8–18 yr

Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1-7d (Elmlinger 2002)	524	190.82	28	569	184.44	17	7.2%	-0.23 [-0.84, 0.37]
8-15d (Elmlinger 2002)	311	99.75	20	342	119.39	20	7.1%	-0.28 [-0.90, 0.35]
0 yr (Forest 1978)	312	168	15	448	187	3	5.0%	-0.76 [-2.03, 0.51]
0 yr (Tennes 1973)	435.9	313.6	23	389.8	227.6	17	7.1%	0.16 [-0.47, 0.79]
0 yrs (Garagorri 2008)	249.41	99.32	80	280.31	124.16	58	7.8%	-0.28 [-0.62, 0.06]
0.4 yr (Lashansky 1991)	342	146	13	353	196	401	7.3%	-0.06 [-0.61, 0.50]
0-2 yr (Bailey 2013)	159.8	135.3	261	190.2	186.2	208	8.1%	-0.19 [-0.37, -0.01]
16d-3y (Elmlinger 2002)	253	95.66	42	271	97.44	44	7.7%	-0.18 [-0.61, 0.24]
3 yr (Lashansky 1991)	337	160	14	281	113	8	6.3%	0.37 [-0.51, 1.25]
4-10 yr (Tsvetkova 1977)	518.69	193.1	9	510.42	157.3	8	6.1%	0.04 [-0.91, 1.00]
4-6y (Elmlinger 2002)	253	143.88	28	218	117.86	23	7.3%	0.26 [-0.29, 0.81]
6.8 yr (Soriano-Rodriguez 2010)	441.4	13.8	106	408.3	13.8	117	7.8%	2.39 [2.05, 2.74]
7-8y (Elmlinger 2002)	400	141.58	26	294	123.47	24	7.2%	0.78 [0.21, 1.36]
2-8 yr (Bailey 2013)	134.41	75.57	110	129.91	71.48	86	7.9%	0.06 [-0.22, 0.34]
Total (95% CI)			**775**			**1053**	**100.0%**	**0.17 [-0.28, 0.63]**

Heterogeneity: Tau² = 0.67; Chi² = 194.87, df = 13 (P < 0.00001); I² = 93%
Test for overall effect: Z = 0.75 (P = 0.46)

C. Saliva <8 yr

Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
8 yr (Apter 1979)	240	80.3	7	226.2	109.5	7	3.6%	0.13 [-0.91, 1.18]
8yr (Lashansky 1991)	254	86	8	229	94	7	3.7%	0.26 [-0.76, 1.28]
8yr (Ong 2004)	210	58	423	215	69	347	4.5%	-0.08 [-0.22, 0.06]
9 yr (Ilias 2009)	112.97	49.14	5	153.26	131.05	6	3.4%	-0.36 [-1.56, 0.84]
9-10y (Elmlinger 2002)	317	165.82	31	328	153.32	40	4.3%	-0.07 [-0.54, 0.40]
6-15 yr (Ross 1986)	74.5	35.9	23	80	44.1	24	4.2%	-0.13 [-0.71, 0.44]
10yr (Stroud 2011)	164.4	5.9	39	146.5	5.99	30	4.1%	2.98 [2.28, 3.68]
11-14 (Tsvetkova 1977)	424.89	171.1	7	408.3	160	7	3.6%	0.09 [-0.95, 1.14]
11y (Elmlinger 2002)	341	113.5	22	364	117.3	23	4.2%	-0.20 [-0.78, 0.39]
12y (Elmlinger 2002)	386	114.29	22	388	117.35	23	4.2%	-0.02 [-0.60, 0.57]
12.5 yr (Susman 1991)	364.7	149.5	56	344.3	148.7	52	4.4%	0.14 [-0.24, 0.51]
13y (Elmlinger 2002)	430	131.12	21	442	103.32	25	4.2%	-0.10 [-0.68, 0.48]
13 yr (Hackney 2003)	397.3	126.9	62	466.3	162.8	63	4.4%	-0.47 [-0.82, -0.11]
13yr (Lashansky 1991)	223	74	10	229	94	11	3.9%	-0.07 [-0.92, 0.79]
14y (Elmlinger 2002)	440	109.18	32	466	131.63	30	4.3%	-0.21 [-0.71, 0.29]
14.6 yr (Ghaziuddin 2003)	1,757.5	1,191.9	12	1,914.7	1,658.2	9	3.9%	-0.11 [-0.97, 0.76]
14.7 yr (Huybrechts 2014)	299.5	1.6	297	305.6	1.7	426	4.5%	-3.67 [-3.91, -3.43]
15 yr (Syme 2008)	546.3	140.7	77	714.6	306.24	82	4.5%	-0.70 [-1.02, -0.38]
15.5yr (Lashansky 1991)	262	80	13	279	77	8	3.9%	-0.21 [-1.09, 0.68]
16y (Elmlinger 2002)	444	111.99	31	388	116.58	40	4.3%	0.48 [0.01, 0.96]
16.6 ± 0,5yr (Reynolds 2013)	551	154	653	590.98	202.79	401	4.6%	-0.23 [-0.35, -0.10]
17 (Stupnicki 1995)	347	130	14	397	125	15	4.1%	-0.38 [-1.12, 0.35]
8-18yr (Bailey 2013)	218.3	98.9	370	215.7	110.7	369	4.5%	0.02 [-0.12, 0.17]
17-19y (Elmlinger 2002)	465	124.74	30	403	103.57	42	4.3%	0.54 [0.07, 1.02]
Total (95% CI)			**2265**			**2087**	**100.0%**	**-0.11 [-0.56, 0.33]**

Heterogeneity: Tau² = 1.12; Chi² = 906.47, df = 23 (P < 0.00001); I² = 97%
Test for overall effect: Z = 0.50 (P = 0.62)

D. Saliva 8–18 yr

Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
3yr (Lundberg 1981)	3.35	1.12	4	2.46	0.44	9	9.9%	1.19 [-0.11, 2.49]
4.5yr (Lundberg 1983)	8.5	1.8	8	11.9	5.5	3	8.6%	-1.02 [-2.45, 0.41]
3-4 yr (Wudy 2007)	20.5	6.63	25	21.2	6.49	25	23.4%	-0.11 [-0.66, 0.45]
5-6 yr (Wudy 2007)	30.5	10.2	25	25.4	7.52	25	23.1%	0.56 [-0.01, 1.13]
5-10 yr (Nakamura 1984)	14.8	3.6	7	9.98	3.2	8	11.6%	1.34 [0.18, 2.50]
7-8 yr (Wudy 2007)	36.9	12.9	25	31.9	10.9	25	23.3%	0.41 [-0.15, 0.97]
Total (95% CI)			**94**			**95**	**100.0%**	**0.39 [-0.10, 0.87]**

Heterogeneity: Tau² = 0.18; Chi² = 11.11, df = 5 (P = 0.05); I² = 55%
Test for overall effect: Z = 1.55 (P = 0.12)

E. Urine <8 yr

Study or Subgroup	Male Mean	SD	Total	Female Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
8yr (Honour 2007)	2,826.86	774.81	244	2,557.71	655.56	188	47.0%	0.37 [0.18, 0.56]	
9-10 yr (Wudy 2007)	35.5	9.26	25	34.9	12.5	25	8.6%	0.05 [-0.50, 0.61]	
4-15yr (Canalis 1982)	15.2	6.3	14	13	5.6	15	5.0%	0.36 [-0.38, 1.09]	
11-12 yr (Wudy 2007)	54.1	13.8	25	43.6	12.2	25	7.9%	0.79 [0.22, 1.37]	
13 ±3yr (Vaindirlis 2000)	94	32	12	69	33	10	3.6%	0.74 [-0.13, 1.61]	
13-14 yr (Wudy 2007)	53.2	21.4	25	49.5	15.1	25	8.5%	0.20 [-0.36, 0.75]	
14-19yr (Nakamura 1984)	40.9	23	6	33.3	3.1	7	2.3%	0.45 [-0.66, 1.56]	
15-16 yr (Wudy 2007)	68.5	24	25	65.1	22	25	8.5%	0.15 [-0.41, 0.70]	
17-18 yr (Wudy 2007)	75	25.3	25	81.1	32.3	25	8.5%	-0.21 [-0.76, 0.35]	
Total (95% CI)			**401**			**345**	**100.0%**	**0.31 [0.14, 0.48]**	

Heterogeneity: Tau² = 0.01; Chi² = 8.71, df = 8 (P = 0.37); I² = 8%
Test for overall effect: Z = 3.59 (P = 0.0003)

F. Urine 8–18 yr

APPENDIX 4

Funnel plots

A. Serum <8 yr

B. Serum 8–18 yr

C. Saliva <8 yr

D. Saliva 8–18 yr

E. Urine <8 yr

F. Urine 8–18 yr

APPENDIX 5.

Overview of age ranges of studies included in meta-analysis

Study	Mean age ± SD*	Study	Mean age ± SD*
Davis 1995	2 days	Mrug 2016	13.36 ± 0.95 yr
Forest 1978	115.3 ± 120.1 days	Vaindirlis 2000	13 yr ± 3.5 yr
Tennes 1973	3 days	Hackney 2003	13.4 ± 0.9 yr
Garagorri 2008	3 days	Belva 2013	girls 14.0 ± 0.5 yr, boys 14.0 ± 0.4 yr
De Bruijn 2009	38.61 ± 9.4 months	Ghaziuddin 2003	14.6 ± 1.5 yr
Lundberg 1981	3 yr	Huybrechts 2014	14.7 ± 1.2 yr
Gunnar 2010	3.81 ± 0.23 yr	Fransson 2014	ranges 14 - 16 yr
Mills 2008	4.14 ± 0.24 yr	Nakamura 1984	5 – 10 yr and 14 - 19 yr
Tout 1998	mean 4.3 yr	Covelli 2012	15.3 ± 1.1 yr
Lundberg 1983	boys mean 52.3 months, girls 54.9 months	West 2010	15.4 ± 0.4 yr
Soriano-Rodriguez 2010	6.8 ± 0.19 yr	Syme 2008	boys 14.4 ± 1.7 yr, girls 14.4 ± 1.9 yr
Michels 2012	boys 8.44 ± 1.18 yr, girls 8.39 ± 1.20 yr	Daughters 2013	16.1 yr ± 1.0 yr
Apter 1979	range 7.5 – 8.5 yr	Reynolds 2013	16.6 yr ± 0.5 yr
Azurmendi 2016	8 yr	Georgopoulos 2011	boys 15.3 ± 2.0 yr, girls 16.0 ± 1.4 yr
Honour 2007	range 8.2 – 8.4 yr	Minckley 2012	17.86 ± (S.E.M.) 0.096 yr
Jones 2006	range 7- 9 yr	Stupnicki 1995	boys 17.3 ± 0.8 yr, girls 16.4 ± 0.6 yr
Martikainen 2013	boys 8.2 ± 0.3 yr, girls 8.1 ± 0.3 yr	Colomina 1997	range 17.5 – 18.5 yr
Ong 2004	8.2 ± 0.1 yr	Elmlinger 2002	ranges 16 days – 3 yr, 11 yr
Cicchetti 2001	9.24 ± 2.33 yr	Tsvetkova 1977	ranges 4-10 yr and 11-14 yr
Turan 2015 Osika 2007 Ilias 2009 Cieslak 2003	9.38 ± 0.62 yr 9.9 ± 0.6 yr boys 9.5 ± 1.9 yr, girls 9.1 ± 1.3 yr 10.4 ± 0.4 yr	Lashansky 1991	Boys 0.42 ± 0.24, 3.2 ± 1.6, 7.4 ± 1.8, 13.1 ± 1.2, 15.2 ± 1.4 yr Girls 0.42 ± 0.2,2.5 ± 1.5, 9.3 ± 2.2, 12.5 ± 0.9, 15.9 ± 0.7 yr
Stroud 2011	10.5 ± 1.7 yr	Bailey 2013	0.41 ± 0.37, 5.29 ± 1.74, 13.48 ± 3.03 yr
Dietrich 2013 Chen 2014	11.1 yr ± 0.55 yr 11.87 ± 0.60	Wudy 2007	ranges 3-4, 5-6, 7-8, 9-10, 11-12, 13-14, 15-16, 17-18 yr
Portnoy 2015 Susman 1991	11.92 ± 0.59 yr mean boys 12.72 yr, girls 11.99 yr	Alghadir 2009	boys 9.3 ± 1.5 and 14.9 ± 3.7 yr girls 8.96 ± 1.8 and 14.82 ± 4.6 yr
Allen 2009 Yu 2009	12.7 yr ± 2.9 yr 12.6 ± 1.8 yr	Törnhage 2002	median girls 7.4 and 10.3 yr, boys 7.1 and 10.2 yr
Ross 1986 Canalis 1982	range 6-15 yr range 4-15 yr	Tzortzi 2009	ranges boys 10 yr and 3 months – 13 yr and 7 months, girls 10 yr and 3 months – 13 yr and 3 months

*Unless otherwise indicated

ONLINE SUPPLEMENTARY FILES

1. Extracted data of studies included in the meta-analysis (https://doi.org/10.6084/
   m9.figshare.11013239.v1)
2. Risk of bias of studies included in the meta-analysis (https://doi.org/10.6084/
   m9.figshare.11014217.v1)

EVA MOUTON 2019

10

Is HPA-axis reactivity in childhood gender-specific? A systematic review

Jonneke J. Hollanders[*],
Bibian van der Voorn[*],
Joost Rotteveel,
Martijn J.J. Finken
[*] Authors contributed equally to this manuscript

ABSTRACT

Background

In adults, hypothalamus–pituitary–adrenal (HPA) axis activity shows sexual dimorphism, and this is thought to be a mechanism underlying sex-specific disease incidence. Evidence is scarce on whether these sex differences are also present in childhood. In a meta-analysis, we recently found that basal (non-stimulated) cortisol in saliva and free cortisol in 24-h urine follow sex-specific patterns. We explored whether these findings could be extended with sex differences in HPA axis reactivity.

Methods

From inception to January 2016, PubMed and EMBASE.com were searched for studies that assessed HPA axis reactivity in healthy girls and boys aged ≤18 years. Articles were systematically assessed and reported in the categories: (1) diurnal rhythm, (2) cortisol awakening response (CAR), (3) protocolled social stress tests similar or equal to the Trier Social Stress Test for children (TSST-C), (4) pharmacological (ACTH and CRH) stress tests, and (5) miscellaneous stress tests.

Results

Two independent assessors selected 109 out of 6158 records for full-text screening, of which 81 studies (with a total of 14,591 subjects) were included. Studies showed that girls had a tendency towards a more variable diurnal rhythm (12 out of 29 studies), a higher CAR (8 out of 18 studies), and a stronger cortisol response to social stress tests (9 out of 21 studies). We found no evidence for sex differences in cortisol response after a pharmacological challenge or to miscellaneous stress tests.

Discussion

Sex differences in HPA axis reactivity appear to be present in childhood, although evidence is not unequivocal. For a better evaluation of sex differences in HPA axis reactivity, standardization of protocols and reports of stress tests is warranted.

BACKGROUND

Marked gender differences exist in the incidence of several diseases. While men are more prone to obesity, cardiovascular disease, and infectious diseases, women are more susceptible to anxiety, depression, and autoimmune diseases. Sex-specific risks for chronic, non-communicable diseases are thought to result from a combination of genotype, phenotype, and environmental influences during life. Whereas adjustment to environmental challenges is healthy in the short term, developmental plasticity can cause sex-specific adverse effects in the long term.[1]

One of the possible explanations for this sexual dimorphism in disease is a sex-specific reactivity of the hypothalamus-pituitary-adrenal (HPA) axis. HPA-axis functioning can be distinguished by on the one hand the maintenance of homeostasis by controlling basal activity as well as the sensitivity to stressors and, on the other hand, coping with, adapting to, and recovery from reactions to stressors. These processes are controlled by mineralocorticoid and glucocorticoid receptors (MRs and GRs). MRs are mainly involved with basal HPA-axis activity, whereas GRs predominantly regulate HPA-axis reactivity.[2] In animals, receptor expression patterns appear to develop in a sex-specific manner, with sex differences already present at birth.[3] In humans, sexually dimorphic HPA-axis reactivity has also been reported in adulthood: men showed a greater cortisol response to acute real-life or controlled laboratory psychological stress compared to women.[4] Additionally, cortisol responses increased with age in both men and women, but the effect was three-fold stronger in women compared to men, which could possibly be attributed to menopause.[5] These patterns closely resemble those of cardiovascular disease mortality and morbidity.[6] While the setting of HPA-axis functioning results from the balance between MR and GR expression,[2] interactions with the hypothalamus–pituitary–gonadal (HPG) axis are thought to mediate sex-specific stress reactions as well as pathophysiology.[7]

It has previously been hypothesized that disease susceptibility can originate in childhood, possibly through permanent alterations in HPA-axis activity to environmental challenges.[1] We recently showed that basal HPA-axis activity, represented by non-stimulated cortisol concentrations in saliva and free cortisol in 24-h urine, show sexual dimorphism, with a sex-specific change induced by puberty.[8] In addition, gender differences in the reactivity of the HPA-axis have also been described in children,[4,9,10] although evidence is scarce and not systematically reviewed. Therefore, we aimed to examine whether sex-specific differences in HPA-axis reactivity are present in childhood.

To study this sex-specific reactivity of the HPA-axis, we performed a systematic review of the literature. The reactivity of the HPA-axis was defined as the response to either exogenous (e.g., pharmacological, physical, or social) or endogenous (e.g., cortisol awakening response (CAR)) stimuli. In addition, we included diurnal rhythm as a marker of the responsiveness of the HPA-axis, although it functions differently from reactions of

the HPA-axis to stressors. We hypothesized that sex-specific HPA-axis reactivity is already present early in life.

METHODS

Search strategy

PubMed and Embase.com were searched from inception up to January 14, 2016 for studies addressing HPA-axis reactivity in serum or saliva in boys and girls aged ≤18 years by reports of either absolute cortisol values, slopes, AUCs, and/or through visualization of the data in figures. The full search strategy is detailed in Appendix 1 and was based on the index terms or free-text words 'cortisol' or 'glucocorticoid', and 'sex difference' or 'sexual characteristics', and 'child' or 'adolescent'. We excluded studies on children with (psycho)pathology, on synthetic glucocorticoids or with a risk of abnormal HPA axis reactivity (e.g., maltreatment). We did not impose restrictions on the year of publication or study design, apart from reviews and case reports, but we did apply an English language restriction. The review protocol was based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement.

Data collection

Two independent assessors (BvdV and JJH) screened 6,158 titles and abstracts for assessment of sex-specific HPA-axis reactivity. Studies were not assessed blindly. Disagreement between assessors was discussed until consensus was reached. One hundred nine were eligible for full-text screening, of which 81 studies were included in the systematic review.

Figure 1 shows the flowchart of the search. When reports of results were unclear, the authors were contacted (n=4); two authors responded. One author did not reply and one replied but could not provide sufficient data, resulting in exclusion of these studies. Additionally, articles were excluded when 1) no statistical analysis of reactivity was performed (n=9), 2) pharmacological stress tests did not use corticotropin releasing hormone (CRH) and/or ACTH (n=2), 3) HPA-axis reactivity was presented stratified by gender, without analyzing gender differences (n=6), 4) gender was analyzed only as a confounder or effect modifier (n=3), 5) analyses of sex differences were performed with cases and controls combined (n=2) or 5) cortisol reactivity was defined as the variability of cortisol concentrations over several days to months (n=3). Several articles reported on the same cohort. Provided that extra information was presented, all articles were included in the review. Two articles were excluded as no new information was provided compared to other articles describing the same cohort. With respect to case-control studies, we included only the control group.
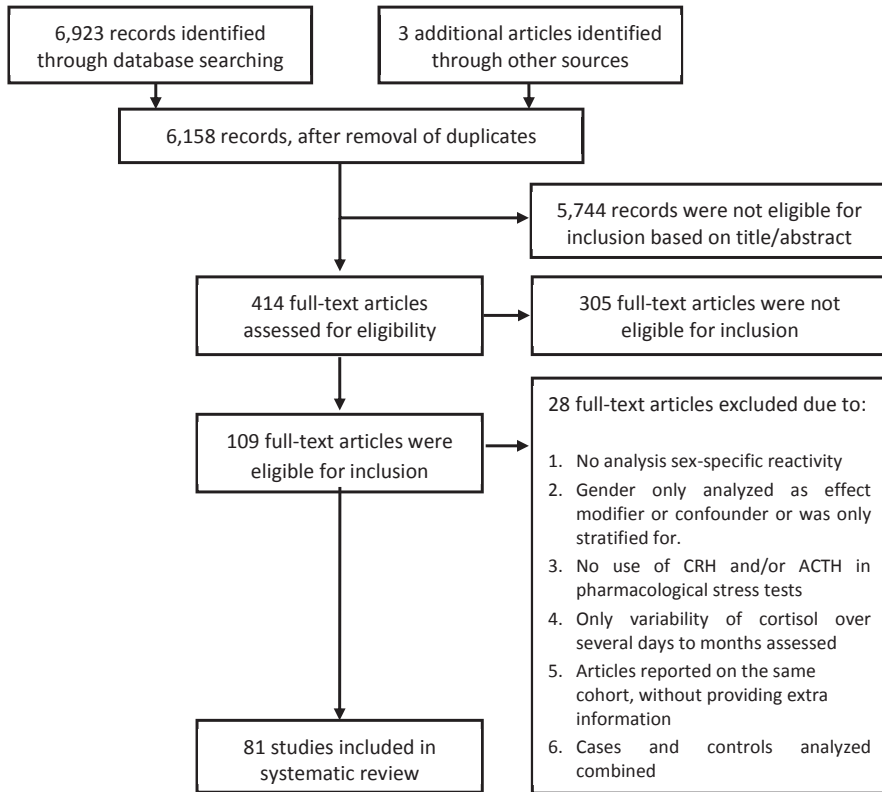
Figure 1: This flowchart presents the different phases of the systematic review, conform the PRISMA-statement. (www.prisma-statement.org)

Data analysis

HPA axis reactivity was classified as follows: 1) diurnal rhythm, 2) CAR, 3) protocolled social stress tests similar or equal to the Trier Social Stress Test for children (TSST-C), 4) pharmacological (ACTH and/or CRH) tests, or 5) miscellaneous stress tests. One assessor (JJH) assessed all the articles and sorted them according to the categories above. Data were extracted from the articles and systematically summarized. If more than one type of reactivity was assessed within one article, the data were included in all applicable categories.

RESULTS

A short overview of all articles is presented in Tables 1 through 5. For a more in-depth summary of the articles, see Online Supplementary File 1. Data on 14,591 subjects were included in this review, with an age range of 31 hours to 18 years.

Table 1: Summary of articles describing sex differences in diurnal rhythmicity

Author (year)	Sample size	Age	Sampling points	Medium	Results
Adam (2010)	230	17.04±0.36 yrs	6x/day on 3 days	Saliva	Lower diurnal cortisol curves in boys
Bae (2015)	138 (70 controls)	10.7±1.7 yrs	3x/day on 3 days	Saliva	Higher levels at awakening, 30 minutes after awakening and higher total daily output in girls; levels in the evening and diurnal slope: no sex differences
Barbosa (2012)	145	8-10yr group: 9.0±0.8 yrs; 11-14yr group: 11.9±1.0 yrs	2x	Saliva	No sex differences, higher diurnal decline in children aged 11-14 yrs old
Bartels (2003)	360	12 yrs	4x/day on 2 days	Saliva	No sex differences; pubertal status not assessed
Carrion (2002)	31	Mean: 10.9 yrs	4x/day on 3 days	Saliva	No sex differences; pubertal status not associated with reactivity
Doom (2013)	110	9.42±0.88 yrs	3x/day on 5 days	Saliva	No sex differences; pubertal status not assessed
Fransson (2014)	157	14-16 yrs	4x (including CAR)	Saliva	Steeper decline in girls
Garcia (1990)	76 (21 controls)	11.2±0.37 yrs	3-hourly during 24 hours	Blood	No sex differences; pubertal status not assessed
Haen (1984)	64	1 mo. to 15 yrs	6 hourly (4x)	Blood	No sex differences; pubertal status not assessed
Jones (2006)	140	7-9 yrs	5x	Saliva	No sex differences; pubertal status not assessed
Kelly (2008)	2995	15.4±0.32 yrs	2x, 30min apart in the morning	Saliva	Steeper decline in girls
Kjolhede (2014)	342	9.5±1.9 yrs	3x/day on 4 days	Saliva	No sex differences; pubertal status not assessed
Knutsson (1997)	235	2.2-18.5 yrs	7x	Blood	No sex differences, except for higher values in girls at pubertal stage 2
Kuhlman (2015)	121	12.8±2.3 yrs	4x/day on 2 days	Saliva	No impact of sex on cortisol at awakening or linear decline, but boys showed less deceleration of the diurnal decline between dinner and bed-time.
Lumeng (2014)	331	3-4 yrs	3x/day on 3 days	Saliva	No sex differences; pubertal status not assessed
Martikainen (2013)	252	8.1±0.3 yrs	7x	Saliva	Higher morning cortisol in girls, no sex difference in nadir

Table 1: Summary of articles describing sex differences in diurnal rhythmicity (continued)

Author (year)	Sample size	Age	Sampling points	Medium	Results
Matchock (2007)	120	Boys: 9, 11 or 13 yrs; girls: 8, 10 or 12 yrs	6x (including CAR)	Saliva	Cortisol peak occurred later in boys than girls during later puberty. Higher morning cortisol in boys at pubertal stage 2. AUCg: no effect of sex, but significant pubertal stage effect
Michels (2012)	385	5-10 yrs	4x (including CAR)	Saliva	No sex differences except somewhat steeper decline in girls (p=0.30)
Morin-Major (2016)	88	14.5±1.8 yrs	4x/day on 2 days	Saliva	Higher AUC in girls
Netherton (2004)	129	12.8±0.19 yrs	2x/day on 4 days	Saliva	Mid-post pubertal girls higher morning cortisol than boys. No sex differences in variance across the four days
Osika (2007)	84	9.9±0.55 yrs	5x (including CAR)	Saliva	No sex differences; pubertal status not assessed
Rosmalen (2005)	1768	11.08±0.55 yrs	3x (including CAR)	Saliva	Higher morning cortisol levels in girls, no sex differences in evening cortisol, already present in pre-pubertal children. Age or pubertal status not associated with cortisol levels.
Ruttle (2013)	346	11, 13 and 15 yrs	3x/day on 3 days	Saliva	Steeper slope in girls at ages 11 and 13, and in longitudinal analyses; higher cortisol levels in girls throughout the day at age 15
Shirtcliff (2012)	357	9, 11, 13 and 15 yrs	3x/day on 3 days	Saliva	Steeper slopes, more curvature in girls. Advancement through puberty: rhythm becomes flatter, especially in girls
Susman (2007)	111	Boys: 9, 11 or 13 yrs; girls: 8, 10 or 12 yrs	6x (including CAR)	Saliva	No sex differences; pubertal status not associated with reactivity
Tzortzi (2009)	21	10-14 yrs	20x (including CAR)	Saliva	No sex differences; pubertal status not assessed
Vaillancourt (2008)	154	147±9.07 mo.	2x/day on 3 days	Saliva	Higher morning levels in girls on Saturday, multilevel regression: consistently higher production in girls
Vanaelst (2013)	355	5-10 yrs	4x/day on 2 days (including CAR)	Saliva	No sex differences; pubertal status not assessed
Williams (2013)	27	9.13±1.41 yrs	3x/day on 2 days (including CAR)	Saliva	Boys exhibited flatter slopes than girls

Table 2: Summary of articles describing sex differences in cortisol awakening response (CAR)

Author (year)	Sample size	Age	Sampling points	Medium	Results
Adam (2010)	230	17.04 ± 0.36 yrs	0 and 40 min after awakening	Saliva	No sex differences; pubertal status not assessed
Bae (2015)	138 (70 controls)	10.7±1.7 yrs	0 and 30 min after awakening	Saliva	Higher levels in girls at awakening and 30 min after awakening, no sex differences in awakening response
Bouma (2009)	644	16.13±0.59 yrs	0 and 30 min after awakening	Saliva	Higher basal levels in girls, no difference in awakening responses
Bright (2014)	47	12-24 months	0 and 30 min after awakening	Saliva	No sex differences; pubertal status not assessed
Dietrich (2013)	1604	11.1±0.55 yrs	0 and 30 min after awakening	Saliva	AUCg and absolute cortisol values higher in girls, AUCi no sex differences
Fransson (2014)	157	14-16 yrs	0, 30 and 60 min after awakening	Saliva	Higher CAR in girls
Hatzinger (2007)	102	4.91±0.44 yrs	0, 10, 20 and 30 min after awakening	Saliva	Higher CAR in girls
Jones (2006)	140	7-9 yrs	0 and 30 min after awakening	Saliva	CAR present in boys, not girls
Kuhlman (2015)	121	12.8±2.3 yrs	0 and 45 min after awakening	Saliva	No sex differences; pubertal status not assessed
Martikainen (2013)	252	8.1±0.3 yrs	0, 15 and 30 min after awakening	Saliva	Higher AUCg in girls, same increase and AUCi
Michels (2012)	385	5-10 yrs	0, 30 and 60 min after awakening	Saliva	No sex differences; pubertal status not assessed
Morin-Major (2016)	88	14.5±1.8 yrs	0 and 30 min after awakening	Saliva	Correlated to sex, higher CAR in girls
Osika (2007)	84	9.9±0.55 yrs	0 and 15 min after awakening	Saliva	No sex differences; pubertal status not assessed
Pruessner (1997)	42	11.16±1.99 yrs	On 3 days: 0, 10, 20 and 30 min after awakening	Saliva	Marginal differences: higher in girls
Susman (2007)	111	Boys: 9, 11 or 13 yrs; girls: 8, 10 or 12 yrs	0, 20 and 40 min after awakening	Saliva	No sex differences; pubertal status not associated with reactivity
Tzortzi (2009)	21	10-14 yrs	From waking: every 20 min until 3 hours after awakening	Saliva	No sex differences; pubertal status not assessed
Vanaelst (2013)	355	5-10 yrs	0, 30 and 60 min after awakening	Saliva	No sex differences; pubertal status not assessed
Williams (2013)	27	9.13±1.41 yrs	0 and 30 min after awakening	Saliva	No sex differences; pubertal status not assessed

Table 3: Summary of articles describing sex differences in protocolled social stress test similar or equal to the TSST-C

Author (year)	Sample size	Age	Sampling points	Medium	Results
Bae (2015)	169 (81 controls)	10.8±1.8 yrs	8x (3 before, 5 after)	Saliva	No sex differences; pubertal status not associated with reactivity
Bouma (2009)	644	16.13±0.59 yrs	5x (2 before, 3 after) (Groningen Social Stress test)	Saliva	Cortisol responses were stronger in boys
Bouma (2011)	553	16.07±0.90 yrs	4x (1 before, 3 after) (Groningen Social Stress Test)	Saliva	Boys had higher cortisol levels on sample 2
De Veld (2012)	158	10.61±0.52 yrs	7x (2 before, 5 after)	Saliva	Cortisol response stronger in girls
Dockray (2009)	111	Boys: 9, 11 or 13 yrs; girls: 8, 10 or 12 yrs	5x, 2 before, 3 after	Saliva	No sex differences; age but not pubertal stage associated with reactivity in girls, no associations in boys.
Evans (2013)	707	13.77±3.56 yrs	After each period/ task, at the middle of the documentary and at the end of it (in fig. 2: 6 samples, 2 before, 4 during/ after) (Social stress tests based on TSST)	Saliva	In children (7-12): lower cortisol reactivity in boys experiencing less emotional warmth Adolescents (13-20): no sex differences
Gunnar (2009)	82	4 ages groups: 9 (9.79±0.16), 11 (11.57±0.15), 13 (13.55±0.46) 15 (15.55±0.47)	10x, 3 before, 7 after	Saliva	No sex differences, except higher cortisol reactivity in girls at age 13
Hostinar (2014)	191	14.4±1.93 yrs	6x (2 before, 4 after) (TSST for groups)	Saliva	No sex differences; higher intercepts and greater anticipatory responses with increasing age, pubertal status not assessed
Hostinar (2015)	81 (40 children, 41 adolescents)	Children: 9.97±0.52 yrs, adolescents: 16.05±0.39 yrs	4x (1 before, 3 after)	Saliva	Stronger response in 9-10 year old girls, no sex differences among adolescents
Ji (2016)	135	Boys: 9, 11 or 13 yrs; girls: 8, 10 or 12 yrs	5x (2 before, 3 after)	Saliva	At wave 3 (each waves separated by 6 months): girls stronger reaction to stressor, no sex differences in recovery
Jones (2006)	140	7-9 yrs	7x (3 before, 4 after)	Saliva	Anticipatory rise in both, further increment in girls
Kudielka (2004)	31	12.1±0.3 yrs	5x, 1 before, 4 after	Saliva	No sex differences; pubertal status not assessed

Table 3: Summary of articles describing sex differences in protocolled social stress test similar or equal to the TSST-C (continued)

Author (year)	Sample size	Age	Sampling points	Medium	Results
Lu (2014)	87	12.7±0.3 yrs	9x, not specified when	Saliva	More negative logAUCi in girls (less increase)
Martikainen (2013)	252	8.1±0.3 yrs	7x (2 before, 5 after)	Saliva	Higher peak, AUCg and AUCi in girls
Martin (2011)	40	16-18 yrs	7x (1 before, 6 after)	Saliva	No sex differences; pubertal status not assessed
Mrug (2016)	84	13.36±0.95 yrs	3x, 1 before, 2 after	Saliva	Higher post-test cortisol and AUCi in girls
Peckins (2012)	124	10.49±1.68 yrs; boys: 9, 11 or 13 yrs; girls: 8, 10 or 12 yrs	5x, 2 before, 3 after	Saliva	No sex differences; pubertal status not associated with reactivity
Portnoy (2015)	446	11.92±0.59 yrs	4x, 1 before, 3 after	Saliva	No sex differences in AUCg; pubertal status not associated with reactivity
Raikkonen (2010)	292	8.1±0.3 yrs	7x (2 before, 5 after)	Saliva	Boys lower than girls
Strahler (2010)	62	6-10 yrs	4x, 1 before, 3 after	Saliva	No sex differences; pubertal status not assessed
Trickett (2014)	303 maltreated, 151 control	Maltreated: 10.84±1.16 yrs; comparison: 11.11±1.15 yrs	6x (2 before, 4 after)	Saliva	Cortisol response blunted in girls compared to boys

Table 4: Summary of articles describing sex differences in pharmacological stress tests

Author (year)	Sample size	Age	Study protocol	Sampling points	Sampling medium	Results
Dahl (1992)	25	10.3±1.6 yrs	CRH challenge: 1µg/kg iv in the late afternoon	9x, 3 before, 6 after	Blood	Greater peak in boys
Dorn (1996)	20 control subjects	15.1±1.0 yrs	CRH challenge: 1µg/kg iv in the evening	12x, 6 before, 6 after	Blood	No sex differences; groups matched for pubertal status, effect not analyzed
Forest (1978)	20 infants, 35 prepubertal children	Infants: 5-365 days, children: 1-12.6 yrs	ACTH test: 500µg/m^2 im at 8:00 and 20:00 on 3 days)	2x, 1 before, 1 after	Blood	No sex differences; pubertal status not assessed
Lashansky (1991)	102	2 months – 17 yrs	ACTH test: 0.25mg iv in the morning	2x, 1 before, 1 after	Blood	No sex differences; decrease in stimulated cortisol levels with puberty, more pronounced in boys
Ross (1986)	21	6-15 yrs	CRH challenge: 1µg/kg iv in the evening	7x, 2 before, 5 after	Blood	No sex differences; pubertal status not associated with reactivity
Stroud (2011)	68	11.6±1.9 yrs	CRH challenge: 1µg/kg iv in the late afternoon	9-10x, 3 before, 6-7 after	Blood	Sex by Tanner differences: girls increases and boys decreases in cortisol with pubertal maturation, girls decreases and boys stable in reactivity. Boys larger peak change.
Tsvetkova (1977)	31	4-14 yrs	ACTH test: 0.5mg im in the morning	2x, 1 before, 1 after	Blood	No sex differences; pubertal status not assessed

Table 5: Summary of articles describing sex differences in miscellaneous stress tests

	Author (year)	Sample size	Age	Study protocol	Sampling points	Medium	Results
0-1 year old	Davis (1995)	36	30.99±8.09 hours	Neonatal Behavior Assessment Scale	5x, 1 before, 4 after test	Saliva	Higher reactivity in boys
	Eiden (2015)	217	9 months	Laboratory Temperament Assessment Battery	4x, 1 before, 3 after test	Saliva	Cortisol increase in boys, not in girls
	Grunau (2010)	32	4.2±1.0 months	Cortisol response after vaccination	3x, 1 before, 2 after	Saliva	No sex differences; pubertal status not assessed
1-7 years old	De Weerth (2013)	42	68.0±4.3 months	CREST paradigm	6x (2 before, 4 after)	Saliva	No sex differences; pubertal status not assessed
	Gunnar (2010)	151	3.81±0.23 years	Daycare attendance	2x/day on 2 days	Saliva	No sex differences; pubertal status not assessed
	Hatzinger (2007)	102	4.91±0.44 years	MSSB	5x (2 before, 3 after)	Saliva	Higher reactivity in girls
	Kryski (2013)	409	40.72±3.51 months	Matching task	6x (1 before, 5 after)	Saliva	No sex differences; pubertal status not assessed
	Mills (2008)	214	4.14±0.24 years	Easy and difficult matching tasks	6x, 1 before, 5 after	Saliva	Further decreases in boys after initial decrease for both sexes
	Plusquellec (2011)	376	18.85±0.74 months	Two unfamiliar situations (clown and robot)	2x, 1 before, 1 after	Saliva	No sex differences; pubertal status not assessed
	Spinrad (2009)	84	54.07±0.97 months	Preschool Laboratory Assessment Battery	3x, 1 before, 2 after	Saliva	No sex differences; pubertal status not assessed
	Yong Ping (2014)	94	29.9±1.1 months	Maternal separation	4x (2 before, 2 after)	Saliva	No sex differences; pubertal status not assessed

	Study	N	Age	Stressor	Sampling	Medium	Findings
Psychological stress	Daughters (2013)	132	16.1±1.0 years	Behavioral Indicator of Resiliency to Distress	4x, 1 before, 3 after	Saliva	Boys: higher baseline, greater peak. No sex differences in AUCg.
	Hackman (2012)	180	12-14 years	Parent-Adolescent Conflict Discusison	3x (2 before, 1 after)	Saliva	No sex differences; pubertal status not assessed
	Minkley (2012)	93	17.86±0.096 years	Examination challenge (reproduction of knowledge, or transfer and problem-solving)	2x, 1 before, 1 after	Saliva	Not statistically significant, but higher increases in boys. More in reproduction of knowledge group, but also greater in transfer and problem-solving group.
	Zijlmans (2013)	52	12.5±1.21 years	Social Evaluative Stress Test	7x, 1 before, 6 after	Saliva	Higher reactivity in boys
Physical stress	Allen (2009)	235	12.7±2.9 years	Laboratory Pain Tasks	Saliva: 3x, 1 before, 2 after Blood: 2x (after)	Saliva / blood	No sex differences; pubertal status not associated with reactivity
	Chiodo (2011)	16	Boys: 14±0 years, girls: 13±1 years	Taekwondo competition	5x (2 before, 3 after)	Saliva	Lower overall values in girls, but higher peak.
	Covelli (2012)	106	15.3±1.1 years	Cold water hand immersion	2x, 1 before, 1 after	Saliva	No sex differences; pubertal status not assessed
	Frias (2000)	48	13-17 years	Acute alcohol intoxication	1x (after); controls as reference	Blood	More pronounced increase in girls
	Gecgelen (2012)	40	10.9-14.7 years	Rapid maxillary expansion	13x, 1 before, 3 after, and 9 during a period of treatment	Saliva	No sex differences; pubertal status not assessed
	Khilnani (1993)	98	2-20 years	Elective surgery	2x, 1 before, 1 after	Blood	No sex differences; pubertal status not assessed
	Kuhlman (2015)	121	12.8±2.3 years	Socially evaluated cold pressor test	7x (2 before, 5 after)	Saliva	No sex differences; pubertal status not assessed
	Lopez-Duran (2015)	115	12.79±2.26 years	Socially evaluated cold pressor test	8x (2 before, 6 after)	Saliva	No sex differences; pubertal status not assessed
	Stupnicki (1995)	29	Boys: 17.3±0.8, girls 16.4±0.6 years	Exercise	2x, 1 before, 1 after	Blood	Boys decrease in cortisol, girls increase in cortisol after exercise
	Yfanti (2014)	97	89.73±15 months	Dental treatment	5x, 1 before, 4 after	Saliva	No sex differences; pubertal status not assessed

≥7 years old

Diurnal rhythm

Twenty-nine studies (with the data of 8,971 subjects) described diurnal rhythmicity and/ or decline of cortisol throughout the day in children, of which 15 studies reported no significant sex differences.[11-25] Fourteen studies reported significant sex differences, of which 12 reported higher cortisol levels and/or a steeper decline over the day in girls. Both Adam et al. [26] (n=230, age: 17.04±0.36 years) and Williams et al. [27] (n=27, age: 9.13±1.41 years) reported a steeper diurnal cortisol curve in girls. Morin-Major et al. [28] (n=88, age: 14.5±1.8 years) found a higher area under the curve as measured from the ground (AUCg) in girls. Martikainen et al. [29] (n=252, age: 8.1±0.3 years) reported a higher cortisol level at awakening in girls, while there was no difference between sexes at nadir, suggesting a steeper cortisol decline over the day in girls compared to boys. This was also found by Rosmalen et al. [30] (n=1768, age: 11.08±0.55 years) , who found this to be already present pre-pubertally, while age and pubertal status were not associated with diurnal rhythm. Fransson et al. [31] (n=157, age: 14-16 years) found a higher cortisol level at awakening and a steeper diurnal decline in girls. Kelly et al. [32] (n=2,995, age: 15.4±0.3 years) found a greater decrease in cortisol concentration in girls as compared to boys between +/- 9 a.m. and 9:30 a.m. Ruttle et al. [33] (n=346, age: 11, 13 and 15 years) and Shirtcliff et al. [34] (n=357, age: 9, 11, 13 and 15 years) examined the same cohort. Ruttle et al. found a significantly steeper diurnal decline in girls aged 11 and 13 years. At age 15, gender differences in cortisol slope had disappeared, although girls had higher cortisol levels throughout the day. Shirtcliff et al. found similar differences, with higher cortisol and steeper slopes, as well as more curvature, in girls. Moreover, the circadian rhythm became flatter with advancing puberty, particularly among girls. Vaillancourt et al. [35] (n=154, age: 147±9.1 months) examined morning and evening cortisol levels on Monday, Thursday and Saturday. They only found a higher cortisol concentration in girls on Saturday morning. Moreover, after modeling the circadian pattern, they found that girls consistently had higher cortisol levels than boys throughout the day. Bae et al [36] (n=138, age: 10.7±1.7 years) found higher cortisol levels in girls at awakening and 30 min after awakening, as well as a higher total daily output. However, no sex differences were found with regard to diurnal slope or evening levels. Netherton et al. [37] (n=129, age: 12.8±0.19 years) found higher morning cortisol levels in mid- to post-pubertal girls compared to boys, but no sex differences were found in evening cortisol levels. In pre- to early- pubertal children, no sex differences were found in either morning or evening cortisol levels. Contrastingly, Kuhlman et al. [38] (n=121, age: 12.8±2.3 years) reported no sex differences in cortisol levels at awakening or in linear decline, although girls showed more deceleration of the diurnal decline between dinner and bedtime than boys. Matchock et al. [39] (n=120, age: boys: 9, 11 or 13 years; girls: 8, 10 or 12 years) found an earlier cortisol peak in the morning in girls, and at pubertal stage 2 a lower morning cortisol levels in girls. However, although a pubertal stage effect was found, there were no sex differences in the AUCg.

CAR

Eighteen studies (with the data of 3,549 subjects) described the CAR in children. Nine studies did not find differences between boys and girls,[15-18,21,26,27,38,40] although four of these [15-18] studied the CAR as part of the diurnal rhythm, and did not perform separate analyses for the CAR, with therefore limited data available on the CAR. Additionally, Michels et al. [18] (n=385, age: 5-10 years) and Vanaelst et al. [21] (n=355, age: 5-10 years) reported on the same cohort, and Osika et al. [15] (n=84, age: 9.9±0.55 years) only took samples between 0 and 15 minutes after awakening. Nine studies found significant differences in CAR between sexes, of which eight found a higher CAR in girls. Martikainen et al. [29] (n=252, age: 8.1±0.3 years) found a higher peak after awakening in girls, as well as a higher AUCg. However, the awakening response (i.e., the peak value after awakening minus the value immediately after awakening) as well as the AUC increase (AUCi) were not significantly different between the sexes. This was also found by Bouma et al. [41] (n=644, age: 16.1±0.6 years) and Dietrich et al. [42] (n=1604, age: 11.1±0.6 years), who reported on the same cohort (albeit at different ages) and found higher morning cortisol concentrations in girls, but a similar response to awakening in boys and girls, manifesting as a higher AUCg in girls but a similar AUCi between sexes. Additionally, Bae et al. [36] (n=138, 10.7±1.7 years) found higher cortisol levels in girls at awakening and 30 minutes after awakening, although they did not find sex differences in the AUCg. Fransson et al. [31] (n=157, age: 14-16 years) and Hatzinger et al. [43] (n=102, age: 4.9±0.4 years) both found a higher CAR in girls, and Pruessner et al. [44] (n=42, age: 11.2±2.0 years) showed a tendency towards larger increases in girls compared to boys. Morin-Major et al. [28] (n=88, age: 14.5±1.8 years) found a correlation between the CAR and sex, with a higher CAR in girls. Contrastingly, Jones et al. [14] (n=140, age: 7-9 years) found the CAR to be absent in girls, but present in boys.

Protocolled social stress tests similar or equal to the TSST-C

Twenty-one studies (with the data of 3,500 subjects) examined responses to standardized social stress tests. Eighteen used the TSST-C (validated in children aged ≥7 years), while three used other laboratory-based social stress tests that closely resemble the TSST-C [41,45,46]: the Groningen Social Stress Test (GSST) which consisted of a 6-minute speech, a brief interlude and a subtracting task, and a psychosocial stress test which consisted of a mental arithmetic task, a public speaking task and a computer mathematics task. Eight studies, of which two studied the same cohort, did not find sex differences,[36,47-53] while 13 did find sex differences. Ji et al. [54] (n=135, age: boys: 9, 11 or 13 years; girls: 8, 10 or 12 years) reported on the same cohort as Dockray et al. [48] and Peckins et al. [50], who did not find sex differences. However, Ji et al. found that at wave 3, where each wave is separated by six months, girls had a stronger cortisol response to the stressor, although they did not find sex differences with regard to cortisol recovery. Raikkonen et al. [55]

(n=292, age: 8.1±0.3 years) and Martikainen et al. [29] (n=252, age: 8.1±0.3 years) reported on the same cohort, and found a higher peak after stress and higher AUCs (both ground and increase) in girls, while no pre-test differences were found. De Veld [56] (n=158, age: 10.61±0.52 years) found a stronger cortisol response in girls. Jones et al. [14] (n=140, age: 7-9 years) found an anticipatory rise in cortisol in both sexes, but only an additional increase after the TSST-C in girls. Evans et al. [45] (n=707, age: 13.8±3.6 years) found that girls aged ≤12 years displayed higher cortisol reactivity to the psychological stress test, while sex differences were not present in subjects aged 13-20 years. A similar result was found by Hostinar et al. [57] (n=81, age: 9.97±0.52 (children) and 16.05±0.39 (adolescents) years), who found a stronger cortisol response in girls at ages 9 to 10, and no sex differences among the adolescents. Gunnar et al. [58] (n=82, age: 9, 11, 13 and 15 years) found a significantly higher AUCi in girls in response to the TSST-C at age 13, while no sex differences were found at ages 9, 11 and 15 years. Mrug et al. [59] (n=84, age: 13.4±1.0 years) found a higher cortisol 55 min post-test as well as a greater AUCi in girls. On the other hand, Lu et al. [60] (n=87, age: 12.7±0.3 years) found a significantly more negative logAUCi in girls, indicative of a smaller increase in cortisol in girls compared to boys after the TSST-C, and Trickett et al. [61] (n=151 controls, age 11.11±1.15 years) found a blunted cortisol response in girls compared to boys. Additionally, Bouma et al. [41] (n=644, age 16.1±0.6 years), who used the GSST, found lower cortisol responses in girls compared to boys, which was further specified in a study published by Bouma et al. in 2011 [46] (n=553, age: 16.07±0.90 years), who found lower cortisol levels in girls on the first sample after completing the GSST.

Pharmacological stress tests

Seven studies (with the data of 322 subjects) investigated cortisol responses to pharmacological ACTH or CRH. Five studies (3 with ACTH, 2 with CRH) did not find significant sex differences[62-66] and 2 studies found a smaller cortisol increase in girls. Stroud et al. [67] (n=68, age: 11.9±1.9 years), who performed a CRH challenge with 1µg/kg human CRH, found a smaller increase from baseline in girls compared to boys for all Tanner pubertal stages. Additionally, sex-specific pubertal changes were observed, with a baseline cortisol that increased in girls and decreased in boys with advancing puberty. Moreover, girls showed decreases in reactivity/recovery rates (in µg/dL/min), as well as increases in total cortisol response (AUCg) and time to peak cortisol levels with pubertal maturation. Boys, on the other hand, showed little change in reactivity/recovery rates and no changes across puberty for the other parameters. Dahl et al. [68] (n=25, age: 10.3±1.6 years) also performed a 1µg/kg human CRH challenge, and found a smaller increase in cortisol concentration in girls compared to boys.

Miscellaneous stress tests

Twenty-five studies (with the data of 3,004 subjects) performed a wide range of other stress tests.

Three studies were performed in infants aged <1 year (with the data of 285 subjects),[69-71] of which two found a lower cortisol reactivity in girls: Davis and Emory [69] (n=36, age: 31.0±8.1 h), who used the Neonatal Behavior Assessment Scale, and Eiden et al. [70] (n=217, age: 9 months), who used the Laboratory Temperament Assessment Battery.

Eight studies (with the data of 1,472 subjects) were performed in children aged 1-7 years, of which six [72-77] found no sex differences. Hatzinger et al. [43] (n=102, age: 4.9±0.4 years) used the MacArthur Story Stem Battery and found a higher reactivity in girls. Mills et al. [78] (n=214, age: 4.1±0.2 years) used easy and difficult matching tasks with standardized failure and success. They found decreases in cortisol concentrations in both sexes up to 15 min post-stressor but only further decreases in boys.

Fourteen studies (with the data of 1,247 subjects) assessed stress in children aged ≥7 years using miscellaneous protocols. Four studies performed psychological stress tests: one found no sex differences [79], while three found lower reactivity in girls. Zijlmans et al. [80] (n=52, age: 12.5±1.2 years) used a computerized testing paradigm, the social evaluative stress test (SEST), containing elements of social evaluation, unpredictability and uncontrollability. A lower reactivity was found in girls. Daughters et al. [81] (n=132, age: 16.1±1.0 years) used the Behavioral Indicator of Resiliency to Distress (BIRD) and found no cortisol increase and slower cortisol decrease in girls, while there were no sex differences in AUCg. Minkley and Kirchner [82] (n=93, age: 17.9±0.1 years) used two knowledge tests aimed at testing "reproduction of knowledge" or "transfer and problem-solving". A lower reactivity was found in girls, although this was not statistically significant. Ten other studies assessed cortisol reactivity to physical stressors, of which seven did not find sex differences,[38,83-88] of which two reported on the same cohort.[38,88] Chiodo et al[89] (n=16, age: boys: 14±0 years, girls: 13±1 years) used a Taekwondo competitions as stressor, and found lower overall values in girls, although they did exhibit a higher peak compared to boys. Stupnicki et al. [90] (n=29, age: 16-17 years) used physical exercise and found a cortisol increase after physical exercise in girls, compared to a decrease in boys. Frias et al. [91] (n=48, age: 13-17 years) assessed cortisol reactivity after acute alcohol intoxication (AAI). Both boys and girls showed an increase in cortisol concentrations after AAI compared to controls, but this increase was more pronounced in girls, although this was not statistically tested.

DISCUSSION

In this systematic review, we found that sex differences in HPA-axis reactivity are suggested to be present in childhood. In general, with regard to diurnal rhythm, the CAR and social stress tests, around 50% of the studies, notably the larger ones, found sex differences, of which approximately 80% found a more variable diurnal rhythm, a higher CAR, and/or a stronger cortisol response to social stress tests in girls, suggestive of a more variable HPA-axis. We found no evidence for a sex difference in cortisol response after a pharmacological challenge, with only two out of 7 studies reporting a higher cortisol response in boys. Findings from studies addressing sex differences in cortisol response after miscellaneous (social or physical) stress tests were inconsistent, due to different types of stressors applied.

In total, 12 out of 29 studies found a more variable diurnal rhythm in girls, while 2 found this in boys and 15 did not find sex differences. A higher CAR in girls was found in 8 out of 18 studies, although 1 study found a higher CAR in boys and 9 studies found no sex differences. Girls had a stronger cortisol response to social stress tests in 9 out of 21 studies, whereas boys had a stronger response in 4 studies and no sex differences were found in 8 studies. Therefore, although results are suggestive of a more responsive HPA-axis in girls during childhood, these results must be interpreted with caution as the evidence is not unequivocal. However, the sample sizes of the studies that found sex differences were on average larger, while the studies that did not find sex differences more often had a sample size <100.

Our results differed considerably with findings from studies in adults. Notably, psychological stress studies in adults either found no gender difference or a more pronounced cortisol response in men.[4] This difference might be explained by gonadal hormones, more specifically estrogens. In childhood, as we have shown in this review, cortisol reactivity appears to be more pronounced in females. However, other research has shown that in adults, females were found to exhibit attenuated cortisol responses to stress, and males displayed a higher cortisol reactivity.[4] Consequently, it could be hypothesized that post-menopausal women once again show a stronger cortisol response to stress compared to men of the same age. Otte et al.[5], who performed a meta-analysis to evaluate and quantify age-related changes in cortisol response, found a three-fold higher increase in cortisol reactivity with aging in women compared to men. However, studies examining cortisol reactivity in elderly subjects are inconclusive with regard to gender differences.[92-95]

According to the Developmental Origins of Health and Disease (DOHaD) hypothesis, disease susceptibility arises early in development[1] and might be mediated by HPA-axis (re)activity. Dysfunctional (hypo- or hyperreactive) HPA-axis responses have previously been associated with cardiovascular disease risk.[96] In addition, more subtle differences

in early HPA-axis settings can also contribute to sex-specific disease risks throughout life.[10,97]

Sex differences in HPA-axis reactivity might be due to interactions between the HPA- and HPG-axes, and several mechanisms have been proposed. Estradiol has been shown to enhance, while testosterone inhibited CRH gene transcription in the hypothalamus.[9] In addition, estradiol has been found to sensitize the pituitary, thereby increasing the ACTH response, while progesterone seemed to oppose this effect.[9] Moreover, estrogen receptors (ERs) are widely expressed throughout the brain, especially in the limbic system. Although not unequivocal, the distribution of the ER subtypes α and β, which have opposing actions on the HPA-axis,[98] is probably sex-dependent.[99] In rats, gender differences in the expression of ERs were already present early in life.[100] It is possible that sex differences in the balance and distribution of ERα and ERβ in the brain are already present before puberty as a result of priming [1] or genetics, which subsequently change after the onset of puberty. In addition, the sensitivity of the adrenal cortex to ACTH is suggested to be increased in young women,[9] while estrogens were found to increase the production of corticosteroid-binding globulin (CBG),[101] decrease glucocorticoid receptor (GR) expression and activation,[7] and lower hepatic clearance of cortisol by inhibition of A-ring reduction .[102] In contrast, testosterone was found to inhibit the release of ACTH, while progesterone possibly acts as a glucocorticoid antagonist.[9,47,103] However, estrogens seem to have different effects in (postmenopausal) women and men,[104-106] and ACTH responses to a TSST after 2 weeks of DHEA or placebo treatment was found to be equal for women treated with DHEA to those of men, but increased compared to women taking placebos.[4] These HPA-/HPG-axes interactions might explain why the sex differences in HPA-axis reactivity that we found in children are not corroborated by studies in adults. Moreover, some of the included studies in this review took pubertal status into account,[13,24,30,33,34,36,37,39,45,48,50-52,64,65,67,83,94]. Although different (sex-specific) effects of pubertal status on cortisol reactivity were found, HPA-/HPG-axes interactions might nevertheless play a role in the possible sex-specific changes in HPA-axis reactivity throughout puberty. Furthermore, levels of estradiol and progesterone are highly variable in post-menarcheal girls and, there, HPA-HPG-axes interactions might even fluctuate across the menstrual cycle.

The different natures and effects of the applied stressors are something to take into account when assessing HPA-axis reactivity. Different types of stressors activate different levels along the HPA-axis: standard ACTH tests stimulate the adrenals directly, while psychological tests are indirect stimuli of the adrenal cortex through activation of the limbic system. Moreover, the diurnal rhythm and CAR are largely controlled by the suprachiasmatic nucleus, which influences CRH release from the paraventricular nucleus.[107] Additionally, males seem to have a "fight or flight" reaction, with a stronger response when confronted with an achievement challenge (in which you can succeed

or fail at a task), while women show a "tend or befriend" response, and therefore seem to be more sensitive to stress tests that incorporate social rejection or peer pressure.[81,108,109] This might be due to the previously mentioned HPA-/HPG-axes interactions, as well as possible sexually dimorphic site-specific GR and MR expression patterns in the brain.[2,110] Consequently, when designing a study, it is important to realize what type of stress and which level of the HPA-axis is aimed to be tested. Subsequently, the effect of gender on that specific type of stressor should be taken into account. We recommend using standardized protocols, since gender-specific effects on HPA-axis reactivity have been best described with regard to standardized stress protocols.

Additionally, comparing the results of the studies included in our systematic review was hampered by the fact that data were collected and presented in numerous ways. For the same reason, it was impossible to perform a meta-analysis. Moreover, only limited information was often provided, and it is therefore possible that (subtle) sex differences were not found. This was the case for all categories of HPA-axis reactivity discussed in this review. In order to draw more precise conclusions concerning gender differences in HPA-axis reactivity in childhood, we wish to argue using standardized protocols, as well as a standardized presentation of results for future studies on HPA-axis reactivity. Seeman and Robbins [111] have defined stress resiliency as "the overall pattern of HPA response to challenge", which includes the rate of initial response, the magnitude of the response, and the rate of recovery of the HPA-axis. In order to be able to draw conclusions on all of these aspects, and to enable unbiased, quantitative comparisons, reporting data on HPA-axis reactivity should take all of these aspects into account. This can be done by both reporting absolute cortisol values (e.g., minimum and maximum cortisol levels) as well as derived variables (e.g., time to peak/recovery, delta cortisol, ascending/descending slopes and areas under the curve), preferably analyzing sex differences for all these parameters. This will allow a full appreciation and overview of the course followed by cortisol from pre- to post-stressor.

Our review has several strengths and limitations. Our strengths lie in the systematic and extensive search performed, which has resulted in the inclusion of 81 studies. Our review is limited not only by the previously mentioned concerns, but also by the broad range in ages as well as the lack of (reliable) establishment of pubertal stage in the majority of the included articles. Although several studies mention an effect of age or pubertal status on cortisol reactivity,[13,22,24,30,33,34,36,37,39,45,48,50-52,54,57,58,64,65,67,83,94] findings are conflicting between the articles. Moreover, we ourselves were unable to draw any conclusions with regard to age or pubertal status, due to the heterogeneous ways of analyzing these effects as well as limited power within studies. Moreover, pubertal status was often assessed through self-report, which has poor reliability.[112] However, it is possible that the effect of age and/or pubertal status can partly explain our unequivocal conclusions regarding sex differences, as was previously suggested by Jessop and Turner-Cobb.[10]

Aside from standardizing the collection and presentation of data, we therefore urge to also always take age and pubertal status into account. This is in line with a recent study in adults, which showed that adjusting for sex hormones significantly alters sex-specific cortisol profiles.[113]

In conclusion, we found that gender differences in HPA-axis reactivity appear to be present in childhood, suggestive of priming of the HPA-axis during early development, although the evidence is not unequivocal. Overall, girls appear to have a more variable diurnal rhythm, a higher CAR, and a higher cortisol response to social stress tests. These differences are not in line with studies in adults, which might be due to changes in gonadal hormones during puberty impacting on HPA axis reactivity. We found various gender differences depending on the type of stressor applied, which stresses the importance of taking the nature of the stressor into account when designing a new study. Moreover, standardization of protocols and reports of results is warranted.

REFERENCES

1. Hanson MA, Gluckman PD. Early developmental conditioning of later health and disease: physiology or pathophysiology? Physiol Rev 2014; 94:1027-1076

2. De Kloet ER, Vreugdenhil E, Oitzl MS, Joels M. Brain corticosteroid receptor balance in health and disease. Endocr Rev 1998; 19:269-301

3. Slotkin TA, Seidler FJ, Wood CR, Lau C. Development of glucocorticoid receptor regulation in the rat forebrain: implications for adverse effects of glucocorticoids in preterm infants. Brain Res Bull 2008; 76:531-535

4. Kudielka BM, Kirschbaum C. Sex differences in HPA axis responses to stress: a review. Biol Psychol 2005; 69:113-132

5. Otte C, Hart S, Neylan TC, Marmar CR, Yaffe K, Mohr DC. A meta-analysis of cortisol response to challenge in human aging: importance of gender. Psychoneuroendocrinology 2005; 30:80-91

6. Lerner DJ, Kannel WB. Patterns of coronary heart disease morbidity and mortality in the sexes: a 26-year follow-up of the Framingham population. Am Heart J 1986; 111:383-390

7. Bourke CH, Harrell CS, Neigh GN. Stress-induced sex differences: adaptations mediated by the glucocorticoid receptor. Horm Behav 2012; 62:210-218

8. van der Voorn B, Hollanders JJ, Ket JC, Rotteveel J, Finken MJ. Gender-specific differences in hypothalamus-pituitary-adrenal axis activity during childhood: a systematic review and meta-analysis. Biol Sex Differ 2017; 8:3

9. Panagiotakopoulos L, Neigh GN. Development of the HPA axis: where and when do sex differences manifest? Front Neuroendocrinol 2014; 35:285-302

10. Jessop DS, Turner-Cobb JM. Measurement and meaning of salivary cortisol: a focus on health and disease in children. Stress 2008; 11:1-14

11. Haen EH, F.; Cornelissen, G. Cortisol marker rhythmometry in pediatrics and clinical pharmacology. Annual Review of Chronopharmacology 1984; 1:165-168

12. Garcia L, Hermida RC, Ayala DE, Lodeiro C, Iglesias T. Circadian characteristics of plasma cortisol in children with standard and short stature. Chronobiol Int 1990; 7:221-225

13. Knutsson U, Dahlgren J, Marcus C, Rosberg S, Bronnegard M, Stierna P, Albertsson-Wikland K. Circadian cortisol rhythms in healthy boys and girls: relationship with age, growth, body composition, and pubertal development. J Clin Endocrinol Metab 1997; 82:536-540

14. Jones A, Godfrey KM, Wood P, Osmond C, Goulden P, Phillips DI. Fetal growth and the adrenocortical response to psychological stress. J Clin Endocrinol Metab 2006; 91:1868-1871

15. Osika W, Friberg P, Wahrborg P. A new short self-rating questionnaire to assess stress in children. Int J Behav Med 2007; 14:108-117

16. Susman EJ, Dockray S, Schiefelbein VL, Herwehe S, Heaton JA, Dorn LD. Morningness/eveningness, morning-to-afternoon cortisol ratio, and antisocial behavior problems during puberty. Dev Psychol 2007; 43:811-822

17. Tzortzi C, Proff P, Redlich M, Aframian DJ, Palmon A, Golan I, Muessig D, Wichelhaus A, Baumert U. Cortisol daily rhythm in saliva of healthy school children. Int Dent J 2009; 59:12-18

18. Michels N, Sioen I, Huybrechts I, Bammann K, Vanaelst B, De Vriendt T, Iacoviello L, Konstabel K, Ahrens W, De Henauw S. Negative life events, emotions and psychological difficulties as determinants of salivary cortisol in Belgian primary school children. Psychoneuroendocrinology 2012; 37:1506-1515

19. Kjolhede EA, Gustafsson PE, Gustafsson PA, Nelson N. Overweight and obese children have lower cortisol levels than normal weight children. Acta Paediatr 2014; 103:295-299

20. Lumeng JC, Miller A, Peterson KE, Kaciroti N, Sturza J, Rosenblum K, Vazquez DM. Diurnal cortisol pattern, eating behaviors and overweight in low-income preschool-aged children. Appetite 2014; 73:65-72

21. Vanaelst B, Michels N, Clays E, Herrmann D, Huybrechts I, Sioen I, Vyncke K, De Henauw S. The association between childhood stress and body composition, and the role of stress-related lifestyle factors--cross-sectional findings from the baseline ChiBSD survey. Int J Behav Med 2014; 21:292-301

22. Barbosa TS, Castelo PM, Leme MS, Gaviao MB. Associations between oral health-related quality of life and emotional statuses in children and preadolescents. Oral Dis 2012; 18:639-647

23. Bartels M, de Geus EJ, Kirschbaum C, Sluyter F, Boomsma DI. Heritability of daytime cortisol levels in children. Behav Genet 2003; 33:421-433

24. Carrion VG, Weems CF, Ray RD, Glaser B, Hessl D, Reiss AL. Diurnal salivary cortisol in pediatric posttraumatic stress disorder. Biol Psychiatry 2002; 51:575-582

25. Doom JR, Cicchetti D, Rogosch FA, Dackis MN. Child maltreatment and gender interactions as predictors of differential neuroendocrine profiles. Psychoneuroendocrinology 2013; 38:1442-1454

26. Adam EK, Doane LD, Zinbarg RE, Mineka S, Craske MG, Griffith JW. Prospective prediction of major depressive disorder from cortisol awakening responses in adolescence. Psychoneuroendocrinology 2010; 35:921-931

27. Williams SR, Cash E, Daup M, Geronimi EM, Sephton SE, Woodruff-Borden J. Exploring patterns in cortisol synchrony among anxious and nonanxious mother and child dyads: a preliminary study. Biol Psychol 2013; 93:287-295

28. Morin-Major JK, Marin MF, Durand N, Wan N, Juster RP, Lupien SJ. Facebook behaviors associated with diurnal cortisol in adolescents: Is befriending stressful? Psychoneuroendocrinology 2016; 63:238-246

29. Martikainen S, Pesonen AK, Lahti J, Heinonen K, Feldt K, Pyhala R, Tammelin T, Kajantie E, Eriksson JG, Strandberg TE, Raikkonen K. Higher levels of physical activity are associated with lower hypothalamic-pituitary-adrenocortical axis reactivity to psychosocial stress in children. J Clin Endocrinol Metab 2013; 98:E619-627

30. Rosmalen JG, Oldehinkel AJ, Ormel J, de Winter AF, Buitelaar JK, Verhulst FC. Determinants of salivary cortisol levels in 10-12 year old children; a population-based study of individual differences. Psychoneuroendocrinology 2005; 30:483-495

31. Fransson E, Folkesson L, Bergstrom M, Ostberg V, Lindfors P. Exploring salivary cortisol and recurrent pain in mid-adolescents living in two homes. BMC Psychol 2014; 2:46

32. Kelly SJ, Young R, Sweeting H, Fischer JE, West P. Levels and confounders of morning cortisol collected from adolescents in a naturalistic (school) setting. Psychoneuroendocrinology 2008; 33:1257-1268

33. Ruttle PL, Javaras KN, Klein MH, Armstrong JM, Burk LR, Essex MJ. Concurrent and longitudinal associations between diurnal cortisol and body mass index across adolescence. J Adolesc Health 2013; 52:731-737

34. Shirtcliff EA, Allison AL, Armstrong JM, Slattery MJ, Kalin NH, Essex MJ. Longitudinal stability and developmental properties of salivary cortisol levels and circadian rhythms from childhood to adolescence. Dev Psychobiol 2012; 54:493-502

35. Vaillancourt T, Duku E, Decatanzaro D, Macmillan H, Muir C, Schmidt LA. Variation in hypothalamic-pituitary-adrenal axis activity among bullied and non-bullied children. Aggress Behav 2008; 34:294-305

36. Bae YJ, Stadelmann S, Klein AM, Jaeger S, Hiemisch A, Kiess W, Ceglarek U, Gaudl A, Schaab M, von Klitzing K, Thiery J, Kratzsch J, Dohnert M. The hyporeactivity of salivary cortisol at stress test (TSST-C) in children with internalizing or externalizing disorders is contrastively associated with alpha-amylase. J Psychiatr Res 2015; 71:78-88

37. Netherton C, Goodyer I, Tamplin A, Herbert J. Salivary cortisol and dehydroepiandrosterone in relation to puberty and gender. Psychoneuroendocrinology 2004; 29:125-140

38. Kuhlman KR, Geiss EG, Vargas I, Lopez-Duran NL. Differential associations between childhood trauma subtypes and adolescent HPA-axis functioning. Psychoneuroendocrinology 2015; 54:103-114

39. Matchock RL, Dorn LD, Susman EJ. Diurnal and seasonal cortisol, testosterone, and DHEA rhythms in boys and girls during puberty. Chronobiol Int 2007; 24:969-990

40. Bright MA, Frick JE, Out D, Granger DA. Individual differences in the cortisol and salivary alpha-amylase awakening responses in early childhood: relations to age, sex, and sleep. Dev Psychobiol 2014; 56:1300-1315

41. Bouma EM, Riese H, Ormel J, Verhulst FC, Oldehinkel AJ. Adolescents' cortisol responses to awakening and social stress; effects of gender, menstrual phase and oral contraceptives. The TRAILS study. Psychoneuroendocrinology 2009; 34:884-893

42. Dietrich A, Ormel J, Buitelaar JK, Verhulst FC, Hoekstra PJ, Hartman CA. Cortisol in the morning and dimensions of anxiety, depression, and aggression in children from a general population and clinic-referred cohort: An integrated analysis. The TRAILS study. Psychoneuroendocrinology 2013; 38:1281-1298

43. Hatzinger M, Brand S, Perren S, von Wyl A, von Klitzing K, Holsboer-Trachsler E. Hypothalamic-pituitary-adrenocortical (HPA) activity in kindergarten children: importance of gender and associations with behavioral/emotional difficulties. J Psychiatr Res 2007; 41:861-870

44. Pruessner JC, Wolf OT, Hellhammer DH, Buske-Kirschbaum A, von Auer K, Jobst S, Kaspers F, Kirschbaum C. Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. Life Sci 1997; 61:2539-2549

45. Evans BE, Greaves-Lord K, Euser AS, Tulen JH, Franken IH, Huizink AC. Determinants of physiological and perceived physiological stress reactivity in children and adolescents. PLoS One 2013; 8:e61724

46. Bouma EM, Riese H, Nolte IM, Oosterom E, Verhulst FC, Ormel J, Oldehinkel AJ. No associations between single nucleotide polymorphisms in corticoid receptor genes and heart rate and cortisol responses to a standardized social stress test in adolescents: the TRAILS study. Behav Genet 2011; 41:253-261

47. Kudielka BM, Buske-Kirschbaum A, Hellhammer DH, Kirschbaum C. HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: impact of age and gender. Psychoneuroendocrinology 2004; 29:83-98

48. Dockray S, Susman EJ, Dorn LD. Depression, cortisol reactivity, and obesity in childhood and adolescence. J Adolesc Health 2009; 45:344-350

49. Strahler J, Mueller A, Rosenloecher F, Kirschbaum C, Rohleder N. Salivary alpha-amylase stress reactivity across different age groups. Psychophysiology 2010; 47:587-595

50. Peckins MK, Dockray S, Eckenrode JL, Heaton J, Susman EJ. The longitudinal impact of exposure to violence on cortisol reactivity in adolescents. J Adolesc Health 2012; 51:366-372

51. Portnoy J, Raine A, Glenn AL, Chen FR, Choy O, Granger DA. Digit ratio (2D:4D) moderates the relationship between cortisol reactivity and self-reported externalizing behavior in young adolescent males. Biol Psychol 2015; 112:94-106

52. Hostinar CE, McQuillan MT, Mirous HJ, Grant KE, Adam EK. Cortisol responses to a group public speaking task for adolescents: variations by age, gender, and race. Psychoneuroendocrinology 2014; 50:155-166

53. Martin A, Hellhammer J, Hero T, Max H, Schult J, Terstegen L. Effective prevention of stress-induced sweating and axillary malodour formation in teenagers. Int J Cosmet Sci 2011; 33:90-97

54. Ji J, Negriff S, Kim H, Susman EJ. A study of cortisol reactivity and recovery among young adolescents: Heterogeneity and longitudinal stability and change. Dev Psychobiol 2016; 58:283-302

55. Raikkonen K, Matthews KA, Pesonen AK, Pyhala R, Paavonen EJ, Feldt K, Jones A, Phillips DI, Seckl JR, Heinonen K, Lahti J, Komsi N, Jarvenpaa AL, Eriksson JG, Strandberg TE, Kajantie E. Poor sleep and altered hypothalamic-pituitary-adrenocortical and sympatho-adrenal-medullary system activity in children. J Clin Endocrinol Metab 2010; 95:2254-2261

56. de Veld DM, Riksen-Walraven JM, de Weerth C. The relation between emotion regulation strategies and physiological stress responses in middle childhood. Psychoneuroendocrinology 2012; 37:1309-1319

57. Hostinar CE, Johnson AE, Gunnar MR. Parent support is less effective in buffering cortisol stress reactivity for adolescents compared to children. Dev Sci 2015; 18:281-297

58. Gunnar MR, Wewerka S, Frenn K, Long JD, Griggs C. Developmental changes in hypothalamus-pituitary-adrenal activity over the transition to adolescence: normative changes and associations with puberty. Dev Psychopathol 2009; 21:69-85

59. Mrug S, Tyson A, Turan B, Granger DA. Sleep problems predict cortisol reactivity to stress in urban adolescents. Physiol Behav 2016; 155:95-101

60. Lu Q, Tao F, Hou F, Zhang Z, Sun Y, Xu Y, Xu S, Zhao Y. Cortisol reactivity, delay discounting and percent body fat in Chinese urban young adolescents. Appetite 2014; 72:13-20

61. Trickett PK, Gordis E, Peckins MK, Susman EJ. Stress reactivity in maltreated and comparison male and female young adolescents. Child Maltreat 2014; 19:27-37

62. Dorn LD, Burgess ES, Susman EJ, von Eye A, DeBellis MD, Gold PW, Chrousos GP. Response to oCRH in depressed and nondepressed adolescents: does gender make a difference? J Am Acad Child Adolesc Psychiatry 1996; 35:764-773

63. Forest MG. Age-related response of plasma testosterone, delta 4-androstenedione, and cortisol to adrenocorticotropin in infants, children, and adults. J Clin Endocrinol Metab 1978; 47:931-937

64. Lashansky G, Saenger P, Fishman K, Gautier T, Mayes D, Berg G, Di Martino-Nardi J, Reiter E. Normative data for adrenal steroidogenesis in a healthy pediatric population: age- and sex-related changes after adrenocorticotropin stimulation. J Clin Endocrinol Metab 1991; 73:674-686

65. Ross JL, Schulte HM, Gallucci WT, Cutler GB, Jr., Loriaux DL, Chrousos GP. Ovine corticotropin-releasing hormone stimulation test in normal children. J Clin Endocrinol Metab 1986; 62:390-392

66. Tsvetkova V. Adrenocortical function after stimulation with synthetic ACTH. Curr Med Res Opin 1977; 4:635-639

67. Stroud LR, Papandonatos GD, Williamson DE, Dahl RE. Sex differences in cortisol response to corticotropin releasing hormone challenge over puberty: Pittsburgh Pediatric Neurobehavioral Studies. Psychoneuroendocrinology 2011; 36:1226-1238

68. Dahl RE, Siegel SF, Williamson DE, Lee PA, Perel J, Birmaher B, Ryan ND. Corticotropin releasing hormone stimulation test and nocturnal cortisol levels in normal children. Pediatr Res 1992; 32:64-68

69. Davis M, Emory E. Sex differences in neonatal stress reactivity. Child Dev 1995; 66:14-27

70. Eiden RD, Molnar DS, Granger DA, Colder CR, Schuetze P, Huestis MA. Prenatal tobacco exposure and infant stress reactivity: role of child sex and maternal behavior. Dev Psychobiol 2015; 57:212-225

71. Grunau RE, Tu MT, Whitfield MF, Oberlander TF, Weinberg J, Yu W, Thiessen P, Gosse G, Scheifele D. Cortisol, behavior, and heart rate reactivity to immunization pain at 4 months corrected age in infants born very preterm. Clin J Pain 2010; 26:698-704

72. Gunnar MR, Kryzer E, Van Ryzin MJ, Phillips DA. The rise in cortisol in family day care: associations with aspects of care quality, child behavior, and child sex. Child Dev 2010; 81:851-869

73. Plusquellec P, Ouellet-Morin I, Feng B, Perusse D, Tremblay RE, Lupien SJ, Boivin M. Salivary cortisol levels are associated with resource control in a competitive situation in 19 month-old boys. Horm Behav 2011; 60:159-164

74. Spinrad TL, Eisenberg N, Granger DA, Eggum ND, Sallquist J, Haugen RG, Kupfer A, Hofer C. Individual differences in preschoolers' salivary cortisol and alpha-amylase reactivity: relations to temperament and maladjustment. Horm Behav 2009; 56:133-139

75. de Weerth C, Zijlmans MA, Mack S, Beijers R. Cortisol reactions to a social evaluative paradigm in 5- and 6-year-old children. Stress 2013; 16:65-72

76. Kryski KR, Smith HJ, Sheikh HI, Singh SM, Hayden EP. HPA axis reactivity in early childhood: associations with symptoms and moderation by sex. Psychoneuroendocrinology 2013; 38:2327-2336

77. Yong Ping E, Laplante DP, Elgbeili G, Hillerer KM, Brunet A, O'Hara MW, King S. Prenatal maternal stress predicts stress reactivity at 2(1/2) years of age: the Iowa Flood Study. Psychoneuroendocrinology 2015; 56:62-78

78. Mills RS, Imm GP, Walling BR, Weiler HA. Cortisol reactivity and regulation associated with shame responding in early childhood. Dev Psychol 2008; 44:1369-1380

79. Hackman DA, Betancourt LM, Brodsky NL, Hurt H, Farah MJ. Neighborhood disadvantage and adolescent stress reactivity. Front Hum Neurosci 2012; 6:277

80. Zijlmans MA, Beijers R, Mack S, Pruessner JC, de Weerth C. Cortisol responses to social evaluation in 10- to 15-year-old boys and girls. Stress 2013; 16:393-401

81. Daughters SB, Gorka SM, Matusiewicz A, Anderson K. Gender specific effect of psychological stress and cortisol reactivity on adolescent risk taking. J Abnorm Child Psychol 2013; 41:749-758

82. Minkley N, Kirchner WH. Influence of test tasks with different cognitive demands on salivary cortisol concentrations in school students. Int J Psychophysiol 2012; 86:245-250

83. Allen LB, Lu Q, Tsao JC, Worthman CM, Zeltzer LK. Sex differences in the association between cortisol concentrations and laboratory pain responses in healthy children. Gend Med 2009; 6 Suppl 2:193-207

84. Covelli MM, Wood CE, Yarandi HN. Biologic measures as epidemiological indicators of risk for the development of hypertension in an African American adolescent population. J Cardiovasc Nurs 2012; 27:476-484

85. Gecgelen M, Aksoy A, Kirdemir P, Doguc DK, Cesur G, Koskan O, Ozorak O. Evaluation of stress and pain during rapid maxillary expansion treatments. J Oral Rehabil 2012; 39:767-775

86. Khilnani P, Munoz R, Salem M, Gelb C, Todres ID, Chernow B. Hormonal responses to surgical stress in children. J Pediatr Surg 1993; 28:1-4

87. Yfanti K, Kitraki E, Emmanouil D, Pandis N, Papagiannoulis L. Psychometric and biohormonal indices of dental anxiety in children. A prospective cohort study. Stress 2014; 17:296-304

88. Lopez-Duran NL, McGinnis E, Kuhlman K, Geiss E, Vargas I, Mayer S. HPA-axis stress reactivity in youth depression: evidence of impaired regulatory processes in depressed boys. Stress 2015; 18:545-553

89. Chiodo S, Tessitore A, Cortis C, Cibelli G, Lupo C, Ammendolia A, De Rosas M, Capranica L. Stress-related hormonal and psychological changes to official youth Taekwondo competitions. Scand J Med Sci Sports 2011; 21:111-119

90. Stupnicki R, Obminski Z, Klusiewicz A, Viru A. Pre-exercise serum cortisol concentration and responses to laboratory exercise. Eur J Appl Physiol Occup Physiol 1995; 71:439-443

91. Frias J, Rodriguez R, Torres JM, Ruiz E, Ortega E. Effects of acute alcohol intoxication on pituitary-gonadal axis hormones, pituitary-adrenal axis hormones, beta-endorphin and prolactin in human adolescents of both sexes. Life Sci 2000; 67:1081-1086

92. Kudielka BM, Hellhammer J, Hellhammer DH, Wolf OT, Pirke KM, Varadi E, Pilz J, Kirschbaum C. Sex differences in endocrine and psychological responses to psychosocial stress in healthy elderly subjects and the impact of a 2-week dehydroepiandrosterone treatment. J Clin Endocrinol Metab 1998; 83:1756-1761

93. Lekkakou L, Tzanela M, Lymberi M, Consoulas C, Tsagarakis S, Koutsilieris M. Effects of gender and age on hypothalamic-pituitary-adrenal reactivity after pharmacological challenge with low-dose 1-mug ACTH test: a prospective study in healthy adults. Clin Endocrinol (Oxf) 2013; 79:683-688

94. Seeman TE, Singer B, Charpentier P. Gender differences in patterns of HPA axis response to challenge: Macarthur studies of successful aging. Psychoneuroendocrinology 1995; 20:711-725

95. Traustadottir T, Bosch PR, Matt KS. Gender differences in cardiovascular and hypothalamic-pituitary-adrenal axis responses to psychological stress in healthy older adult men and women. Stress 2003; 6:133-140

96. Rosmond R, Bjorntorp P. The hypothalamic-pituitary-adrenal axis activity as a predictor of cardiovascular disease, type 2 diabetes and stroke. J Intern Med 2000; 247:188-197

97. Kajantie E, Hovi P. Is very preterm birth a risk factor for adult cardiometabolic disease? Semin Fetal Neonatal Med 2014; 19:112-117

98. Lund TD, Rovis T, Chung WC, Handa RJ. Novel actions of estrogen receptor-beta on anxiety-related behaviors. Endocrinology 2005; 146:797-807

99. Gillies GE, McArthur S. Estrogen actions in the brain and the basis for differential action in men and women: a case for sex-specific medicines. Pharmacol Rev 2010; 62:155-198

100. Yokosuka M, Okamura H, Hayashi S. Postnatal development and sex difference in neurons containing estrogen receptor-alpha immunoreactivity in the preoptic brain, the diencephalon, and the amygdala in the rat. J Comp Neurol 1997; 389:81-93

101. Moore DE, Kawagoe S, Davajan V, Mishell DR, Nakamura RM. An in vivo system in man for quantitation of estrogenicity. I. Physiologic changes in binding capacity of serum corticosteroid-binding globulin. Am J Obstet Gynecol 1978; 130:475-481

102. Finken MJ, Andrews RC, Andrew R, Walker BR. Cortisol metabolism in healthy young adults: sexual dimorphism in activities of A-ring reductases, but not 11beta-hydroxysteroid dehydrogenases. J Clin Endocrinol Metab 1999; 84:3316-3321

103. Kajantie E, Phillips DI. The effects of sex and hormonal status on the physiological response to acute psychosocial stress. Psychoneuroendocrinology 2006; 31:151-178

104. Kirschbaum C, Schommer N, Federenko I, Gaab J, Neumann O, Oellers M, Rohleder N, Untiedt A, Hanker J, Pirke KM, Hellhammer DH. Short-term estradiol treatment enhances pituitary-adrenal axis and sympathetic responses to psychosocial stress in healthy young men. J Clin Endocrinol Metab 1996; 81:3639-3643

105. Komesaroff PA, Esler MD, Sudhir K. Estrogen supplementation attenuates glucocorticoid and catecholamine responses to mental stress in perimenopausal women. J Clin Endocrinol Metab 1999; 84:606-610

106. Lindheim SR, Legro RS, Bernstein L, Stanczyk FZ, Vijod MA, Presser SC, Lobo RA. Behavioral stress responses in premenopausal and postmenopausal women and the effects of estrogen. Am J Obstet Gynecol 1992; 167:1831-1836

107. Clow A, Hucklebridge F, Stalder T, Evans P, Thorn L. The cortisol awakening response: more than a measure of HPA axis function. Neurosci Biobehav Rev 2010; 35:97-103

108. Stroud LR, Salovey P, Epel ES. Sex differences in stress responses: social rejection versus achievement stress. Biol Psychiatry 2002; 52:318-327

109. Taylor SE, Klein LC, Lewis BP, Gruenewald TL, Gurung RA, Updegraff JA. Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. Psychol Rev 2000; 107:411-429

110. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, Sousa AM, Pletikos M, Meyer KA, Sedmak G, Guennel T, Shin Y, Johnson MB, Krsnik Z, Mayer S, Fertuzinhos S, Umlauf S, Lisgo SN, Vortmeyer A, Weinberger DR, Mane S, Hyde TM, Huttner A, Reimers M, Kleinman JE, Sestan N. Spatio-temporal transcriptome of the human brain. Nature 2011; 478:483-489

111. Seeman TE, Robbins RJ. Aging and hypothalamic-pituitary-adrenal response to challenge in humans. Endocr Rev 1994; 15:233-260

112. Rasmussen AR, Wohlfahrt-Veje C, Tefre de Renzy-Martin K, Hagen CP, Tinggaard J, Mouritsen A, Mieritz MG, Main KM. Validity of self-assessment of pubertal maturation. Pediatrics 2015; 135:86-93

113. Juster RP, Raymond C, Desrochers AB, Bourdon O, Durand N, Wan N, Pruessner JC, Lupien SJ. Sex hormones adjust "sex-specific" reactive and diurnal cortisol profiles. Psychoneuroendocrinology 2016; 63:282-290

APPENDIX 1

Appendix 1A: Search strategy for PubMed (14 January 2016)

Search	Query	Records (*n*)*
#1	"Hydrocortisone"[Mesh] OR "Glucocorticoids"[Mesh] OR "11-beta-Hydroxysteroid Dehydrogenases"[Mesh] OR "Tetrahydrocortisone"[Mesh] OR "Tetrahydrocortisol"[Mesh] OR cortisol*[tiab] OR hydrocortison*[tiab] OR epicortisol*[tiab] OR cortifair*[tiab] OR cortril*[tiab] OR glucocorticoid*[tiab] OR beta hydroxysteroid dehydrogenase*[tiab] OR 11 oxoreductase*[tiab] OR 11 oxidoreductase*[tiab] OR 11 hydroxysteroid dehydrogenase*[tiab] OR 11b hydroxysteroid dehydrogenase*[tiab] OR 11 reductase*[tiab] OR 11beta hydroxysteroid dehydrogenase*[tiab] OR tetrahydrocortiso*[tiab] OR "hsd11b2"[tiab] OR "11bhsd2"[tiab] OR "11betahsd2"[tiab] OR 11beta hsd*[tiab] OR hydroxycortisol*[tiab] OR "Circadian Rhythm"[Mesh] OR "twenty four hour"[tiab] OR circadian*[tiab] OR diurnal*[tiab] OR nyctohemeral*[tiab]	250,920
#2	child*[tw] OR schoolchild*[tw] OR infan*[tw] OR adolescen*[tw] OR pediatri*[tw] OR paediatr*[tw] OR neonat*[tw] OR boy[tw] OR boys[tw] OR boyhood[tw] OR girl[tw] OR girls[tw] OR girlhood[tw] OR youth[tw] OR youths[tw] OR baby[tw] OR babies[tw] OR toddler*[tw] OR "Mental Disorders Diagnosed in Childhood"[MeSH] OR teen[tw] OR teens[tw] OR teenager*[tw] OR newborn*[tw] OR postneonat*[tw] OR postnat*[tw] OR perinat*[tw] OR puberty[tw] OR preschool*[tw] OR suckling*[tw] OR picu[tw] OR nicu[tw] OR "Arthritis, Juvenile"[Mesh] OR "Myoclonic Epilepsy, Juvenile"[Mesh] OR "Leukemia, Myelomonocytic, Juvenile"[Mesh] OR "Xanthogranuloma, Juvenile"[Mesh] OR "Juvenile Delinquency"[Mesh] OR "Corneal Dystrophy, Juvenile Epithelial of Meesmann"[Mesh]	3,664,351
#3	"Sex Characteristics"[Mesh] OR "Sex Factors"[Mesh] OR sex characteristic*[tiab] OR sex difference*[tiab] OR sex dimorphism*[tiab] OR sexual dimorphism*[tiab] OR sexual difference*[tiab] OR sexual characteristic*[tiab] OR sex factor*[tiab] OR sexual factor*[tiab] OR sexual dimorphi*[tiab] OR sex influenc*[tiab] OR sexual influenc*[tiab] OR gender*[tiab]	451,894
#4	(#1 AND #2 AND #3)	2,643

Mesh = Medical subject headings; tiab = words in title OR abstract; tw = words in title, abstract, MeSH and other content related fields

Appendix 1B: Search strategy for Embase.com (14 January 2016)

Search	Query	Records (*n*)*
#1	'hydrocortisone'/exp OR 'glucocorticoid'/de OR '11beta hydroxysteroid dehydrogenase'/exp OR 'tetrahydrocortisone'/exp OR 'tetrahydrocortisol'/exp OR cortisol*:ab,ti OR hydrocortison*:ab,ti OR epicortisol*:ab,ti OR cortifair*:ab,ti OR cortril*:ab,ti OR glucocorticoid*:ab,ti OR ('beta hydroxysteroid' NEAR/3 dehydrogenase*):ab,ti OR (11 NEXT/1 oxoreductase*):ab,ti OR (11 NEXT/1 oxidoreductase*):ab,ti OR ('11 hydroxysteroid' NEXT/1 dehydrogenase*):ab,ti OR ('11b hydroxysteroid' NEXT/1 dehydrogenase*):ab,ti OR (11 NEXT/1 reductase*):ab,ti OR ('11beta hydroxysteroid' NEXT/1 dehydrogenase*):ab,ti OR tetrahydrocortiso*:ab,ti OR 'hsd11b2':ab,ti OR '11bhsd2':ab,ti OR '11betahsd2':ab,ti OR (11beta NEXT/1 hsd*):ab,ti OR hydroxycortisol*:ab,ti OR ('trier social stress' NEXT/1 test*):ab,ti OR tsst:ab,ti OR (stress NEAR/3 hormone*):ab,ti OR (stress NEAR/3 marker*):ab,ti	**235,221**
#2	adolescen*:ab,ti OR 'adolescence'/exp OR 'adolescent coping orientation for problem experiences'/exp OR 'adolescent development'/exp OR 'adolescent disease'/exp OR 'adolescent health'/exp OR 'adolescent parent'/exp OR 'adolescent pregnancy'/exp OR 'adolescent smoking'/exp OR 'adolescent'/exp OR 'adolescent-family inventory of life events and changes'/exp OR babies:ab,ti OR baby:ab,ti OR 'birth weight'/exp OR boy:ab,ti OR boyhood:ab,ti OR boys:ab,ti OR 'brazelton neonatal behavioral assessment scale'/exp OR 'child abuse'/exp OR 'child advocacy'/exp OR 'child behavior checklist'/exp OR 'child behavior'/exp OR 'child care'/exp OR 'child death'/exp OR 'child health care'/exp OR 'child health'/exp OR 'child nutrition'/exp OR 'child parent relation'/exp OR 'child psychology'/exp OR 'child restraint system'/exp OR 'child safety'/exp OR 'child welfare'/exp OR child*:ab,ti OR 'child'/exp OR 'childhood disease'/exp OR 'childhood mortality'/exp OR 'childhood'/exp OR girl:ab,ti OR girlhood:ab,ti OR girls:ab,ti OR 'high risk infant'/exp OR infan*:ab,ti OR 'infant disease'/exp OR 'infant mortality'/exp OR 'infant nutrition'/exp OR 'infant welfare'/exp OR 'infanticide'/exp OR 'infantile diarrhea'/exp OR 'infantile hypotonia'/exp OR 'juvenile delinquency'/exp OR neonat*:ab,ti OR 'neonatal weight loss'/exp OR 'newborn disease'/exp OR 'newborn morbidity'/exp OR 'newborn period'/exp OR newborn*:ab,ti OR 'newborn'/exp OR nicu:ab,ti OR 'only child'/exp OR paediatr*:ab,ti OR pediatri*:de,ab,ti OR 'pediatric advanced life support'/exp OR 'pediatric anesthesia'/exp OR 'pediatric cardiology'/exp OR 'pediatric hospital'/exp OR 'pediatric intensive care nursing'/exp OR 'pediatric nurse practitioner'/exp OR 'pediatric nursing'/exp OR 'pediatric rehabilitation'/exp OR 'pediatric surgery'/exp OR 'newborn hypoxia'/exp OR 'pediatric ward'/exp OR 'pediatrics'/exp OR perinat*:ab,ti OR 'perinatal development'/exp OR 'perinatal period'/exp OR 'persistent hyperinsulinemic hypoglycemia of infancy'/exp OR picu:ab,ti OR postnat*:ab,ti OR 'postnatal care'/exp OR 'postnatal development'/exp OR 'postnatal growth'/exp OR postneonat*:ab,ti OR preschool*:ab,ti OR puberty:ab,ti OR 'runaway behavior'/exp OR 'school child':ab,ti OR schoolchild*:ab,ti OR 'severe myoclonic epilepsy in infancy'/exp OR suckling*:ab,ti OR teen:ab,ti OR teenager*:ab,ti OR teens:ab,ti OR toddler*:ab,ti OR 'transient hypogammaglobulinemia of infancy'/exp OR youth:ab,ti OR youths:ab,ti	**4,477,134**
#3	'sex difference'/exp OR 'sex ratio'/exp OR ('boy'/exp AND 'girl'/exp) OR (sex NEAR/3 characteristic*):ab,ti OR (sex NEAR/3 difference*):ab,ti OR (sex NEAR/3 dimorphism*):ab,ti OR (sexual NEAR/3 dimorphism*):ab,ti OR (sexual NEAR/3 difference*):ab,ti OR (sexual NEAR/3 characteristic*):ab,ti OR (sex NEAR/3 factor*):ab,ti OR (sexual NEAR/3 factor*):ab,ti OR (sexual NEAR/3 dimorphi*):ab,ti OR (sex NEAR/3 influenc*):ab,ti OR (sexual NEAR/3 influenc*):ab,ti OR gender*:ab,ti OR (boy*:ab,ti AND girl*:ab,ti) OR sex:ab,ti	**1,014,014**
#4	(#1 AND #2 AND #3)	**4,280**

/exp = EMtree keyword with explosion; /de = EMtree keyword without explosion; :ab,ti = words in title or abstract; NEXT/x = words in that order next to each other, x places apart; NEAR/x = words near to each other, x places apart
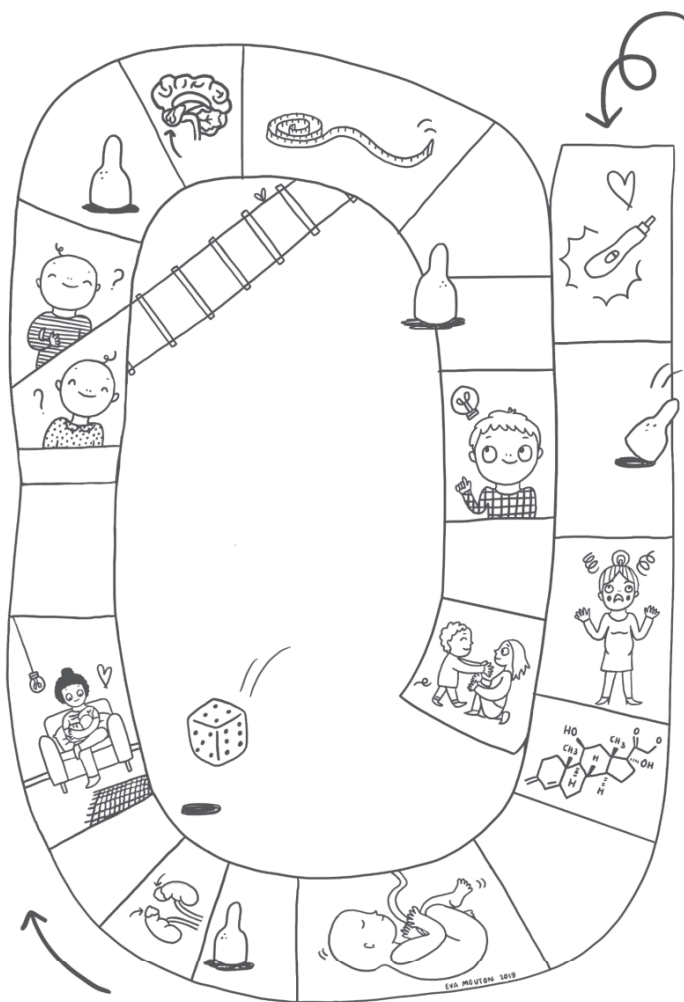
ONLINE SUPPLEMENTARY FILE

1. Extracted data of the studies included in the systematic review (https://doi.org/10.6084/m9.figshare.11014775.v1)

Part 3

Early-life thyroid regulation in preterm infants

EVA MOUTON 2019

11
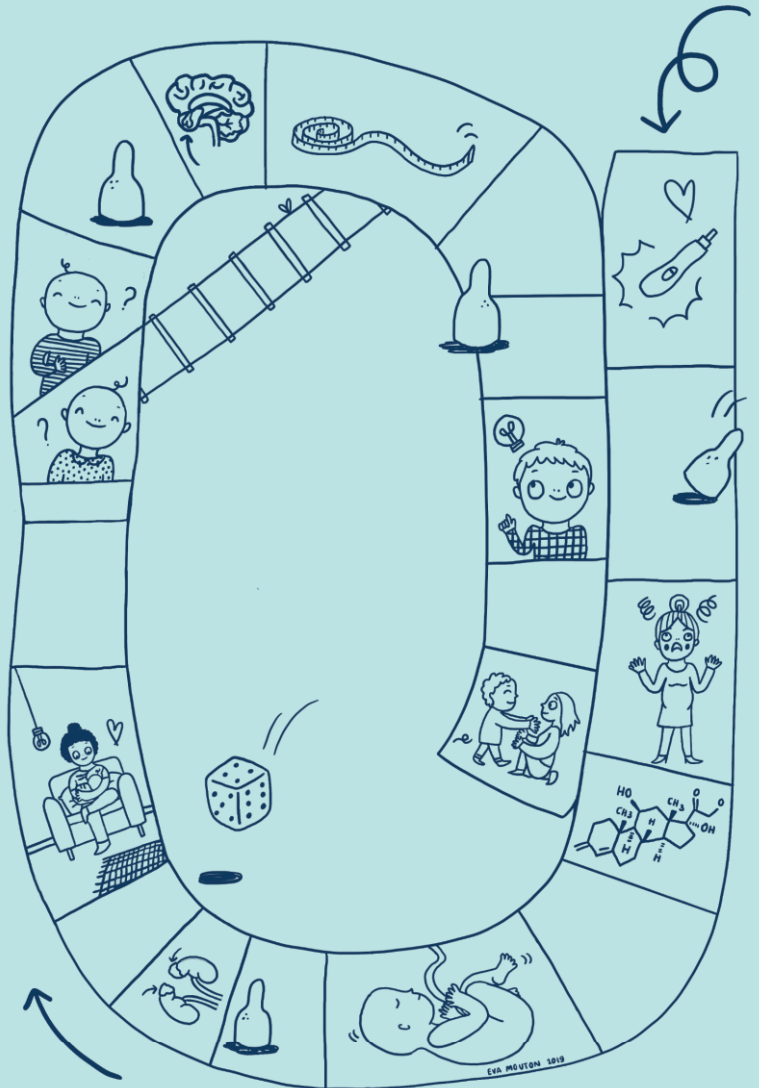
No association between transient hypothyroxinemia of prematurity and neurodevelopmental outcome in young adulthood

Jonneke J. Hollanders,
Joël Israëls,
Sylvia M. van der Pal,
Paul H. Verkerk,
Joost Rotteveel,
Martijn J. J. Finken

ABSTRACT

Context

Transient hypothyroxinemia of prematurity (THoP) has been associated with neuro-developmental impairment in infancy and childhood. It is not known whether these relations persist into adulthood.

Objective

The objective was to examine whether there is an effect of THoP on intelligence quotient (IQ) score and motor functioning at a young adult age.

Design

This study was part of the 19-year follow-up of the Project On Preterm and Small-for-gestational-age birth (POPS) cohort, which included infants born very preterm (ie, <32 wk) and/or with a very low birth weight (ie, <1500 g).

Setting

This was a multicenter study.

Patients

There were 398 19-year-old participants of the POPS cohort, of whom 120 had THoP.

Exposure

T4 concentrations were obtained through the national neonatal screening program for congenital hypothyroidism. THoP was defined as a total T4 concentration < −3 SD of the daily mean (approximately 60 nmol/L).

Main Outcome Measures

Main outcome measures were IQ and motor functioning, measured with the digital Multicultural Capacities Test-Intermediate Level and a revised version of Touwen's examination of minor neurological dysfunction, respectively.

Results

THoP was not associated with IQ score (mean difference, 0 [95% confidence interval, −3.8 to 3.8] points) or motor function (mean difference, 0.6 [95% confidence interval, −1.3 to 2.5] points) after adjustment for demographic and perinatal characteristics.

Conclusions

No associations between THoP and neurodevelopmental outcome at age 19 years were found.

INTRODUCTION

Preterm infants often develop transient hypothyroxinemia of prematurity (THoP). This is characterized by a temporary reduction in thyroxine (T4) that may last for 6-8 weeks,[1,2] whereas TSH remains low to normal.[3] After the severance of the umbilical cord, the transplacental supply of maternal T4 stops immediately.[4] Other mechanisms contributing to THoP are hypothalamus-pituitary-thyroid axis immaturity, reduced thyroidal iodine reserves, and acute illnesses.[2,5-8] Continuing debate exists about whether THoP is harmful for the developing brain.

THoP has been associated with delayed nerve conduction velocity,[9] later achievement of developmental milestones,[10] lower scores in cognitive tests,[11] and increased risks of school failure[12] and cerebral palsy.[13] However, there are currently no evidence-based guidelines on the screening for THoP. Moreover, the few trials that have addressed neurodevelopmental outcome after levothyroxine supplementation in preterm newborns were negative.[14-16] Nevertheless, a *post hoc* analysis suggested that the effects of this therapy were dependent on the degree of prematurity.[15] More specifically, it was found that infants of 25–26 weeks' gestation had a higher score on the Mental Developmental Index of the Bayley Scales of Infant Development at 2 years of age if treated with levothyroxine when compared with untreated controls. By contrast, treated infants of 27–29 weeks' gestation scored, on average, 10 points lower than untreated infants. Similar gestational age-dependent effects were observed for motor functioning, [15] and all of the associations found at 2 years were reported to persist at ages 5.7 and 10 years.[17,18]

It is not known whether the neurodevelopmental effects of THoP persist into adulthood. We therefore studied the effects of a low T4 concentration, obtained during a T4-based neonatal screening program for congenital hypothyroidism, on intelligence quotient (IQ) and neuromotor function at 19 years of age in a large cohort of very preterm (<32 weeks' gestation) and/or very low birth weight (<1,500 g) infants in The Netherlands. Based on previous observations in this cohort,[12] we expected to find worse neurodevelopmental outcomes after THoP.

METHODS

Population

The Project On Preterm and Small-for-gestational-age infants (POPS) cohort comprised 94% of the infants born alive in The Netherlands in 1983 with a gestational age <32 weeks and/or a birth weight <1,500 g.[19] The original cohort consisted of 1,338 infants, of whom 959 (72%) survived to age 19 years.

From April 1983 onward, neonatal screening results for congenital hypothyroidism were prospectively collected.[12] In addition, the screening results of 54 subjects born before April 1983 could also be retrieved. T4 was therefore known for 745 of the surviving subjects (78%). In line with previous analyses in the POPS cohort, we excluded the data of subjects whose T4 concentrations were measured before postnatal day 5 or after day 17 (n=66).[10,12] Subjects were also excluded if they received thyroid hormone supplementation during their stay in the hospital (n=5), as were subjects with severe congenital malformations, such as Down's syndrome, central nervous system defects or inborn errors of metabolism (n=10), severe sensory handicap (n=8), and congenital hypothyroidism (n=1). This left 655 eligible subjects for our study, of whom 398 (61%) underwent a neurological examination and/or IQ testing at one of the 10 participating centers. For the analyses of neuromotor function, we also excluded subjects taking drugs with a high risk of extrapyramidal side effects (n=2). The flowchart of the study sample is shown in Figure 1.

The study was approved by the medical ethics committees of the participating centers, and written informed consent was obtained from all participants.

Laboratory investigations

T4 concentrations from filter paper eluates were determined in duplicate by radioimmunoassay.[20] Five accredited laboratories processed an average of 125 samples per day; they were all under permanent quality control.[21] Samples were not analyzed continuously. T4 levels in the eluates were expressed as standard deviations (SD) from the mean, which was calculated on a daily basis.[22] The intra-assay and interassay coefficients of variation in the eluates were 8 and 10%, respectively. Sampling time of day was not taken into account. Consistent with previous analyses in this cohort, hypothyroxinemia was defined as a T4 concentration of >3 SD below the daily reference mean (approximately 60 nmol/L).[12] TSH was measured only in infants with the lowest 20% T4 values. These values were not used because they do not aid in the identification of infants with THoP.[14]

Intelligence quotient

Intellectual functioning was assessed with the use of the computerized version of the Multicultural Capacity Test (MCT)–Intermediate Level.[23] In all, this test provides an overview of a person's capacities and skills: i.e., verbal and numerical intelligence, spatial visualization, speech fluency, memory, reasoning, and speed of perception. The MCT is validated for individuals aged ≥16 years from different backgrounds, whose level of education ranges from 5 years of secondary school to university level. In the Dutch norm population, the MCT reports an IQ score of 100±15.
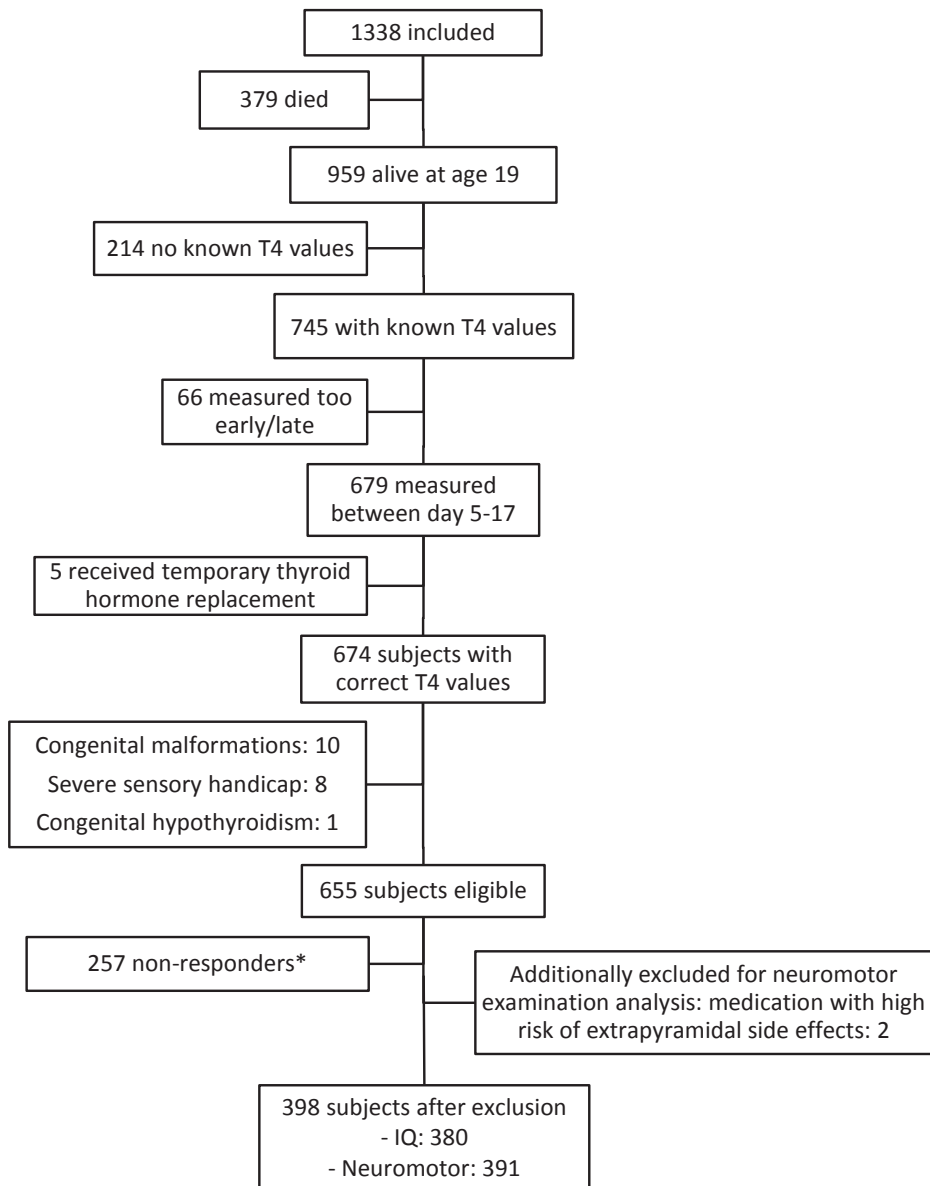
Figure 1: Flowchart of the inclusion of POPS subjects at age 19 years.
* Included in this number are subjects who only returned the (parental) questionnaires

Neuromotor performance

Neuromotor function assessment was based on a revised version of Touwen's examination of minor neurological dysfunction.[24,25] This examination focuses on five subcategories of function: hand function, quality of walking, coordination, posture, and passive muscle tone. The test comprises 34 items, each of which is scored on a 3-point scale. Two points are assigned for optimal performance, 1 point for slightly reduced performance, and 0 points for poor performance. The maximum total score is 68.

Statistical analysis

All outcomes showed fairly normal distributions. Linear regression analysis was used to study the effects of hypothyroxinemia and of T4 standard deviation score (SDS) across the entire range on continuous outcomes. Logistic regression analysis was used to study the effects of neonatal thyroid function parameters on the odds of having an IQ <85 points. Analyses were repeated after adjustment for the demographic characteristics gender, socioeconomic status (SES), ethnicity, and parity. Next, perinatal characteristics, including gestational age, being born small-for-gestational-age (SGA), and neonatal illnesses like infant respiratory distress syndrome, intraventricular hemorrhage, and sepsis, were added as covariates to the model. A *P* value of ≤0.05 was considered statistically significant.

Analyses were repeated after including only the subjects who participated in the POPS follow-up at age 5 (n=377 for IQ, n=387 for neuromotor functioning).[12] Additionally, analyses were repeated after stratification of gestational age into <and ≥ 29 weeks. This cut-off point was based on studies showing gestational age-dependent effects of levothyroxine treatment in preterm newborns.[26]

Our sample size enabled us to detect a difference of 5.0 IQ points, assuming an SD value of 15 points,[23] and a difference of 3.1 points on the neuromotor examination, assuming an SD of 9.5 points,[27] with a power of 80% and a significance level of 0.05.

RESULTS

Table 1 shows the general and perinatal characteristics of responders and nonresponders. Nonresponse was associated with male gender, non-Caucasian ethnicity, lower SES, and younger maternal age at birth. It was unrelated to the perinatal characteristics, the T4 concentration, or the proportion that exibited THoP. Hypothyroxinemia was associated with a lower gestational age and birth weight and the presence of neonatal morbidities. However, hypothyroxinemic infants were less often born SGA.

Table 1: General and perinatal characteristics of hypothyroxinemic vs. non-hypothyroxinemic groups, and of responders vs. non-responders

	Hypothyroxinemic (n=104)	Non-hypothyroxinemic (n=294)	P value*	Non-responders (n=257)	P value†
General					
Male sex (%)	50 (48.1)	129 (43.9)	0.46	154 (59.9)	<0.001
White (%)	87 (86.1)	265 (90.4)	0.23	208 (80.9)	0.002
Low socio-economic status (%)	37 (35.9)	105 (36.0)	0.99	134 (54.3)	<0.001
First child (%)	55 (53.4)	167 (56.8)	0.55	136 (53.1)	0.48
Perinatal					
Maternal age (yrs)	26.3±4.8	27.4±6.1	0.102	26.6±5.1	0.22
Gestational age (wks)	29.4±2.2	31.3±2.4	<0.001	31.1±2.6	0.23
Birth weight (g)	1,170±241	1,297±261	<0.001	1,297±250	0.11
SGA birth (%)	27 (26.2)	127 (43.2)	0.002	100 (38.9)	0.98
Apgar score ≥7 after 5 min (%)	81 (77.9)	254 (86.4)	0.09	213 (82.9)	0.55
Part of multiple pregnancy (%)	31 (29.8)	66 (22.4)	0.13	51 (19.8)	0.18
Respiratory distress syndrome (%)	60 (57.7)	101 (34.4)	<0.001	101 (39.3)	0.77
Intraventricular hemorrhage (%)	32 (30.8)	34 (11.6)	<0.001	51 (19.8)	0.29
Sepsis (%)	45 (43.3)	80 (27.3)	0.003	86 (33.5)	0.60
Necrotizing enterocolitis (%)	7 (6.7)	17 (5.8)	0.73	10 (3.9)	0.23
Total T4 (SD)	-3.5±0.3	-2.0±0.7	<0.001	-2.4±1.0	0.75
THoP (%)	104 (100%)	0 (0%)	<0.001	72 (28.0%)	0.60

Values represent mean±SD or n (%). Continuous variables were compared with the unpaired t test. Dichotomous variables were compared with the Chi square test.

* *P* value between hypothyroxinemic and non-hypothyroxinemic groups

† *P* value between responders and non-responders

Table 2: Associations between neonatal thyroid function parameters and neurodevelopmental outcomes at age 19 years

	Unadjusted	P value	Adjusted (1)	P value	Adjusted (2)	P value
T4 SDS						
IQ						
IQ total score	-1.4 (-3.1 to 0.2)	0.084	-1.3 (-2.9 to 0.4)	0.129	-1.4 (-3.3 to 0.4)	0.131
Linguistic capacity z score	-0.01 (-0.10 to 0.08)	0.806	-0.01 (-0.10 to 0.08)	0.819	-0.04 (-0.14 to 0.05)	0.389
Mathematical capacity z score	-0.09 (-0.20 to 0.01)	0.088	-0.08 (-0.19 to 0.02)	0.120	-0.11 (-0.23 to 0.01)	0.074
Logical reasoning z score	-0.08 (-0.18 to 0.02)	0.098	-0.09 (-0.18 to 0.01)	0.088	-0.11 (-0.22 to 0.00)	0.054
Spatial visualization z score	-0.03 (-0.13 to 0.06)	0.501	-0.02 (-0.11 to 0.08)	0.690	-0.03 (-0.14 to 0.07)	0.547
Neuromotor function						
Neuromotor sum score	0.3 (-0.6 to 1.1)	0.524	0.2 (-0.6 to 1.1)	0.603	-0.1 (-1.0 to 0.9)	0.904
Hand function	0.08 (-0.03 to 0.18)	0.137	0.07 (-0.03 to 0.17)	0.164	0.06 (-0.05 to 0.18)	0.280
Quality of walking	-0.02 (-0.14 to 0.10)	0.736	-0.03 (-0.15 to 0.10)	0.685	-0.08 (-0.22 to 0.06)	0.276
Coordination	0.13 (-0.34 to 0.61)	0.579	0.12 (-0.36 to 0.61)	0.614	-0.01 (-0.56 to 0.55)	0.986
Posture	0.02 (-0.11 to 0.16)	0.765	0.03 (-0.10 to 0.17)	0.636	0.00 (-0.16 to 0.15)	0.958
Passive muscle tone	0.00 (-0.18 to 0.17)	0.936	-0.03 (-0.21 to 0.15)	0.769	-0.06 (-0.27 to 0.14)	0.544

Table 2: Associations between neonatal thyroid function parameters and neurodevelopmental outcomes at age 19 years (continued)

	Unadjusted	P value	Adjusted (1)	P value	Adjusted (2)	P value
Hypothyroxinemia						
IQ						
IQ total score	0.0 (-3.5 to 3.5)	0.980	0.1 (-3.4 to 3.6)	0.972	0.0 (-3.8 to 3.8)	0.995
Linguistic capacity *z* score	-0.03 (-0.23 to 0.16)	0.741	0.02 (-0.20 to 0.17)	0.861	0.04 (-0.16 to 0.24)	0.692
Mathematical capacity *z* score	0.06 (-0.17 to 0.29)	0.584	0.07 (-0.17 to 0.30)	0.583	0.10 (-0.15 to 0.36)	0.429
Logical reasoning *z* score	0.07 (-0.15 to 0.28)	0.533	0.10 (-0.11 to 0.31)	0.353	0.13 (-0.10 to 0.36)	0.158
Spatial visualization *z* score	-0.10 (-0.30 to 0.10)	0.323	-0.10 (-0.30 to 0.10)	0.314	-0.08 (-0.29 to 0.14)	0.470
Neuromotor function						
Neuromotor sum score	-0.13 (-1.9 to 1.6)	0.887	0.0 (-1.8 to 1.8)	0.998	0.6 (-1.3 to 2.5)	0.544
Hand function	-0.12 (-0.34 to 0.10)	0.276	-0.10 (-0.32 to 0.12)	0.392	-0.05 (-0.29 to 0.18)	0.653
Quality of walking	0.05 (-0.21 to 0.32)	0.686	0.06 (-0.22 to 0.33)	0.683	0.13 (-0.16 to 0.43)	0.381
Coordination	0.20 (-0.82 to 1.21)	0.700	0.22 (-0.82 to 1.26)	0.676	0.50 (-0.63 to 1.63)	0.384
Posture	0.00 (-0.29 to 0.29)	0.987	-0.01 (-0.30 to 0.29)	0.970	0.09 (-0.22 to 0.41)	0.561
Passive muscle tone	-0.09 (-0.47 to 0.28)	0.626	-0.07 (-0.46 to 0.31)	0.715	-0.01 (-0.43 to 0.41)	0.967

Values represent beta (95%CI).
Adjusted (1): gender, SES, ethnicity and parity
Adjusted (2): adjusted (1) + gestational age, SGA birth and neonatal illnesses like IRDS, IVH and sepsis

The hypothyroxinemic group had an IQ score of 100.8±14.9 points and a neuromotor score of 58.4±8.4 points. These scores were not different from those of the non-hypothyroxinemic group, which were 100.7±15.4 and 58.5±7.6 points, respectively. Fifty-three subjects had an IQ score <85 points.

Table 2 presents the associations between neonatal thyroid function parameters and continuous outcomes at age 19 years. No associations with total scores or subscores were found. Furthermore, neonatal thyroid function parameters were not associated with the odds of having an IQ <85 points (data not shown).

Analyses when including only the subjects who participated in the POPS follow-up at age 5 did not change our results (data not shown). Moreover, stratified analyses provided no evidence for gestational age-dependent effects of neonatal thyroid function parameters on outcomes (data not shown).

DISCUSSION

The main finding from our study is that previous observations linking THoP to neurodevelopmental outcome in infancy and childhood were not confirmed in young adulthood.

A limitation of our study is that only total T4 concentrations obtained during a single measurement were available for analysis. It is therefore possible that several participants in our study were misclassified as being hypothyroxinemic, because local tissue concentrations of unbound T4 can still be adequate despite a low circulating total T4 level.[2] Moreover, the reported SDS were based on the Dutch norm population. Although these scores do not reflect normality for prematurity, it was still possible to differentiate between lower and higher concentrations of T4. However, samples were not analyzed continuously, with means being calculated on a daily basis, which could lead to day-to-day fluctuation in the absolute level of total T4. However, laboratories were under permanent quality control,[21] and therefore these fluctuations were probably minimal.

Another limitation of our study is the loss to follow-up that is almost inevitable in life-course studies. Nonresponse was associated with male gender, nonwhite ethnicity, and lower parental socioeconomic class, but not with any of the perinatal characteristics. Moreover, both total T4 levels (-2.4 SD vs. -2.4 SD, $P=0.75$) and the proportion of children with THoP (26.1 vs. 28.0%; $P=.60$) did not differ between responders and nonresponders. Furthermore, our results did not change after statistical adjustment for many of the differing factors, as well as after analyzing the data while only including those subjects who participated at age 5. Therefore, we believe that response bias is unlikely to explain our associations. Hypothyroxinemic and non-hypothyroxinemic groups differed significantly in many perinatal characteristics, with hypothyroxinemia being associated with a greater proportion of neonatal morbidities. An explanation for these associations is

nonthyroidal illness.[2,7,8] Statistical adjustments for many of these factors did not change our results.

A recent meta-analysis showed that a birth weight <2,500 g was associated with a 4.98-point reduction in IQ at adolescence or young adulthood after taking publication bias into account.[28] This difference became smaller with increasing age. Among the studies included in the meta-analysis were several that had included only subjects born very preterm, and their results were similar. Another meta-analysis demonstrated that very preterm birth was associated with motor impairment in childhood.[29] Our data suggest that these relations cannot be explained by THoP.

A possible explanation for the lack of association in our study is that the total T4 concentration does not always reflect the availability of free T4 in tissues. However, previous analyses in the POPS cohort showed that a total T4 concentration <-3 SD was associated with adverse outcomes in childhood.[10,12] Alternatively, it could be possible that neurodevelopmental impairment after THoP improves with age, although our results should be interpreted carefully due to losses to follow-up. Therefore, replication of our findings in an independent sample is warranted, preferably in a prospectively designed study with serial measurements of free T4.[3]

It has been demonstrated in a placebo-controlled randomized trial that treatment with levothyroxine (at a dose of 8 μg/kg/d during the first 6 postnatal weeks) of infants born <30 weeks gestation does not improve long-term neurodevelopmental outcome.[15,17,18] However, it was suggested from a small subgroup analysis that treatment could be beneficial for infants in the extremely preterm range. Our study showed no evidence for a gestational age-dependent effects of THoP on IQ or neuromotor function at 19 years of age.

In a recent study among infants born <28 weeks gestation, three different doses of levothyroxine (of 4, 8, and 16 μg/kg/d for 6 wk), either continuous or as bolus, as well as iodide, were compared to placebo.[16] The groups on levothyroxine also received triiodothyronine (T3) continuously, at a dose of 1 μg/kg/d, during the first 14 days. Although mental and motor performance at age 3 years did not differ between treated and untreated children, or between the treatment arms, it was concluded that "further trials are warranted." Our findings, showing that the neurodevelopmental sequelae of THoP could not be extrapolated to young adult age, question whether new trials are necessary.

In conclusion, we did not find an association between THoP and neurodevelopmental outcome in young adulthood.

REFERENCES

1. Mercado M, Yu VY, Francis I, Szymonowicz W, Gold H. Thyroid function in very preterm infants. Early Hum Dev 1988; 16:131-141

2. van Wassenaer AG, Kok JH, Dekker FW, de Vijlder JJ. Thyroid function in very preterm infants: influences of gestational age and disease. Pediatr Res 1997; 42:604-609

3. Williams FL, Simpson J, Delahunty C, Ogston SA, Bongers-Schokking JJ, Murphy N, van Toor H, Wu SY, Visser TJ, Hume R, Collaboration from the Scottish Preterm Thyroid G. Developmental trends in cord and postpartum serum thyroid hormones in preterm infants. J Clin Endocrinol Metab 2004; 89:5314-5320

4. Fisher DA. Thyroid function in premature infants. The hypothyroxinemia of prematurity. Clin Perinatol 1998; 25:999-1014, viii

5. Ares S, Escobar-Morreale HF, Quero J, Duran S, Presas MJ, Herruzo R, Morreale de EG. Neonatal hypothyroxinemia: effects of iodine intake and premature birth. J Clin Endocrinol Metab 1997; 82:1704-1712

6. Murphy N, Hume R, van TH, Matthews TG, Ogston SA, Wu SY, Visser TJ, Williams FL. The hypothalamic-pituitary-thyroid axis in preterm infants; changes in the first 24 hours of postnatal life. J Clin Endocrinol Metab 2004; 89:2824-2831

7. Pavelka S, Kopecky P, Bendlova B, Stolba P, Vitkova I, Vobruba V, Plavka R, Houstek J, Kopecky J. Tissue metabolism and plasma levels of thyroid hormones in critically ill very premature infants. Pediatr Res 1997; 42:812-818

8. Simpson J, Williams FL, Delahunty C, van TH, Wu SY, Ogston SA, Visser TJ, Hume R. Serum thyroid hormones in preterm infants and relationships to indices of severity of intercurrent illness. J Clin Endocrinol Metab 2005; 90:1271-1279

9. De Vries LS, Heckmatt JZ, Burrin JM, Dubowitz LM, Dubowitz V. Low serum thyroxine concentrations and neural maturation in preterm infants. Arch Dis Child 1986; 61:862-866

10. Meijer WJ, Verloove-Vanhorick SP, Brand R, van den Brande JL. Transient hypothyroxinaemia associated with developmental delay in very preterm infants. Arch Dis Child 1992; 67:944-947

11. Delahunty C, Falconer S, Hume R, Jackson L, Midgley P, Mirfield M, Ogston S, Perra O, Simpson J, Watson J, Willatts P, Williams F. Levels of neonatal thyroid hormone in preterm infants and neurodevelopmental outcome at 5 1/2 years: millennium cohort study. J Clin Endocrinol Metab 2010; 95:4898-4908

12. Den Ouden AL, Kok JH, Verkerk PH, Brand R, Verloove-Vanhorick SP. The relation between neonatal thyroxine levels and neurodevelopmental outcome at age 5 and 9 years in a national cohort of very preterm and/or very low birth weight infants. Pediatr Res 1996; 39:142-145

13. Reuss ML, Paneth N, Pinto-Martin JA, Lorenz JM, Susser M. The relation of transient hypothyroxinemia in preterm infants to neurologic development at two years of age. N Engl J Med 1996; 334:821-827

14. Osborn DA. Thyroid hormones for preventing neurodevelopmental impairment in preterm infants. Cochrane Database Syst Rev 2001:CD001070

15. van Wassenaer AG, Kok JH, de Vijlder JJ, Briet JM, Smit BJ, Tamminga P, van BA, Dekker FW, Vulsma T. Effects of thyroxine supplementation on neurologic development in infants born at less than 30 weeks' gestation. N Engl J Med 1997; 336:21-26

16. van Wassenaer-Leemhuis A, Ares S, Golombek S, Kok J, Paneth N, Kase J, LaGamma EF. Thyroid hormone supplementation in preterm infants born before 28 weeks gestational age and neurodevelopmental outcome at age 36 months. Thyroid 2014; 24:1162-1169

17. van Wassenaer AG, Westera J, Houtzager BA, Kok JH. Ten-year follow-up of children born at <30 weeks' gestational age supplemented with thyroxine in the neonatal period in a randomized, controlled trial. Pediatrics 2005; 116:e613-e618

18. Briet JM, van Wassenaer AG, Dekker FW, de Vijlder JJ, van Baar A, Kok JH. Neonatal thyroxine supplementation in very preterm children: developmental outcome evaluated at early school age. Pediatrics 2001; 107:712-718

19. Verloove-Vanhorick SP, Verwey RA, Brand R, Gravenhorst JB, Keirse MJ, Ruys JH. Neonatal mortality risk in relation to gestational age and birthweight. Results of a national survey of preterm and very-low-birthweight infants in the Netherlands. Lancet 1986; 1:55-57

20. Chopra IJ. A radioimmunoassay for measurement of thyroxine in unextracted serum. J Clin Endocrinol Metab 1972; 34:938-947

21. Schopman WE, E; de Kock, HW; Rechsteiner, J; Tertoolen, JFW. Analytische vergelijking van de CHT-laboratoria. Rapporten over het 1e t/m 4e kwartaal 1983 1984;

22. Schopman W. Het L-thyroxine gehalte in ponsjes van bloedvlekken op PKU-CHT kaarten in de Nederlandse proefregio. Nulcleair Geneeskundig Bulletin 1979; 1:4-10

23. Bleichrodt N, Berg RH. Multicultural Capacity Test: Intermediate Level (MCT-M) - Manual. Amsterdam: NOA.

24. Samsom JF, de GL, Cranendonk A, Bezemer D, Lafeber HN, Fetter WP. Neuromotor function and school performance in 7-year-old children born as high-risk preterm infants. J Child Neurol 2002; 17:325-332

25. Touwen BC. The Examination of the Child With Minor Neurological Dysfunction: Clinics in Developmental Medicine Series. Vol 71. London, England: Heinemann.

26. van Wassenaer AG, Kok JH. Trials with thyroid hormone in preterm infants: clinical and neurodevelopmental effects. Semin Perinatol 2008; 32:423-430

27. Hille ET, Weisglas-Kuperus N, van Goudoever JB, Jacobusse GW, Ens-Dokkum MH, de GL, Wit JM, Geven WB, Kok JH, de Kleine MJ, Kollee LA, Mulder AL, van Straaten HL, De Vries LS, van Weissenbruch MM, Verloove-Vanhorick SP. Functional outcomes and participation in young adulthood for very preterm and very low birth weight infants: the Dutch Project on Preterm and Small for Gestational Age Infants at 19 years of age. Pediatrics 2007; 120:e587-e595

28. Kormos CE, Wilkinson AJ, Davey CJ, Cunningham AJ. Low birth weight and intelligence in adolescence and early adulthood: a meta-analysis. J Public Health (Oxf) 2014; 36:213-224

29. de Kieviet JF, Piek JP, Aarnoudse-Moens CS, Oosterlaan J. Motor development in very preterm and very low-birth-weight children from birth to adolescence: a meta-analysis. JAMA 2009; 302:2235-2242

12

Transient hypothyroxinemia of prematurity and problem behavior in young adulthood

Jonneke J. Hollanders,
Sylvia M. van der Pal,
Paul H. Verkerk,
Joost Rotteveel,
Martijn J.J. Finken

ABSTRACT

Introduction

Preterm newborns are at risk of developing transient hypothyroxinemia of prematurity (THoP), which has been associated with subsequent neurodevelopmental impairments. Behavioral outcomes at adult age after THoP have never been reported.

Aim

To examine whether there is an association between THoP and problem behavior at young adult age.

Methods

This study was part of the follow-up of 19-year-old subjects born very preterm (i.e., <32 weeks) and/or with a very low birth weight (i.e.,<1500g) from the Project On Preterm and Small-for-gestational-age infants (POPS) cohort. We included 468 subjects of the POPS cohort; of whom 123 had THoP. Thyroxine (T4) concentrations were obtained through the national neonatal screening program for congenital hypothyroidism. THoP was defined as a T4 concentration <-3 SD (approximately 60nmol/L). At age 19, behavior was assessed using the Young Adult Self Report and the Young Adult Behavioral Checklist for parents.

Results

THoP was associated with a 1.8 (95% confidence interval (CI): 1.01-3.4) -fold increased odds of self-reported Internalizing behavior, as well as with a 1.9 (95% CI: 1.1-3.1) -fold increased odds of parent-reported Total problem behavior. These relations persisted after correction for demographic and perinatal variables. Similar associations were absent for the other self-reported and parent-reported syndrome and problem scales.

Conclusions

THoP was associated with more internalizing and total problem behavior at age 19. While our observations warrant more awareness of problem behavior in preterm infants, at present, it is unclear whether these associations are causal and screening for THoP does not seem necessary.

INTRODUCTION

Thyroid hormones are crucial for the developing brain, where they help to control cell migration, proliferation and differentiation.[1] According to the construct of Zoeller and Rovet (2004) [1] in the first half of pregnancy, thyroid hormones play a role in the development of visual attention and processing, and of fine motor skills. During the second half of pregnancy, thyroid hormones are proposed to be involved in the development of memory, visuospatial skills, and fine and gross motor skills [1] Because the fetal thyroid starts to become functional from the 12th week of gestation,[2] adequate maternal-fetal transfer of thyroxine (T4) in the first trimester is essential for early brain development. However, maternal T4 remains a major fraction of fetal serum T4 after the onset of fetal thyroid hormone production, and continues to play a role in fetal neurodevelopment until birth.[3]

Small reductions in the early supply of thyroid hormones might lead to permanent alterations in behavioral patterns. In animals, disruptions in the transplacental supply of thyroid hormones resulted in morphological changes in the cerebral cortex and hippocampus of the pups.[4,5] In humans, functional changes in these structures have been proposed to underlie attention deficit/hyperactivity disorder (ADHD).[6,7] Indeed, small reductions in the maternal thyroid function during early pregnancy were associated with ADHD symptoms in 8-year-old offspring.[8,9] Similarly, the risk of developing attention problems was also increased in subjects with early-treated congenital hypothyroidism (CHT).[10]

Preterm birth has been associated with attention problems and internalizing behavior. These patterns have been reported to persist into adulthood.[11] Nowadays, of all liveborn children, 11.1 (range: 5-18) % are born <37 week of gestation, and 1.7% are born <32 weeks of gestation.[12]

After preterm birth, a transient reduction in the thyroid hormone level, known as transient hypothyroxinemia of prematurity (THoP), has been estimated to occur in approximately 20% of infants, although it is even more common with increasing degrees of prematurity.[13,14] It can be attributed to the sudden disruption of the transplacental T4 supply.[15-17] Additionally, hypothalamus-pituitary-thyroid axis immaturity, reduced thyroidal iodine reserves, acute illnesses, and treatment with dopamine also contribute to the development of THoP.[15,18-21] T4 concentrations are therefore lower in extremely preterm infants than in fetuses of the same post-conceptional age[22] THoP usually restores spontaneously within 6 to 8 weeks.[23] There is conflicting evidence with regard to the effects of THoP on long-term neurodevelopmental outcomes. Although THoP was associated with adverse neurodevelopment in infancy and childhood,[13,24,25] the only study that had provided follow-up into adulthood was negative.[26] Whether THoP is associated with problem behavior, in particular attention problems, has not been addressed to date.

Therefore, we aimed to investigate whether there is an association between THoP and problem behavior at young adult age. Here, we provide a prospective follow-up of a well-described cohort of males and females born very preterm (i.e., <32 weeks) and/or with a very low birth weight (i.e., <1,500 g) in whom behavioral outcomes were assessed at age 19 years and whose T4 levels were determined during a T4-based national screening program for CHT. Based on previous studies that addressed the effects of early disruptions in the supply of thyroid hormones,[8-10] we expected to find more problem behavior in subjects with THoP, especially attention problems.

METHODS

Study population

The Project On Preterm and Small-for-gestational-age infants (POPS) cohort is a nation-wide birth cohort study, which comprised 94% (n=1,338) of infants who were born alive in the Netherlands in January-December 1983 with a gestational age of less than 32 weeks and/or with a birth weight below 1,500 grams.[27] In 1983, 101 out of 115 level 1 to level 3 hospitals throughout the Netherlands collected data. At age 1 year, 975 subjects (73%) were still alive, and they were followed up throughout childhood. At age 19 years, another follow-up was scheduled; 959 subjects (72%) were still alive at that point. Of these, 745 subjects had known neonatal T4 concentrations. In keeping with previous analyses in the POPS cohort with regard to THoP,[24,25] we excluded subjects whose T4 concentrations were measured before postnatal day 5 or after day 17 (n=66) or who received thyroid hormone supplementation during their hospital stay (n=5). We also excluded subjects with severe congenital malformations, such as Down syndrome, central nervous system defects or inborn errors of metabolism (n=10), severe sensory handicaps (n=8) and congenital hypothyroidism (n=1). 655 subjects were therefore eligible for our study. Figure 1 shows the flowchart of our study sample.

Iodine intake of the infants and/or their mothers was not measured in our cohort. However, iodine supplementation guidelines were intensified in The Netherlands in 1982, and iodine status was subsequently considered sufficient in a survey among school children in the 1990s.[28]

The study was approved by the medical ethical committees of the participating centers, and written informed consent was obtained from all participants.

Laboratory analysis

T4 was measured in the context of the neonatal screening for CHT. From April 1983 onward, results were prospectively collected,[24] although the T4 values of 54 subjects born before April 1983 could be acquired retrospectively. T4 concentrations from filter paper

eluates were measured in duplicate by radioimmunoassay [29] in the five laboratories connected to the national screening program. These laboratories were under permanent quality control.[30] T4 levels in the eluates were expressed as standard deviations from the mean, which was calculated on a daily basis.[31] In our sample, T4 SDS was normally distributed. The intra-assay and inter-assay coefficients of variation in the eluates were 8 and 10%, respectively. In line with previous analyses in this cohort,[24,25] THoP was defined as a T4 concentration of <-3SD (approximately 60 nmol/L).[25]
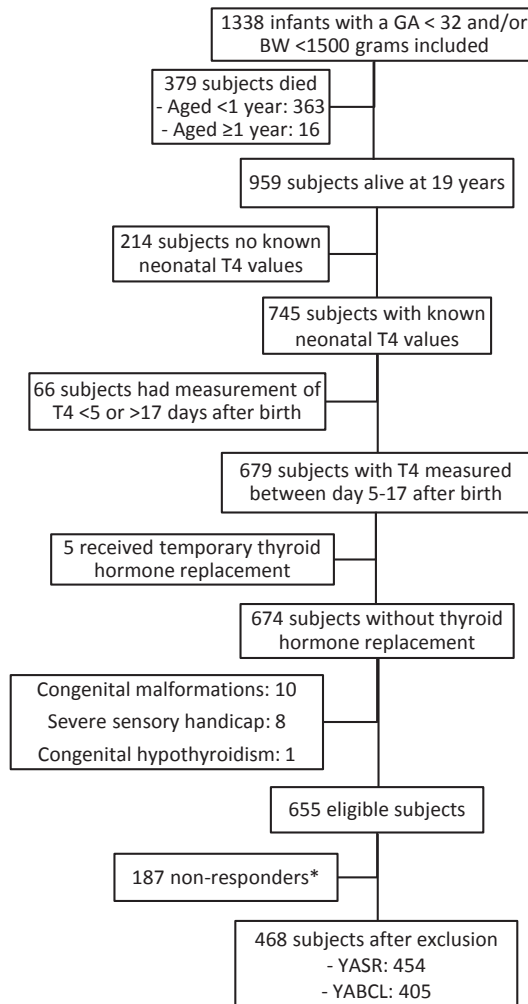
Figure 1: Flowchart of the inclusion of POPS subjects at age 19
*included in this number: subjects who were followed up but did not return the YASR or YABCL
YASR, Young Adult Self Report; YABCL, Young Adult Behavior Checklist

In a number of subjects re-tests were performed in order to ascertain the transient nature of hypothyroxinemia. Whether and when re-tests were done was at the discretion of the treating clinician, and T4 was analyzed in the same manner as the initial T4 measurements.

Study procedure

Behavior at age 19 was studied using the Young Adult Self Report (YASR) and/or the Young Adult Behavior Checklist (YABCL). The YASR was used to assess problem behavior from the perspective of the adolescent, while the YABCL assessed problem behavior from the parent's or caregiver's perspective. Both questionnaires were developed by Achenbach and provide a standardized description of behavior, feelings, thoughts and competences in people aged 18 to 30 years.[32] The YASR and YABCL contain 130 and 109 items, respectively. Each item is scored according to a 3-step scale, where 0 = "not true", 1 = "somewhat or sometimes true", and 2 = "very often or often true". The questions pertain to the preceding 6 months. From these items, 8 syndrome scales are derived: Anxious/Depressed, Withdrawn, Somatic complaints, Thought problems, Attention problems, Intrusive behavior, Aggressive behavior, and Delinquent behavior. Although the syndrome scale Attention problems assesses some aspects of ADHD, confirmation of ADHD diagnoses in childhood and/or young adulthood was not known for our cohort. Subsequently, three problem scales are calculated: the problem scale 'Internalizing behavior' is the sum of the syndrome scales Anxious/Depressed and Withdrawn, the problem scale 'Externalizing behavior' is the sum of the syndrome scales Aggressive behavior, Delinquent behavior and Intrusive behavior, and the Total problems scale is the sum of all individual items. For each syndrome and problem scale a clinical cut-off, as well as a borderline clinical cut-off has been determined on the basis of a non-referred population.[32] For this study, we used the borderline clinical cut-off, since we aimed at exploring associations even in the subclinical range. For the syndrome scales, the 95th percentile is considered the borderline clinical cut-off point. For the problem scales, the 83rd percentile is considered as the borderline clinical cut-off point. These cut-off points are gender-specific.

Statistical analysis

Multivariate logistic regression was used to study the effect of THoP on behavioral outcomes. Behavioral outcomes were dichotomized according to the borderline clinical cut-off point. The model was first adjusted for the demographic characteristics gender, socio economic status (SES), ethnicity and parity. Analyses were repeated after also adjusting for the perinatal characteristics gestational age (GA) and being born small-for-gestational-age (SGA), and for neonatal illnesses, such as infant respiratory distress syndrome (diagnosed clinically and/or radiographically), intraventricular hemorrhage

(diagnosed clinically and/or ultrasonographically) and sepsis (diagnosed hematologically and/or through blood culture). Next, analyses were repeated with T4 as a continuous variable, and after stratifying the data according to gestational age into < and ≥ 29 weeks. We based this cut-off point on a randomized placebo-controlled trial, showing gestational age-dependent effects of early levothyroxine treatment on behavior.[33] A *P* value of ≤0.05 was considered statistically significant.

Our study was not designed specifically to address the effect of THoP on behavioral outcomes. Our sample size enabled us to detect a difference of 1.3 points on the attention scale of both the YASR and YABCL, assuming a maximum SD of 4 points, with a power of 80% and a significance level of 0.05.

RESULTS

Descriptives

Of the 655 eligible subjects, 468 subjects and/or their parents (71.5%) filled in the YASR and/or the YABCL. Non-responders were more likely to be male and to have a non-Caucasian ethnicity, a lower SES and a younger maternal age at birth than responders (Table 1). Non-response was unrelated to perinatal characteristics, the T4 concentration (-2.4 SD vs. -2.4 SD, P=1.00), or the proportion of subjects with THoP (26.3 % vs. 28.3%, P=0.59).

Among the 468 participants, 391 returned both the YASR and the YABCL, 14 returned the YASR only, and 63 returned the YABCL only. Of these, 123 developed THoP, which was diagnosed in 67% of the infants with a GA <28 weeks, in 27% of the infants with a GA 28-31 weeks, and in 12% of the infants with a GA ≥32 weeks. Table 1 details the baseline characteristics of subjects with and without THoP.

Re-testing was performed in 427 of 468 subjects. None of the re-tested subjects had a T4 concentration <-3SD after 53 days postpartum, which is consistent with the known duration of THoP.[23]

YASR

Table 2 presents the self-reported behavioral outcomes at age 19. THoP was associated with a higher odds of Internalizing behavior. This relation persisted after correction for demographic and perinatal characteristics. Hypothyroxinemic subjects showed a non-significant tendency towards more Withdrawn behavior and Total problem behavior; these associations only became significant after adjustment for confounders. There were no associations between THoP and other self-reported syndrome or problem scales.

Table 1: General and perinatal characteristics

	Subjects with THoP (n=123, 26.3%)	Subjects without THoP (n=345, 73.7%)	P value*	Non-responders (n=187)	P value†
General					
Male sex (%)	64 (52.0)	154 (44.6)	0.158	115 (61.5)	0.001
White (%)	102 (85.0)	316 (91.9)	0.030	142 (75.9)	<0.001
Low socio-economic status (%)	45 (36.9)	123 (36.0)	0.856	108 (60.7)	<0.001
First child (%)	64 (52.9)	197 (57.1)	0.422	97 (51.9)	0.337
Perinatal					
Maternal age (years)	26.7±5.0	27.4±6.1	0.279	26.1±4.9	0.021
Gestational age (weeks)	29.4±2.1	31.4±2.4	<0.001	31.0±2.7	0.549
Birth weight (grams)	1,191±254	1,306±257	<0.001	1,279±251	0.876
SGA birth (%)	30 (24.6)	148 (42.9)	<0.001	76 (40.6)	0.549
Apgar score ≥7 after 5 min (%)	93 (75.6)	298 (86.4)	0.010	157 (84.0)	0.987
Part of multiple pregnancy (%)	33 (26.8)	78 (22.6)	0.345	37 (19.8)	0.277
Respiratory distress syndrome (%)	73 (59.3)	122 (35.4)	<0.001	67 (35.8)	0.168
Intraventricular hemorrhage (%)	45 (36.6)	42 (12.2)	<0.001	30 (16.0)	0.442
Sepsis (%)	54 (43.9)	88 (25.6)	<0.001	69 (36.9)	0.109
Necrotizing enterocolitis (%)	7 (5.7)	19 (5.5)	0.939	8 (4.3)	0.506
Total T4 (SD)	-3.5±0.4	-1.9±0.7	<0.001	-2.4±1.0	0.995

Values represent mean±SD or n (%). Continuous variables were compared with the unpaired t test. Dichotomous variables were compared with the Chi square test.

*P value when comparing subjects with and without THoP

† P value when comparing responders (subjects with and without THoP combined) and non-responders

YABCL

Table 3 displays the parent-reported behavioral outcomes at age 19. THoP was associated with a higher score for Total problem behavior (29.0±22.6 vs. 23.3±20.6; p=0.023). Moreover, THoP was associated with a higher odds of Total problem behavior, which remained significant after adjusting for all potential confounders. THoP also increased the odds of Thought problems and Internalizing problem behavior. However, these associations lost significance after adjustment for demographic and perinatal variables. There were no associations between THoP and other parent-reported syndrome or problem scales.

When analyzing T4 SDS as a continuous variable, no association was found with any of the self- or parent-reported syndrome or problem scales (Supplementary Table 1). Stratification by gestational age did not change our results (data not shown).

Correlations between self- and parent reported behaviors ranged between 0.40 and 0.61 (all P values <0.001).

Table 2: Risk of problem behavior according to the Young Adult Self Report (YASR) after THoP.

	Subjects with THoP	Subjects without THoP	OR (95% CI)					
	N = 118	n = 336	Unadjusted	P value	Adjusted (1)	P value	Adjusted (2)	P value
Syndrome scales								
Anxious/depressed	7.6±7.1	6.3±6.2	2.5 (0.95-6.4)	0.064	2.9 (1.1-7.6)	0.034*	2.3 (0.7-7.1)	0.150
Withdrawn	2.9±2.9	2.6±2.4	1.7 (0.9-3.5)	0.112	2.0 (1.02-4.1)	0.045*	2.4 (1.1-5.2)	0.031*
Somatic complaints	3.7±3.7	3.3±3.6	1.1 (0.5-2.6)	0.806	0.9 (0.4-2.4)	0.907	1.5 (0.5-4.0)	0.459
Thought problems	0.5±1.3	0.3±0.9	1.3 (0.6-3.1)	0.528	1.5 (0.6-3.7)	0.332	1.7 (0.6-4.7)	0.283
Attention problems	2.7±2.5	2.7±2.2	1.4 (0.6-3.3)	0.455	1.7 (0.7-4.1)	0.261	1.6 (0.6-4.4)	0.371
Intrusive behavior	2.1±2.1	1.7±1.9	1.9 (0.3-11.7)	0.476	1.9 (0.3-12.2)	0.511	5.7 (0.7-47.4)	0.107
Aggressive behavior	2.7±2.7	2.5±2.9	0.5 (0.1-2.4)	0.402	0.6 (0.1-2.9)	0.544	0.9 (0.2-5.7)	0.950
Delinquent behavior	1.0±1.6	1.1±1.7	0.3 (0.0-2.3)	0.234	0.3 (0.0-2.8)	0.315	0.3 (0.0-3.3)	0.355
Problem scales								
Internalizing behavior	10.5±9.5	8.9±8.0	1.8 (1.01-3.4)	0.045*	2.0 (1.1-3.7)	0.032*	2.4 (1.2-5.0)	0.016*
Externalizing behavior	5.8±5.2	5.4±5.2	0.5 (0.2-1.3)	0.134	0.6 (0.2-1.5)	0.253	0.8 (0.3-2.4)	0.687
Total problem behavior	34.8±25.4	30.9±23.0	1.7 (0.9-3.1)	0.082	1.8 (0.98-3.5)	0.058	2.4 (1.2-5.0)	0.016*

Values represent mean±SD or OR (95% CI).
Adjusted (1): gender, SES, ethnicity and parity
Adjusted (2): adjusted (1) + gestational age, SGA-status and neonatal illnesses (IRDS, IVH and sepsis)
* P <0.05

Table 3: Risk of problem behavior according to the Young Adult Behavioral Checklist (YABCL) after THoP.

	Subjects with THoP	Subjects without THoP	OR (95% CI)					
	n = 107	n = 298	Unadjusted	P value	Adjusted (1)	P value	Adjusted (2)	P value
Syndrome scales								
Anxious/depressed	5.9±5.4	5.0±4.9	1.3 (0.6-2.7)	0.517	1.3 (0.6-2.8)	0.508	1.4 (0.6-3.4)	0.441
Withdrawn	2.1±2.3	1.6±1.8	1.4 (0.7-3.0)	0.381	1.4 (0.7-3.0)	0.378	1.5 (0.6-3.5)	0.391
Somatic complaints	2.4±2.5	2.2±2.6	0.7 (0.3-1.8)	0.506	0.8 (0.3-1.9)	0.807	0.9 (0.3-2.5)	0.858
Thought problems	0.9±1.7	0.6±1.4	2.1 (1.01-4.2)	0.048*	2.1 (0.97-4.4)	0.059	1.8 (0.7-4.2)	0.200
Attention problems	5.3±4.3	4.4±4.0	1.5 (0.7-3.1)	0.297	1.7 (0.8-3.5)	0.192	1.8 (0.8-4.4)	0.177
Intrusive behavior	2.2±2.0	1.8±2.3	0.4 (0.1-1.9)	0.267	0.4 (0.1-2.0)	0.278	0.9 (0.2-4.7)	0.927
Aggressive behavior	4.2±5.1	3.5±4.6	1.5 (0.6-3.7)	0.435	1.5 (0.6-3.9)	0.409	1.9 (0.6-5.6)	0.241
Delinquent behavior	0.9±1.4	0.8±1.6	0.6 (0.1-4.9)	0.603	0.6 (0.1-5.0)	0.602	1.0 (0.1-12.0)	0.988
Problem scales								
Internalizing behavior	8.0±6.6	6.6±6.1	1.7 (1.03-2.7)	0.039*	1.6 (0.97-2.6)	0.068	1.7 (0.96-3.0)	0.068
Externalizing behavior	7.3±7.0	6.2±7.5	1.4 (0.8-2.5)	0.281	1.5 (0.8-2.7)	0.208	1.8 (0.9-3.6)	0.110
Total problem behavior	29.0±22.6	23.3±20.6*	1.9 (1.1-3.1)	0.018*	1.8 (1.1-3.1)	0.025*	2.4 (1.3-4.5)	0.005*

Values represent mean±SD or OR (95% CI).
Adjusted (1): gender, SES, ethnicity and parity
Adjusted (2): adjusted (1) + gestational age, SGA-status and neonatal illnesses (IRDS, IVH and sepsis)
* P <0.05

DISCUSSION

We found that THoP was associated with increased odds of internalizing and total problem behavior at age 19 years. THoP was not associated with attention problems.

A recent meta-analysis found that internalizing behavior and attention problems were more common after preterm birth.[11] In several cohorts of preterm infants, including the POPS cohort, these behavioral patterns were demonstrated to persist into young adulthood.[34,35] Whether THoP has an influence on adult behavior has not been addressed to date, with the exception of our study. However, there are some data available regarding behavior after subtle impairments in the early supply of thyroid hormones. The Generation R study, which follows a community-based sample of mothers and their children, found that hypothyroxinemia in women early in gestation was associated with an increased risk for ADHD symptoms in their 8-year-old children.[8] Another study found higher rates of ADHD in children born to mothers with higher TSH levels early in their pregnancies.[9]

We found that THoP was associated with increases in internalizing and total problem behavior at age 19. However, we did not find evidence for an effect of THoP on attention problems. Since thyroid hormones appear to influence different domains of brain development at different stages during gestation,[1] it is possible that the developmental window in which a transient decrease in the T4 supply could have an influence on attention might be earlier than in (the period equivalent to) the third trimester of pregnancy,[8,36] although children with treated CHT were also at risk of attention problems.[10] An alternative explanation is that the causes of attention problems after preterm birth are complex.[37] Therefore, in our sample, a possible effect of THoP on attention problems might go unnoticed.

A randomized placebo-controlled trial in infants born before 30 weeks of gestation showed no effect of treatment with levothyroxine (at a dose of 8 mcg/kg/d during the first 6 postnatal weeks) on behavior at the ages of 5.7 and 10 years.[38,39] However, in *post hoc* analyses benefit was evident only among children of 25-26 weeks' gestation at age 5.7, but not at age 10. In our study, we did not find evidence for gestational age-dependency of our associations.

Currently, the POPS cohort is the only preterm population that has been followed up into young adulthood with regard to THoP as far as we are aware. It was previously shown that THoP is not associated with IQ or neuromotor development at age 19.[26] Presently, preterm infants are not routinely screened for THoP, and our results do not indicate this practice should be altered. On the other hand, behavioral problems were more prevalent in subjects with THoP. While the causality of these associations is still unclear, more awareness of problem behavior in preterm infants is warranted, especially since the risk of behavioral problems is already increased in these populations.[11]

Our study has several strengths and limitations. The major strength of our study is the long follow-up period, which extends into adulthood. Thus far, the long-term effects of THoP have been studied up until childhood.[13,24] Another strength of our study is the use of both the YASR and the YABCL, enabling us to study both the adolescent's and parent's, or caregiver's, perspective. Both yielded similar results. Moreover, 80% of the subjects in our cohort were still living with their parents, and previous research has shown that parents of children who were born preterm provide highly valid data,[40] suggesting a reliable parental evaluation in our sample.

Our study has several weaknesses. First, only total T4 concentrations, obtained during a single measurement, were available for analysis. It is possible that several participants in our study were misclassified as hypothyroxinemic, since the systemic total T4 level does not always parallel the tissue free T4 concentration.[20] Moreover, the reported standard deviation scores were based on the Dutch norm population. Despite this limitation, it was still possible to differentiate between lower and higher concentrations of T4. Second, thyroid function tests were not part of the assessment at age 19 years. Recent evidence in this research field indicates that school performance is influenced for an important part by the current thyroid hormone status.[36] However, in a small subset of the POPS cohort, thyroid function tests in young adulthood were not different compared to age-matched controls born at term.[41] Third, the losses to follow-up throughout the years might have introduced bias. However, the numbers lost to follow-up were acceptably low for our study. While non-response was associated with male gender, non-white ethnicity and lower socio-economic status, it was not associated with any of the perinatal characteristics.[42] Moreover, we found no differences between responders and non-responders with regard to both total T4 levels and the proportion of subjects with THoP. Furthermore, our results did not change after statistical adjustment for many of the differing factors, and we therefore believe that response bias is unlikely to explain our associations. Additionally, subjects with and without THoP differed significantly in many perinatal characteristics. Adjusting for many of these factors in our analyses did not change our results. This makes bias less likely, but the possibility cannot be excluded completely.

Although the sudden disruption of the maternal thyroid supply is considered an important factor in the development of THoP, it is unknown whether higher maternal T4 levels prior to preterm delivery protect against it. In term-born infants, maternal TSH and FT4 at the end of the first trimester were positively correlated to cord blood TSH and FT4.[43] The correlation with thyroid function in the first days of life was not studied. However, these findings cannot simply be extrapolated to preterm infants, whose thyroid function is also determined by factors related to illness, treatment and immaturity. Future research should elucidate this.

CONCLUSION

We found that THoP was associated with increases in internalizing and total problem behavior at age 19. While our observations warrant more awareness of problem behavior during the follow-up of preterm infants, at present, it is unclear whether these associations are causal. Therefore, combined with recent findings reporting on the lack of association between THoP and neurodevelopment outcomes once adulthood is reached,[26] at present, there is no indication to screen preterm infants for THoP.

REFERENCES

1. Zoeller RT, Rovet J. Timing of thyroid hormone action in the developing brain: clinical observations and experimental findings. J Neuroendocrinol 2004; 16:809-818

2. Calvo RM, Jauniaux E, Gulbis B, Asuncion M, Gervy C, Contempre B, Morreale de EG. Fetal tissues are exposed to biologically relevant free thyroxine concentrations during early phases of development. J Clin Endocrinol Metab 2002; 87:1768-1777

3. Morreale de EG, Obregon MJ, Escobar del RF. Role of thyroid hormone during early brain development. Eur J Endocrinol 2004; 151 Suppl 3:U25-U37

4. Auso E, Lavado-Autric R, Cuevas E, Del Rey FE, Morreale de EG, Berbel P. A moderate and transient deficiency of maternal thyroid function at the beginning of fetal neocorticogenesis alters neuronal migration. Endocrinology 2004; 145:4037-4047

5. Lavado-Autric R, Auso E, Garcia-Velasco JV, Arufe Mdel C, Escobar del Rey F, Berbel P, Morreale de Escobar G. Early maternal hypothyroxinemia alters histogenesis and cerebral cortex cytoarchitecture of the progeny. J Clin Invest 2003; 111:1073-1082

6. Castellanos FX, Giedd JN, Marsh WL, Hamburger SD, Vaituzis AC, Dickstein DP, Sarfatti SE, Vauss YC, Snell JW, Lange N, Kaysen D, Krain AL, Ritchie GF, Rajapakse JC, Rapoport JL. Quantitative brain magnetic resonance imaging in attention-deficit hyperactivity disorder. Arch Gen Psychiatry 1996; 53:607-616

7. Riccio CA, Hynd GW, Cohen MJ, Gonzalez JJ. Neurological Basis of Attention-Deficit Hyperactivity Disorder. Except Children 1993; 60:118-124

8. Modesto T, Tiemeier H, Peeters RP, Jaddoe VW, Hofman A, Verhulst FC, Ghassabian A. Maternal Mild Thyroid Hormone Insufficiency in Early Pregnancy and Attention-Deficit/Hyperactivity Disorder Symptoms in Children. JAMA Pediatr 2015; 169:838-845

9. Pakkila F, Mannisto T, Pouta A, Hartikainen AL, Ruokonen A, Surcel HM, Bloigu A, Vaarasmaki M, Jarvelin MR, Moilanen I, Suvanto E. The impact of gestational thyroid hormone concentrations on ADHD symptoms of the child. J Clin Endocrinol Metab 2014; 99:E1-E8

10. Rovet JF. Congenital hypothyroidism: an analysis of persisting deficits and associated factors. Child Neuropsychol 2002; 8:150-162

11. Aarnoudse-Moens CS, Weisglas-Kuperus N, van Goudoever JB, Oosterlaan J. Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. Pediatrics 2009; 124:717-728

12. Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller AB, Narwal R, Adler A, Vera GC, Rohde S, Say L, Lawn JE. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. Lancet 2012; 379:2162-2172

13. Delahunty C, Falconer S, Hume R, Jackson L, Midgley P, Mirfield M, Ogston S, Perra O, Simpson J, Watson J, Willatts P, Williams F. Levels of neonatal thyroid hormone in preterm infants and neurodevelopmental outcome at 5 1/2 years: millennium cohort study. J Clin Endocrinol Metab 2010; 95:4898-4908

14. Reuss ML, Paneth N, Lorenz JM, Susser M. Correlates of low thyroxine values at newborn screening among infants born before 32 weeks gestation. Early Hum Dev 1997; 47:223-233

15. Fisher DA. Thyroid function in premature infants. The hypothyroxinemia of prematurity. Clin Perinatol 1998; 25:999-1014, viii

16. Glinoer D, Delange F. The potential repercussions of maternal, fetal, and neonatal hypothyroxinemia on the progeny. Thyroid 2000; 10:871-887

17. Vulsma T, Gons MH, de Vijlder JJ. Maternal-fetal transfer of thyroxine in congenital hypothyroidism due to a total organification defect or thyroid agenesis. N Engl J Med 1989; 321:13-16

18. Murphy N, Hume R, van TH, Matthews TG, Ogston SA, Wu SY, Visser TJ, Williams FL. The hypothalamic-pituitary-thyroid axis in preterm infants; changes in the first 24 hours of postnatal life. J Clin Endocrinol Metab 2004; 89:2824-2831

19. Ares S, Escobar-Morreale HF, Quero J, Duran S, Presas MJ, Herruzo R, Morreale de EG. Neonatal hypothyroxinemia: effects of iodine intake and premature birth. J Clin Endocrinol Metab 1997; 82:1704-1712

20. van Wassenaer AG, Kok JH, Dekker FW, de Vijlder JJ. Thyroid function in very preterm infants: influences of gestational age and disease. Pediatr Res 1997; 42:604-609

21. Williams FL, Ogston SA, van TH, Visser TJ, Hume R. Serum thyroid hormones in preterm infants: associations with postnatal illnesses and drug usage. J Clin Endocrinol Metab 2005; 90:5954-5963

22. Morreale de EG, Ares S. The hypothyroxinemia of prematurity. J Clin Endocrinol Metab 1998; 83:713-716

23. Mercado M, Yu VY, Francis I, Szymonowicz W, Gold H. Thyroid function in very preterm infants. Early Hum Dev 1988; 16:131-141

24. Den Ouden AL, Kok JH, Verkerk PH, Brand R, Verloove-Vanhorick SP. The relation between neonatal thyroxine levels and neurodevelopmental outcome at age 5 and 9 years in a national cohort of very preterm and/or very low birth weight infants. Pediatr Res 1996; 39:142-145

25. Meijer WJ, Verloove-Vanhorick SP, Brand R, van den Brande JL. Transient hypothyroxinaemia associated with developmental delay in very preterm infants. Arch Dis Child 1992; 67:944-947

26. Hollanders JJ, Israels J, van der Pal SM, Verkerk PH, Rotteveel J, Finken MJ. No Association Between Transient Hypothyroxinemia of Prematurity and Neurodevelopmental Outcome in Young Adulthood. J Clin Endocrinol Metab 2015; 100:4648-4653

27. Verloove-Vanhorick SP, Verwey RA, Brand R, Gravenhorst JB, Keirse MJ, Ruys JH. Neonatal mortality risk in relation to gestational age and birthweight. Results of a national survey of preterm and very-low-birthweight infants in the Netherlands. Lancet 1986; 1:55-57

28. Wiersinga WM, Podoba J, Srbecky M, van VM, van Beeren HC, Platvoet-Ter Schiphorst MC. A survey of iodine intake and thyroid volume in Dutch schoolchildren: reference values in an iodine-sufficient area and the effect of puberty. Eur J Endocrinol 2001; 144:595-603

29. Chopra IJ. A radioimmunoassay for measurement of thyroxine in unextracted serum. J Clin Endocrinol Metab 1972; 34:938-947

30. Schopman WE, E; de Kock, HW; Rechsteiner, J; Tertoolen, JFW. Analytische vergelijking van de CHT-laboratoria. Rapporten over het 1e t/m 4e kwartaal 1983 1984;

31. Schopman W. Het L-thyroxine gehalte in ponsjes van bloedvlekken op PKU-CHT kaarten in de Nederlandse proefregio. Nulcleair Geneeskundig Bulletin 1979; 1:4-10

32. Achenbach TM. Manual for the young adult self-report and young adult behavioral checklist. Burlington: University of Vermont Department of Psychiatry.

33. van Wassenaer AG, Kok JH. Trials with thyroid hormone in preterm infants: clinical and neurodevelopmental effects. Semin Perinatol 2008; 32:423-430

34. Hack M, Youngstrom EA, Cartar L, Schluchter M, Taylor HG, Flannery D, Klein N, Borawski E. Behavioral outcomes and evidence of psychopathology among very low birth weight infants at age 20 years. Pediatrics 2004; 114:932-940

35. Hille ET, Dorrepaal C, Perenboom R, Gravenhorst JB, Brand R, Verloove-Vanhorick SP. Social lifestyle, risk-taking behavior, and psychopathology in young adults born very preterm or with a very low birthweight. J Pediatr 2008; 152:793-800, 800

36. Pakkila F, Mannisto T, Hartikainen AL, Ruokonen A, Surcel HM, Bloigu A, Vaarasmaki M, Jarvelin MR, Moilanen I, Suvanto E. Maternal and Child's Thyroid Function and Child's Intellect and Scholastic Performance. Thyroid 2015; 25:1363-1374

37. Sucksdorff M, Lehtonen L, Chudal R, Suominen A, Joelsson P, Gissler M, Sourander A. Preterm Birth and Poor Fetal Growth as Risk Factors of Attention-Deficit/ Hyperactivity Disorder. Pediatrics 2015; 136:e599-e608

38. Briet JM, van Wassenaer AG, Dekker FW, de Vijlder JJ, van Baar A, Kok JH. Neonatal thyroxine supplementation in very preterm children: developmental outcome evaluated at early school age. Pediatrics 2001; 107:712-718

39. van Wassenaer AG, Westera J, Houtzager BA, Kok JH. Ten-year follow-up of children born at <30 weeks' gestational age supplemented with thyroxine in the neonatal period in a randomized, controlled trial. Pediatrics 2005; 116:e613-e618

40. Saigal S, Rosenbaum PL, Feeny D, Burrows E, Furlong W, Stoskopf BL, Hoult L. Parental perspectives of the health status and health-related quality of life of teen-aged children who were extremely low birth weight and term controls. Pediatrics 2000; 105:569-574

41. Mollee TS, Finken MJ, van Weissenbruch MM, Rotteveel J. Normal thyroid function in young adults who were born very preterm. J Pediatr Endocrinol Metab 2011; 24:887-891

42. Hille ET, Elbertse L, Gravenhorst JB, Brand R, Verloove-Vanhorick SP. Nonresponse bias in a follow-up study of 19-year-old adolescents born as preterm infants. Pediatrics 2005; 116:e662-e666

43. Medici M, de Rijke YB, Peeters RP, Visser W, SM dMK-S, Jaddoe VV, Hofman A, Hooijkaas H, Steegers EA, Tiemeier H, Bongers-Schokking JJ, Visser TJ. Maternal early pregnancy and newborn thyroid hormone parameters: the Generation R study. J Clin Endocrinol Metab 2012; 97:646-652

Supplementary Table 1: Risk of problem behavior across the entire range of T4 SDS

	Unadjusted	*P* value	Adjusted (1)	*P* value	Adjusted (2)	*P* value
YASR						
Syndrome scales						
Anxious	0.7 (0.4 - 1.1)	0.140	0.6 (0.4 - 1.1)	0.077	0.7 (0.4 - 1.4)	0.356
Withdrawn	1.0 (0.7 - 1.4)	1.000	0.9 (0.7 - 1.3)	0.732	0.9 (0.6 - 1.4)	0.680
Somatic complaints	1.3 (0.8 - 1.9)	0.269	1.3 (0.9 - 2.0)	0.165	1.2 (0.7 - 1.9)	0.488
Thought problems	1.1 (0.7 - 1.7)	0.664	1.0 (0.7 - 1.6)	0.881	1.0 (0.6 - 1.7)	0.967
Attention problems	0.9 (0.6 - 1.4)	0.572	0.8 (0.5 - 1.3)	0.433	0.9 (0.5 - 1.5)	0.590
Intrusive behavior	0.8 (0.3 - 2.1)	0.669	0.8 (0.3 - 2.2)	0.689	0.4 (0.1 - 1.5)	0.169
Aggressive behavior	1.7 (0.96 - 3.0)	0.068	1.8 (0.95 - 3.3)	0.070	1.6 (0.7 - 3.3)	0.231
Delinquent behavior	1.6 (0.9 - 2.9)	0.138	1.3 (0.7 - 2.6)	0.442	1.3 (0.6 - 3.1)	0.508
Problem scales						
Internal problems	0.9 (0.6 - 1.2)	0.310	0.8 (0.6 - 1.1)	0.240	0.7 (0.5 - 1.1)	0.180
External problems	1.2 (0.9 - 1.8)	0.263	1.1 (0.8 - 1.6)	0.584	0.8 (0.5 - 1.4)	0.485
Total problems	0.8 (0.6 - 1.1)	0.234	0.8 (0.6 - 1.1)	0.175	0.7 (0.5 - 1.01)	0.056
YABCL						
Syndrome scales						
Anxious	0.9 (0.6 - 1.2)	0.428	0.9 (0.6 - 1.2)	0.409	0.8 (0.5 - 1.3)	0.318
Withdrawn	0.9 (0.6 - 1.3)	0.431	0.9 (0.6 - 1.3)	0.456	0.9 (0.5 - 1.3)	0.495
Somatic complaints	1.1 (0.7 - 1.6)	0.673	1.0 (0.7 - 1.5)	0.827	1.0 (0.6 - 1.6)	0.968
Thought problems	0.8 (0.6 - 1.2)	0.249	0.8 (0.5 - 1.2)	0.241	0.9 (0.6 - 1.4)	0.579
Attention problems	0.8 (0.6 - 1.2)	0.262	0.8 (0.5 - 1.1)	0.148	0.7 (0.4 - 1.1)	0.113
Intrusive behavior	1.2 (0.7 - 2.0)	0.572	1.1 (0.7 - 2.0)	0.617	0.7 (0.4 - 1.4)	0.371
Aggressive behavior	0.9 (0.6 - 1.4)	0.595	0.9 (0.5 - 1.4)	0.565	0.8 (0.4 - 1.4)	0.381
Delinquent behavior	1.2 (0.5 - 2.8)	0.635	1.2 (0.5 - 2.8)	0.643	1.2 (0.4 - 4.1)	0.760
Problem scales						
Internal problems	0.8 (0.7 - 1.04)	0.105	0.8 (0.6 - 1.1)	0.125	0.8 (0.6 - 1.1)	0.110
External problems	0.9 (0.7 - 1.2)	0.563	0.9 (0.7 - 1.2)	0.462	0.8 (0.6 - 1.2)	0.241
Total problems	0.8 (0.6 - 1.1)	0.150	0.8 (0.6 - 1.1)	0.164	0.7 (0.5 - 0.9)	0.038*

Values represent OR (95% CI).
Adjusted (1): gender, SES, ethnicity and parity
Adjusted (2): adjusted (1) + gestational age, SGA-status and neonatal illnesses (IRDS, IVH and sepsis)
* *P* <0.05

Part 4

Early-life growth and neurodevelopment

13

Growth pattern and final height of very preterm vs. very low birth weight infants

Jonneke J. Hollanders,
Sylvia M. van der Pal,
Paula van Dommelen,
Joost Rotteveel,
Martijn J.J. Finken

ABSTRACT

Background

Both very preterm (VP; i.e., gestational age <32 weeks) and very low birth weight (VLBW; i.e., birth weight <1,500 g) are used as inclusion criteria by studies on preterm birth. We aimed to quantify the impact of these entities on postnatal growth until final height.

Methods

Subjects born VP and/or with VLBW from the Project On Preterm and Small-for-gestational-age infants cohort were classified as follows: (1) VP+/VLBW+ (n=495), (2) VP+/VLBW- (n=207), or (3) VP-/VLBW+ (n=296) infants. Anthropometric data were collected at birth, 3, 6, 12, and 24 months' corrected age, and at 5 and 19 years. At 19 years, 590/998 (59%) of the subjects enrolled in 1983 were followed up.

Results

Birth size was smallest in the VP-/VLBW+ group compared with the VP+/VLBW+ and VP+/VLBW- groups. During childhood, length, weight, and head circumference SD scores increased in the VP-/VLBW+ group, whereas SD scores in the VP+/VLBW+ and VP+/VLBW- groups either remained stable or decreased. Despite catch-up growth, VP-/VLBW+ infants remained the shortest and lightest at age 19.

Conclusion

Classification on the basis of VP and VLBW impacts growth, causing different growth patterns for infants born VP+/VLBW+, VP+/VLBW-, or VP-/VLBW+. For future studies, we recommend, at least for industrialized countries, including preterm infants based on gestational age.

INTRODUCTION

Infants born very preterm (i.e., VP; <32 weeks of gestation) and/or with very low birth weight (i.e., VLBW; <1,500 grams) require admission to a Neonatal Intensive Care Unit (NICU). Early postnatal growth in such infants is often characterized by extrauterine growth retardation (EUGR),[1] caused by a combination of factors, including acute illnesses, glucocorticoid therapy and feeding difficulties. Although the majority of these infants exhibit late postnatal catch-up growth, they often remain short and thin during childhood and adolescence.[2] Moreover, postnatal growth in VP and/or VLBW infants has been associated with a variety of short- and long-term outcomes, such as cognitive functioning,[3,4] motor performance,[4] and body composition.[5]

However, most of the evidence on the long-term consequences of preterm birth on growth comes from studies in infants with VLBW,[6-11] which can be attributed to prematurity, intrauterine growth retardation (IUGR), or both. Before the 1980s, prematurity was often based on birth weight in lieu of reliable ways to estimate gestational age. After the widespread application of ultrasound as a tool to accurately measure pregnancy duration,[12] the majority of studies on preterm infants still used birth weight as an estimate of the degree of prematurity.[6-11]

Although it is reasonable to assume that being born VP or with VLBW has different impacts on outcomes, surprisingly, this has been quantified only once before.[13] In comparing VLBW with VP infants, VLBW was associated with an overrepresentation of SGA births and with lower numbers experiencing neonatal morbidities. However, this study did not provide long-term follow-up and, therefore, it is not known whether the differences between these groups could be extended with growth after hospital discharge. Findings from studies in cohorts of infants born either VP or with VLBW that have provided follow-up until adulthood are incomparable because of the different standards of care at the time of birth.[5,7,8,14] The past 30 years were characterized by major changes including widespread use of antenatal glucocorticoid therapy, introduction of synthetic surfactant and more aggressive feeding policies. Such changes have resulted in improved chances for survival. Therefore, caution must be exercised in the comparison of populations born either VP or with VLBW from different birth eras. Consequently, the long-term impact of being born VP or with VLBW can only be studied within the same birth cohort.

Therefore, in this study, we aimed to quantify the impact of being born VP and/or with a VLBW on growth until adulthood. For this purpose, we used the data of the Dutch Project On Preterm and Small-for-gestational-age infants (POPS) cohort, which is, to our knowledge, the only study to date that followed up children born both VP and/or with a VLBW into young adulthood.

METHODS

Study population

The POPS cohort comprised 94% (n=1,334) of the infants born alive in the Netherlands in 1983 with a gestational age <32 weeks (VP) and/or with a birth weight <1,500 grams (VLBW). Gestational age was based on last menstrual period, pregnancy testing, and/or ultrasound. Subjects were excluded if they died during their hospital stay (n=340). Thus, 998 subjects were included in this study. The inclusion criteria enabled us to compare groups of: (1) VP+/ VLBW+ infants (n=495), (2) VP+/VLBW- infants (n=207), and (3) VP-/ VLBW+ infants (n=296) (Figure 1). Approval of the medical ethical committees of all participating centers was obtained.
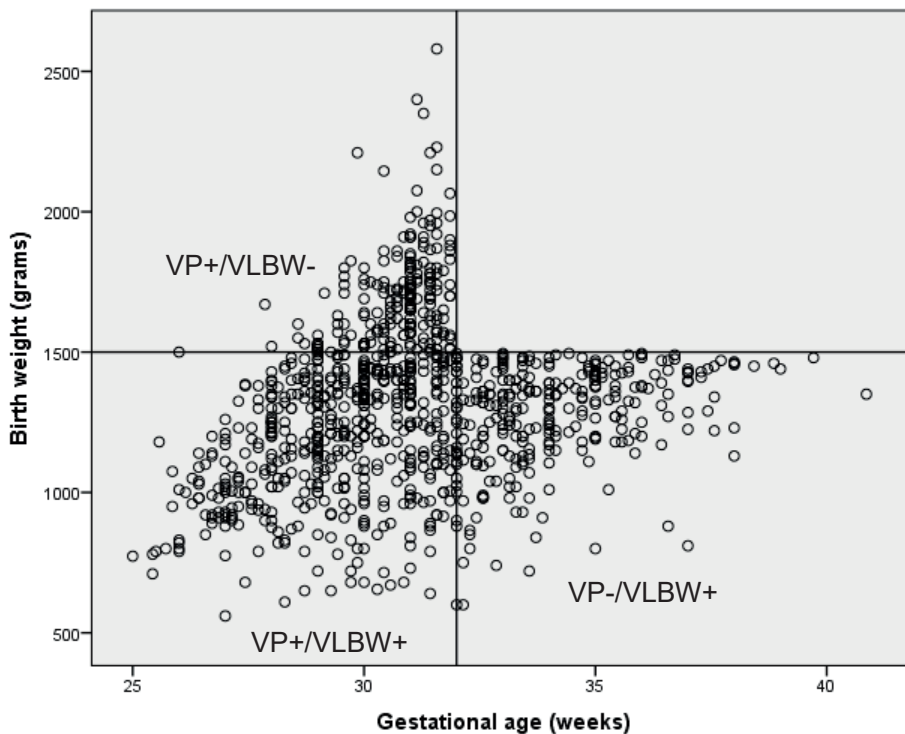
Figure 1: Distribution of our study population, with the three different groups (i.e., VP+/VLBW+, VP+/VLBW- and VP-/VLBW+) indicated

Growth assessment

Subjects underwent growth assessment at birth, at 3, 6, 12 and 24 months of corrected age, and again at the chronological ages of 5 and 19 years, when all participants had reached final height. Follow-up was done by trained research nurses and/or physicians

according to standardized procedures. At ages 3 months to 2 years, anthropometric data were collected at outpatient clinics. At age 5, research staff made house visits, whereas at age 19 follow-up took place at 1 of the 10 involved research centers. Until the age of 2, length was measured in supine position to the nearest 1 cm. From age 5 onward, standing height was measured to the nearest 1 mm. Weight was measured to the nearest 5 grams at birth, and during the follow-up visits to the nearest 0.1 kg on a balance scale. Head circumference (HC) was measured to the nearest 1 cm up to the age of 5.

Standard deviation scores (SDSs) for length/height, weight, and HC were calculated.[15,16] Subjects with a birth weight and/or length of less than -2 SD were classified as SGA. Body mass index (BMI) was calculated as (weight (kg) / (length (m))^2) and converted to SDS.[17]

Statistical analysis

SDS of length/height, weight, BMI, and HC were compared between groups using a generalized estimating equation (GEE). The different measurement points were used as an interaction term with the different groups. A P value of ≤0.05 was considered as a significant difference, and 95% confidence intervals (95% CIs) were calculated as described by Figueiras, et al.[18]

RESULTS

Perinatal characteristics of participants were significantly different between the VP+/ VLBW+, VP+/VLBW- and VP-/VLBW+ groups (Table 1). In general, the VP-/VLBW+ group showed significantly less neonatal morbidity and had a better Apgar score after 5 minutes, whereas they were significantly more often SGA and born to mothers with (pre-existent) hypertension and/or who smoked during pregnancy. The VP+/VLBW+ group had significantly more neonatal morbidity, a longer hospital stay, and more days on ventilation compared with the other two groups. The VP+/VLBW- group had the shortest hospital stay. However, ethnicity, marital status, and socio-economic status (SES) were not significantly different between the three groups, and target height SD and maternal height also did not differ between the three groups.

The response rate was different for the several follow-up visits (Figure 2). At age 19, 59% of the infants included for this study were followed up; for the other visits, the response rate was adequate. At age 19, responders were significantly different compared with the non-responders with regard to gender distribution, target height, pre-existent hypertension, ethnicity, and SES, whereas all other perinatal characteristics as well as the distribution of the subjects between the three groups did not differ (Table 2).

Table 1: Perinatal characteristics of the surviving subjects in the three groups

	VP+/VLBW+ n=495	VP+/VLBW- n=207	VP-/VLBW+ n=296	Overall *P* value
Male	250 (50.5)	134 (64.7)	139 (47.0)	<0.001*‡
Birth weight (grams)	1,173±214	1,705±178	1,276±178	<0.001*†‡
Gestational age (weeks)	29.3±1.6	30.8±0.9	34.1±1.7	<0.001*†‡
PROM	115 (23.2)	56 (27.1)	12 (4.1)	<0.001†‡
Born via caesarian section	194 (39.2)	52 (25.1)	223 (75.3)	<0.001*†‡
Apgar score >7 after 5 minutes	384 (77.6)	175 (84.5)	264 (89.2)	<0.001*†
Duration of hospital stay (d)	79±32	50±20	62±33	<0.001*†‡
Days of ventilation (d)	7.5±11.0	3.2±5.1	1.4±7.5	<0.001*†‡
IRDS	254 (51.3)	99 (47.8)	36 (12.2)	<0.001†‡
Sepsis	196 (39.8)	60 (29.0)	77 (26.0)	<0.001*†
IVH	129 (26.1)	28 (13.5)	18 (6.1)	<0.001*†‡
NEC	31 (6.3)	9 (4.3)	17 (5.7)	0.608
SGA-status	48 (9.7)	2 (1.0)	235 (79.4)	<0.001*†‡
Target height (SD)[a]	-0.1±0.8	0.0±0.8	-0.1±0.9	0.203
Maternal height (cm)	166±6.1	167±6.5	166±7.0	0.306
Pre-existent hypertension	16 (3.2)	3 (1.5)	25 (8.6)	<0.001†‡
Hypertension during pregnancy	104 (21.0)	11 (5.3)	152 (51.4)	<0.001*†‡
Smoking during pregnancy	145 (29.3)	55 (26.6)	112 (37.8)	0.023†‡
Caucasian maternal ethnicity	417 (84.4)	177 (87.2)	252 (85.7)	0.628
Married (parents)	438 (88.8)	192 (92.8)	249 (84.1)	0.114
Low parental SES	214 (45.1)	75 (37.9)	123 (42.4)	0.228

Values represent mean±SD or n (%). Continuous variables were compared with the one-way ANOVA test when comparing the three groups, and the independent t-test when comparing two groups. Dichotomous variables were compared with the Chi square test.
* *P* value <0.05 for VP+/VLBW+ vs. VP+/VLBW-
† *P* value <0.05 for VP+/VLBW+ vs. VP-/VLBW+
‡ *P* value <0.05 for VP+VLBW- vs. VP-/VLBW+
VP: very preterm; VLBW: very low birth weight; PROM: premature rupture of membranes; IRDS: infants respiratory distress syndrome; IVH: intraventricular hemorrhage; NEC: necrotizing enterocolitis; SGA: small-for-gestational-age
[a] Girls: target height = [(height father in cm + height mother in cm – 13) / 2] + 4.5; boys: target height = [(height father in cm + height mother in cm + 13) / 2] + 4.5

Length/height was significantly different between groups at all ages, except between the VP-/VLBW+ and VP+/VLBW+ group at age 5 (Figure 3A). The VP+/VLBW- group was the tallest at all ages, whereas the VP-/VLBW+ group remained the shortest.

Weight was significantly different between groups at all ages (Figure 3B), except for age 19 between the VP-/VLBW+ and the VP+/VLBW+ groups. Once again, the VP+/VLBW- group consistently had the highest weight, whereas the VP-/VLBW+ group had the lowest.
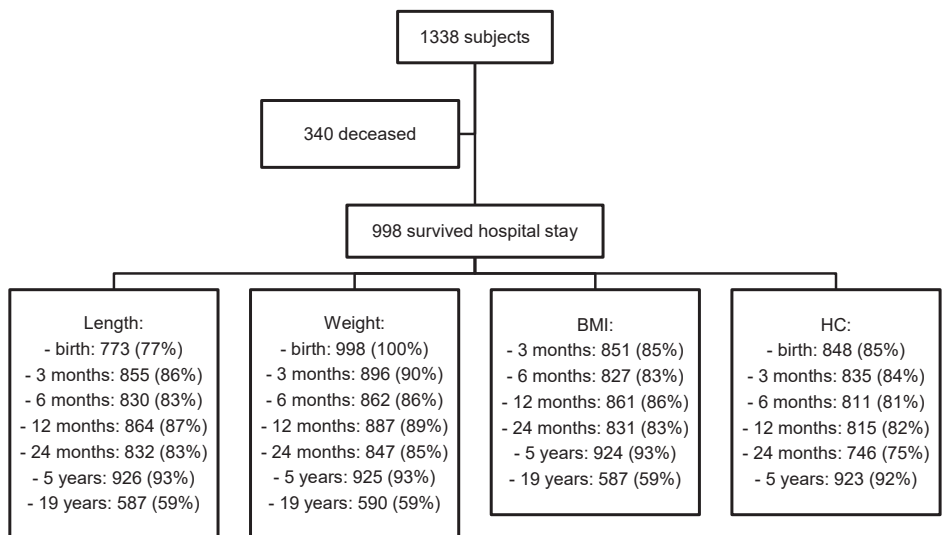
Figure 2: Follow-up response for the different growth parameters at all follow-up visits
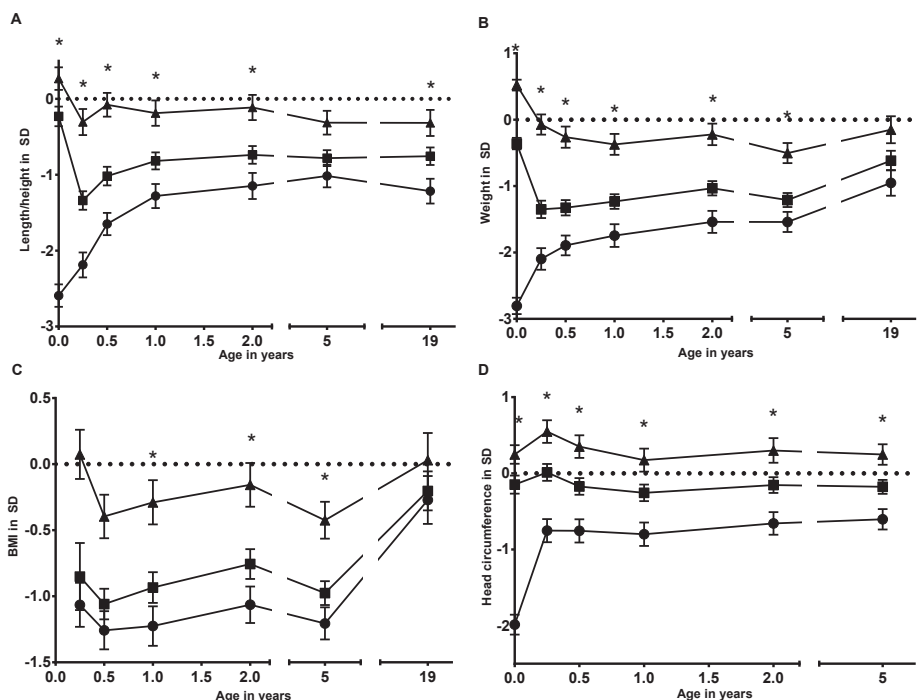

Figure 3: Growth for length/height (A), weight (B), BMI (C) and head circumference (D), expressed as SD scores, for all three groups
Squares=VP+/VLBW+, triangles=VP+/VLBW-, circles=VP-/VLBW+; the error bars represent 95%CI's.
* *P* value <0.05 between all three group, analyzed per measurement point

Table 2: Perinatal characteristics of the responders vs. non-responders at age 19

	Responders at age 19 n=590	Non-responders at age 19 n=408	P value
Male	265 (44.9)	258 (63.2)	<0.001
Birth weight (grams)	1301±299	1331±261	0.095
Gestational age (weeks)	31.0±2.5	31.2±2.6	0.265
PROM	105 (17.8)	78 (19.1)	0.596
Born via caesarian section	289 (49.0)	180 (44.1)	0.130
Apgar score >7 after 5 minutes	492 (83.4)	331 (81.1)	0.618
Duration of hospital stay (d)	67.3±30.3	69.9±35.4	0.464
Days of ventilation (d)	4.8±9.5	4.7±9.4	0.840
IRDS	232 (39.3)	157 (38.5)	0.789
Sepsis	193 (32.8)	140 (34.4)	0.592
IVH	95 (16.1)	80 (19.6)	0.152
NEC	36 (6.1)	21 (5.1)	0.523
SGA-status	170 (28.8)	115 (28.3)	0.848
Target height (SD)[a]	0.0±0.8	-0.2±0.9	0.001
Maternal height (cm)	167±6.3	165±6.6)	0.006
Pre-existent hypertension	33 (5.6)	11 (2.7)	0.030
Hypertension during pregnancy	168 (28.5)	99 (24.3)	0.140
Smoking during pregnancy	178 (30.2)	134 (32.8)	0.644
Caucasian maternal ethnicity	519 (88.9)	327 (80.3)	<0.001
Married (parents)	529 (89.7)	350 (86.2)	0.324
Low parental SES	208 (35.5)	204 (54.1)	<0.001
VP/VLBW status	302 (51.2)	193 (47.3)	0.392
- VP+/VLBW+	115 (19.5)	92 (22.5)	
- VP+/VLBW-	173 (29.3)	123 (30.1)	
- VP-/VLBW+			

Values represent mean±SD or n (%). Continuous variables were compared with the one-way ANOVA test when comparing the three groups, and the independent t-test when comparing two groups. Dichotomous variables were compared with the Chi square test.
* P value <0.05 for VP+/VLBW+ vs. VP+/VLBW-
[†] P value <0.05 for VP+/VLBW+ vs. VP-/VLBW+
[‡] P value <0.05 for VP+VLBW- vs. VP-/VLBW+
VP: very preterm; VLBW: very low birth weight; PROM: premature rupture of membranes; IRDS: infants respiratory distress syndrome; IVH: intraventricular hemorrhage; NEC: necrotizing enterocolitis; SGA: small-for-gestational-age
[a] Girls: target height = [(height father in cm + height mother in cm – 13) / 2] + 4.5; boys: target height = [(height father in cm + height mother in cm + 13) / 2] + 4.5

Table 3: Changes in SD scores within the groups for length/height, weight, BMI, and HC over time

	VP+/VLBW+		VP+/VLBW-		VP-/VLBW+	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Length/height						
Birth - 19 years	*-0.5 (-0.7 to -0.4)*	*<0.001*	*-0.6 (-0.8 to -0.4)*	*<0.001*	*1.4 (1.2 to 1.6)*	*<0.001*
Birth - 3 months	-1.1 (-1.3 to -1.0)	<0.001	-0.6 (-0.8 to -0.4)	<0.001	0.4 (0.2 to 0.6)	<0.001
3 months - 1 year	0.5 (0.4 to 0.6)	<0.001	0.1 (-0.1 to 0.3)	0.20	0.9 (0.8 to 1.0)	<0.001
1 year - 5 years	0.0 (-0.1 to 0.1)	0.43	-0.1 (-0.3 to 0.0)	0.09	0.3 (0.2 to 0.4)	<0.001
5 years - 19 years	0.0 (-0.1 to 0.1)	0.57	0.0 (-0.1 to 0.1)	0.97	-0.2 (-0.3 to -0.1)	0.005
Weight						
Birth - 19 years	*-0.3 (-0.4 to -0.1)*	*0.001*	*-0.7 (-0.9 to -0.5)*	*<0.001*	*1.9 (1.6 to 2.1)*	*<0.001*
Birth - 3 months	-1.0 (-1.1 to -0.9)	<0.001	-0.6 (-0.7 to -0.4)	<0.001	0.7 (0.5 to 0.9)	<0.001
3 months - 1 year	0.1 (0.0 to 0.2)	0.04	-0.3 (-0.5 to -0.1)	<0.001	0.4 (0.2 to 0.5)	<0.001
1 year - 5 years	0.0 (-0.1 to 0.1)	0.59	-0.1 (-0.3 to 0.0)	0.05	0.2 (0.1 to 0.3)	0.001
5 years - 19 years	0.6 (0.5 to 0.7)	<0.001	0.4 (0.2 to 0.5)	<0.001	0.6 (0.4 to 0.8)	<0.001
BMI						
3 months - 19 years	*0.6 (0.3 to 1.0)*	*<0.001*	*0.0 (-0.3 to 0.2)*	*0.72*	*0.8 (0.6 to 1.0)*	*<0.001*
3 months - 1 year	-0.1 (-0.3 to 0.2)	0.52	-0.4 (-0.6 to -0.2)	<0.001	-0.2 (-0.3 to 0.0)	0.06
1 year - 5 years	0.0 (-0.1 to 0.1)	0.43	-0.1 (-0.3 to 0.0)	0.10	0.0 (-0.1 to 0.2)	0.78
5 years - 19 years	0.8 (0.6 to 0.9)	<0.001	0.5 (0.3 to 0.7)	<0.001	0.9 (0.8 to 1.1)	<0.001
HC						
Birth - 5 years	*0.0 (-0.2 to 0.1)*	*0.66*	*0.0 (-0.2 to 0.2)*	*0.99*	*1.4 (1.2 to 1.5)*	*<0.001*
Birth - 3 months	0.2 (0.0 to 0.3)	0.02	0.3 (0.1 to 0.5)	<0.001	1.2 (1.1 to 1.4)	<0.001
3 months - 1 year	-0.3 (-0.4 to -0.2)	<0.001	-0.4 (-0.5 to -0.3)	<0.001	0.0 (-0.1 to 0.0)	0.32
1 year - 5 years	0.1 (0.0 to 0.2)	0.04	0.1 (0.0 to 0.2)	0.16	0.2 (0.1 to 0.3)	<0.001

Values represent beta (95%CI), as compared with the GEE
VP: very preterm; VLBW: very low birth weight; BMI: body mass index; HC: head circumference

BMI in the VP+/VLBW- group was significantly higher at all ages compared with the VP+/VLBW+ and VP-/VLBW+ groups, except for age 19 (Figure 3C). The BMI in the VP+/VLBW+ and VP-/VLBW+ groups was significantly different at ages 1, 2 and 5.

HC was significantly different between groups at all ages (Figure 3D). The VP+/VLBW- group consistently had the largest HC, whereas the VP-/VLBW+ group had the smallest.

Table 3 shows the changes in SDS within the groups over time. Between birth and age 19, length/height SDS significantly increased in the VP-/VLBW+ group, while the VP+/VLBW+ and VP+/VLBW- groups showed a decrease in length/height SDS. A similar pattern was observed for weight SDS between birth and age 19. The greatest changes in SDS took place between birth and 3 months of corrected age, and between ages 5 and 19. BMI did not change in the VP+/VLBW- group between 3 months and age 19, but there was a significant increase in SDS in the VP+/VLBW+ and VP-/VLBW+ groups

between ages 5 and 19. Between birth and age 5, HC SDS increased significantly in the VP-/VLBW+ group, whereas there was no change in SDS in the VP+/VLBW+ and the VP+/VLBW- groups. The greatest SDS change took place between birth and 3 months.

DISCUSSION

In this study, we found significantly different growth patterns for length/height, weight, BMI, and HC between VP+/VLBW+, VP+/VLBW- and VP-/VLBW+ infants. This indicates that the terms VP and VLBW describe two different entities, which impact differently on growth.

We observed that VP-/VLBW+ infants were severely growth restricted at birth, as evidenced by an SGA rate of 79.4%, and exhibited rapid catch-up growth postnatally. However, they remained shorter and lighter, and with a smaller HC, as compared with the other two groups. In contrast, the VP+/VLBW+ and VP+/VLBW- groups showed sharp decreases in length/height SDS and weight SDS between birth and 3 months, which could most likely be attributed to the greater percentage of children with neonatal illnesses or to inadequate nutritional support for their degree of illness, although this was not measured in our study.[19] In a recent systematic review, poor early postnatal growth after preterm birth was associated with adverse neurodevelopment.[20] Conversely, yet another study found that rapid early postnatal growth in subjects born SGA was associated with cardiometabolic disease propensity.[21] HC SDS did not change for both groups. Nevertheless, despite this parallelism in growth patterns, the VP+/VLBW+ group remained below the means of the norm population for length/height, weight and HC. In contrast, the length/height, weight and HC of the VP+/VLBW- group remained close to the population reference mean.

At age 19 years, all three groups had a BMI that was comparable to the norm population, with significant increases in BMI SDS for the VP+/VLBW+ and VP-/VLBW+ groups between ages 5 and 19. This pattern has previously been associated with an increased risk of coronary events later in life.[22]

A previous study found significantly different early neonatal outcomes between VP and VLBW infants.[13] This was replicated in our study, in which we observed a greater percentage of neonatal morbidities in the VP+/VLBW+ and VP+/VLBW- groups. VP-/VLBW+ infants, in turn, were more often born SGA. In addition, they were more often born to mothers who had (pre-existent) hypertension or who smoked during pregnancy, which are known risk factors for IUGR [23]. This suggests that short-term outcomes are significantly different depending on the classification that is used. Our current findings indicate that the differences between these groups could be extended with growth up until adulthood, with the VP-/VLBW+ group, and to a lesser extent the VP+/VLBW+ group,

showing a growth pattern that has previously been associated with cardiometabolic disease.[21]

Because of the different neonatal outcomes and growth patterns found in our study between children born VP and/or with a VLBW, we argue that these terms cannot be used interchangeably. Therefore, findings from studies in children born with VLBW cannot be automatically extrapolated to children born VP, and vice versa.

For future studies on preterm infants, we suggest that researchers strive to use the same inclusion criteria, enabling comparisons between different cohorts. Notably, it is desirable to use inclusion criteria that result in representative samples of preterm infants. The definition of prematurity is based on pregnancy duration and not on birth weight, and with current technologies, pregnancy dating can be determined very accurately, especially in the first trimester.[24,25] Therefore, we would strongly recommend that, at least for industrialized countries, the inclusion of preterm subjects is based on gestational age rather than on birth weight.

Our study has several strengths and limitations. The major strengths of our study are its large sample size, long-term follow-up and completeness of data, including growth parameters.

Our study also has some limitations. First, the follow-up rate between ages 5 and 19 dropped substantially. A non-response analysis (Table 2) showed that non-response was associated with male gender, non-Caucasian ethnicity and a lower SES. However, aside from gender, these factors were not significantly different between the three study groups. Moreover, responders had a 0.2 SD higher target height SDS and their mothers were more often hypertensive compared with non-responders. However, perinatal characteristics, including SGA status, did not differ between response groups. Moreover, the distribution of subjects among the three study groups was not different between responders and non-responders. Therefore, although small differences were present, we think that our results at age 19 are not subject to attrition bias. Second, subjects in this cohort were born in 1983. Therefore, our data are not representative for the current generation of preterm infants,[26] as neonatal care has advanced significantly in the past three decades. However, preterm infants still show EUGR despite improved neonatal care,[27] and VP and/or VLBW subjects can therefore still be expected to have different growth patterns. Nevertheless, these growth patterns will presumably be more favorable than the growth patterns observed in this study. As IUGR pregnancies are nowadays strictly monitored and labor is induced if necessary[23], infants will be less likely severely growth restricted at birth, leading to lower numbers being VLBW but not VP. Moreover, because of improved nutritional strategies, which include earlier introduction and rapid increases in protein intake [28,29], as well as an increased awareness of the benefits of human breast milk [30], VP+/VLBW+ and VP+/VLBW- infants will likely show a smaller decrease in SDS, whereas VP-/VBW+ infants will show improved catch-up growth. Third, although post-
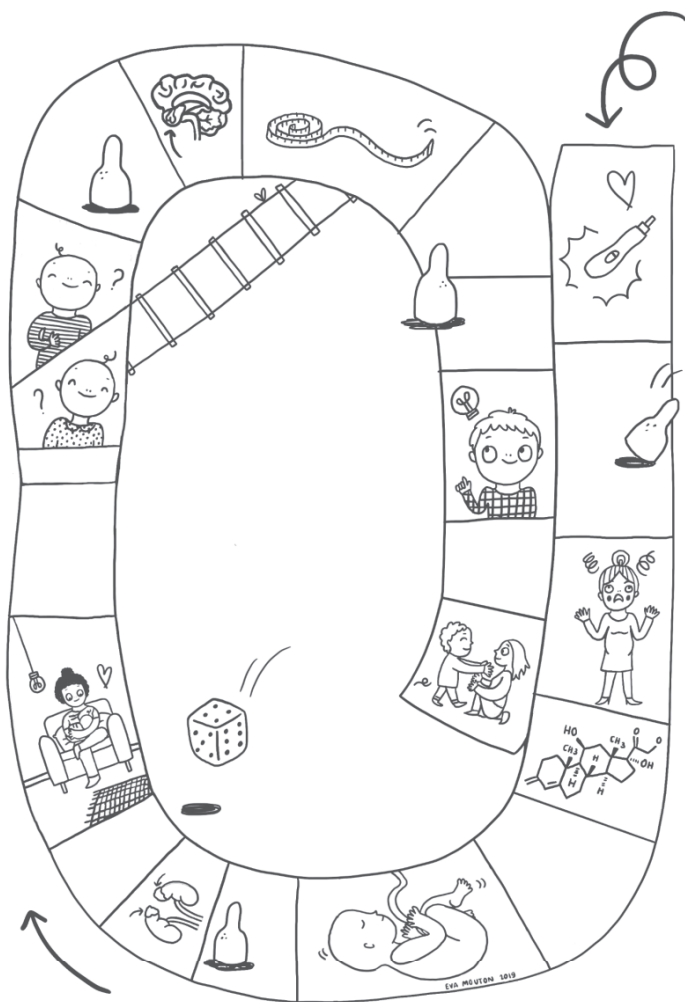
natal growth charts have been designed specifically for preterm infants, we did not use these reference data, as they apply to the current generation of NICU-treated infants. Instead, we have used reference charts that are applicable to Dutch children born in the 1980s.[16,17] Although these charts were not designed specifically for preterm infants, they do allow comparison between the study groups. Finally, the observed growth patterns, especially those of the VP-/VLBW+ group, could also be partly ascribed to regression to the mean. However, this phenomenon is difficult to quantify and to separate from catch-up growth. Nonetheless, despite possible regression to the mean, the three groups still show significantly different growth at age 19.

In conclusion, infants born VP+/VLBW+, VP+/VLBW- or VP-/VLBW+ appear to have significantly different growth patterns, and, therefore, the terms VP and VLBW cannot be used interchangeably. Because pregnancy dating can be reliably assessed nowadays, we recommend, at least for industrialized countries, to use gestational age instead of birth weight as the inclusion criterion for future studies in preterm infants.

REFERENCES

1. Wit JM, Finken MJ, Rijken M, de Zegher F. Preterm growth restraint: a paradigm that unifies intrauterine growth retardation and preterm extrauterine growth retardation and has implications for the small-for-gestational-age indication in growth hormone therapy. Pediatrics 2006; 117:e793-e795

2. Euser AM, de Wit CC, Finken MJ, Rijken M, Wit JM. Growth of preterm born children. Horm Res 2008; 70:319-328

3. Casey PH, Whiteside-Mansell L, Barrett K, Bradley RH, Gargus R. Impact of prenatal and/or postnatal growth problems in low birth weight preterm infants on school-age outcomes: an 8-year longitudinal evaluation. Pediatrics 2006; 118:1078-1086

4. Cooke RW, Foulder-Hughes L. Growth impairment in the very preterm and cognitive and motor performance at 7 years. Arch Dis Child 2003; 88:482-487

5. Euser AM, Finken MJ, Keijzer-Veen MG, Hille ET, Wit JM, Dekker FW. Associations between prenatal and infancy weight gain and BMI, fat mass, and fat distribution in young adulthood: a prospective cohort study in males and females born very preterm. Am J Clin Nutr 2005; 81:480-487

6. Doyle LW, Faber B, Callanan C, Ford GW, Davis NM. Extremely low birth weight and body size in early adulthood. Arch Dis Child 2004; 89:347-350

7. Ford GW, Doyle LW, Davis NM, Callanan C. Very low birth weight and growth into adolescence. Arch Pediatr Adolesc Med 2000; 154:778-784

8. Hack M, Schluchter M, Cartar L, Rahman M, Cuttler L, Borawski E. Growth of very low birth weight infants to age 20 years. Pediatrics 2003; 112:e30-e38

9. Latal-Hajnal B, von SK, Kovari H, Bucher HU, Largo RH. Postnatal growth in VLBW infants: significant association with neurodevelopmental outcome. J Pediatr 2003; 143:163-170

10. Lucas A, Morley R, Cole TJ, Gore SM, Lucas PJ, Crowle P, Pearse R, Boon AJ, Powell R. Early diet in preterm babies and developmental status at 18 months. Lancet 1990; 335:1477-1481

11. Saigal S, Stoskopf B, Streiner D, Paneth N, Pinelli J, Boyle M. Growth trajectories of extremely low birth weight infants from birth to young adulthood: a longitudinal, population-based study. Pediatr Res 2006; 60:751-758

12. Campbell S. The prediction of fetal maturity by ultrasonic measurement of the biparietal diameter. J Obstet Gynaecol Br Commonw 1969; 76:603-609

13. Lapeyre D, Klosowski S, Liska A, Zaoui C, Gremillet C, Truffert P. [Very preterm infant (< 32 weeks) vs very low birth weight newborns (1500 grammes): comparison of two cohorts]. Arch Pediatr 2004; 11:412-416

14. Brandt I, Sticker EJ, Gausche R, Lentze MJ. Catch-up growth of supine length/height of very low birth weight, small for gestational age preterm infants to adulthood. J Pediatr 2005; 147:662-668

15. Niklasson A, Ericson A, Fryer JG, Karlberg J, Lawrence C, Karlberg P. An update of the Swedish reference standards for weight, length and head circumference at birth for given gestational age (1977-1981). Acta Paediatr Scand 1991; 80:756-762

16. Fredriks AM, van BS, Burgmeijer RJ, Meulmeester JF, Beuker RJ, Brugman E, Roede MJ, Verloove-Vanhorick SP, Wit JM. Continuing positive secular growth change in The Netherlands 1955-1997. Pediatr Res 2000; 47:316-323

17. Fredriks AM, van BS, Wit JM, Verloove-Vanhorick SP. Body index measurements in 1996-7 compared with 1980. Arch Dis Child 2000; 82:107-112

18. Figueiras A, Domenech-Massons JM, Cadarso C. Regression models: calculating the confidence interval of effects in the presence of interactions. Stat Med 1998; 17:2099-2105

19. Embleton NE, Pang N, Cooke RJ. Postnatal malnutrition and growth retardation: an inevitable consequence of current recommendations in preterm infants? Pediatrics 2001; 107:270-273

20. Ong KK, Kennedy K, Castaneda-Gutierrez E, Forsyth S, Godfrey KM, Koletzko B, Latulippe ME, Ozanne SE, Rueda R, Schoemaker MH, van der Beek EM, van BS, Fewtrell M. Postnatal growth in preterm infants and later health outcomes: a systematic review. Acta Paediatr 2015; 104:974-986

21. Leunissen RW, Kerkhof GF, Stijnen T, Hokken-Koelega A. Timing and tempo of first-year rapid growth in relation to cardiovascular and metabolic risk profile in early adulthood. JAMA 2009; 301:2234-2242

22. Barker DJ, Osmond C, Forsen TJ, Kajantie E, Eriksson JG. Trajectories of growth among children who have coronary events as adults. N Engl J Med 2005; 353:1802-1809

23. American College of O, Gynecologists. ACOG Practice bulletin no. 134: fetal growth restriction. Obstet Gynecol 2013; 121:1122-1133

24. Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. Am J Obstet Gynecol 1996; 174:278-281

25. Neilson JP. Ultrasound for fetal assessment in early pregnancy. Cochrane Database Syst Rev 2000:CD000182

26. Stoelhorst GM, Rijken M, Martens SE, Brand R, den Ouden AL, Wit JM, Veen S. Changes in neo-natology: comparison of two cohorts of very preterm infants (gestational age <32 weeks): the Project On Preterm and Small for Gestational Age Infants 1983 and the Leiden Follow-Up Project on Prematurity 1996-1997. Pediatrics 2005; 115:396-405

27. Ehrenkranz RA, Younes N, Lemons JA, Fanaroff AA, Donovan EF, Wright LL, Katsikiotis V, Tyson JE, Oh W, Shankaran S, Bauer CR, Korones SB, Stoll BJ, Stevenson DK, Papile LA. Longitudinal growth of hospitalized very low birth weight infants. Pediatrics 1999; 104:280-289

28. Kashyap S, Schulze KF, Forsyth M, Dell RB, Ramakrishnan R, Heird WC. Growth, nutrient retention, and metabolic response of low-birth-weight infants fed supplemented and unsupplemented preterm human milk. Am J Clin Nutr 1990; 52:254-262

29. Miller J, Makrides M, Gibson RA, McPhee AJ, Stanford TE, Morris S, Ryan P, Collins CT. Effect of increasing protein content of human milk fortifier on growth in preterm infants born at <31 wk gestation: a randomized controlled trial. Am J Clin Nutr 2012; 95:648-655

30. Underwood MA. Human milk for the premature infant. Pediatr Clin North Am 2013; 60:189-207

14

Long-term neurodevelopmental and functional outcomes of infants born very preterm versus with a very low birth weight

Jonneke J. Hollanders,
Nina Schaéfer,
Sylvia M. van der Pal,
Jaap Oosterlaan,
Joost Rotteveel,
Martijn J.J. Finken

ABSTRACT

Background

Birth weight (BW) is often used as a proxy for gestational age (GA) in studies on preterm birth. Recent findings indicate that, in addition to perinatal outcomes, subjects born very preterm (VP; GA <32 weeks) differ from those with very-low-birth-weight (VLBW; BW <1,500 g) in postnatal growth up until final height.

Objective

To study whether neurodevelopmental and functional outcomes at age 19 are different between VP and/or VLBW subjects.

Methods

705 19-year-old subjects from the Project On Preterm and Small-for-gestational-age infants cohort were classified as (1) VP+/VLBW+ (n=354), (2) VP+/VLBW- (n=144) or (3) VP-/VLBW+ (n=207), and compared with regard to intelligence quotient (IQ) assessed with the Multicultural Capacity Test-Intermediate Level; neuromotor function using Touwen's examination of mild neurologic dysfunction; hearing loss; self- and parent reported behavioral and emotional functioning; educational achievement and occupation; and self-assessed health using the Health Utilities Index and the London Handicap Scale.

Results

VP+/VLBW- infants on average had 3.8 points higher IQ scores (95% confidence interval (CI): 0.5-7.1), a trend towards higher educational achievement, 3.3 dB better hearing (95%CI: 1.2-5.4), and less anxious behavior, attention problems and internalizing behavior compared to VP+/VLBW+ subjects. VP-/VLBW+ infants reported 1.8 increased odds (95%CI: 1.2-2.6) of poor health compared to VP+/VLBW+ subjects.

Conclusions

At age 19 years, subjects born VP+/VLBW+, VP+/VLBW- or VP-/VLBW+ have different neurodevelopmental and functional outcomes, although effect sizes are small. Hence, the terms VP and VLBW are not interchangeable. We recommend, at least for industrialized countries, to base inclusion for future studies in preterm populations on GA instead of BW.

INTRODUCTION

Being born very preterm (VP; i.e., gestational age <32 weeks) and/or with a very low birth weight (VLBW; i.e., birth weight <1,500 grams) requires admission to a neonatal intensive care unit (NICU). Both entities have previously been associated with neurodevelopmental and functional problems in adolescence.[1-9] Despite their close resemblance, in contrast to VP birth, VLBW can be attributed to prematurity, intrauterine growth restriction (IUGR), or both.

Results of studies in infants with VLBW are often extrapolated to preterm populations, and vice versa. However, previous research has shown that short-term outcomes are significantly different between children born VP and/or with VLBW, with more neonatal morbidities in VP infants, and more small-for-gestational-age (SGA) births among those with VLBW.[10] Moreover, long-term outcomes also appear to differ, as we recently found that VP and VLBW subjects have significantly different growth patterns and final height.[11] Subjects born VP without VLBW attained a height close to the population reference mean, whereas those with VLBW remained approximately 1 SD shorter despite initial catch-up growth. Whether such differences between VP and VLBW subjects also translate into different long-term neurodevelopmental and functional outcomes is unknown.

In the past three decades, NICU care has improved dramatically and survival rates of infants born VP and/or with VLBW have increased substantially.[12] Among the improvements are the widespread application of antenatal glucocorticoid therapy, the introduction of synthetic surfactant and a tendency towards more aggressive feeding strategies, although regional differences in the treatment of VP and VLBW infants do exist.[13] Therefore, the entities VP and VLBW can only be compared between populations that have received the same care.

We aimed to compare neurodevelopmental and functional outcomes in adolescence between subjects born VP and/or with VLBW, using the data from the Project on Preterm and Small-for-gestational-age infants (POPS) cohort. This cohort project is, to our knowledge, the only one which studied subjects born both VP and/or with VLBW into adolescence.

METHODS

Population

The POPS cohort included 94% (n=1,338) of the infants born alive in 1983 in The Netherlands who were VP and/or had a VLBW. We could therefore distinguish between: 1) VP+/VLBW+, 2) VP+/VLBW-, and 3) VP-/VLBW+ infants. Subjects were followed up throughout childhood until the age of 19 years, when the data for this study were collected. Ethical approval of all participating centers was obtained.

Neurodevelopmental outcomes

Cognitive functioning

Cognitive functioning was quantified with the intelligence quotient (IQ) as measured with the Multicultural Capacity Test (MCT)-intermediate level.[14] The MCT has been validated for individuals aged ≥16 years from different ethnic backgrounds with an education ranging from five years of secondary school to university level. It assesses verbal and numerical intelligence, spatial visualization, speech fluency, memory, reasoning, and speed of perception. Four subscales (linguistic capacity, mathematical capacity, logical reasoning, and spatial visualization) and a total score can be derived. Normative scores were expressed on a scale with a mean of 100 and a standard deviation (SD) of 15, based on the Dutch norm population.

Neuromotor function

Neuromotor function was assessed with the revised version of Touwen's examination of minor neurologic dysfunction.[15,16] It examines 5 subcategories (hand function, quality of walking, coordination, posture, and passive muscle tone), and comprises 34 items, which are scored on a 3-point scale where 2="optimal performance", 1="slightly reduced performance" and 0="poor performance". Total scores range between 0 and 68.

Hearing

Hearing was assessed with pure-tone audiometry with a hand-held audiometer for each ear separately. Auditory sensitivity was determined as the mean of the threshold levels at 500, 1,000, 2,000 and 4,000 Hz. Hearing loss in the best and worst ear was recorded.

Behavioral and emotional functioning

Behavior was studied with the self-reported Young Adult Self Report (YASR), and the parent/caretaker-reported Young Adult Behavior Checklist (YABCL). Both questionnaires were developed by Achenbach, and provide standardized scores on behavior, feelings, thoughts and competences in people aged 18 to 30 years.[17] The YASR contains 130 items, and the YABCL contains 109 items. Informants are required to rate items pertaining to the past six months, scored as 0 = "not true", 1 = "sometimes true", and 2 = "very often or often true". Eight syndrome scales can be derived: Anxious/Depressed, Withdrawn, Somatic complaints, Thought problems, Attention problems, Intrusive behavior, Aggressive behavior and Delinquent behavior. In addition, 3 problem scales can be calculated. "Internalizing behavior" is the sum of the syndrome scales Anxious/Depressed and Withdrawn. Aggressive behavior, Delinquent behavior and Intrusive behavior comprise the problem scale "Externalizing behavior"; and the "Total problems scale" is the sum of all individual items.

Functional outcomes

Educational achievement

A self-report was used to assess past and current education. Responses were coded according to the highest level of education achieved or currently enrolled, using a revised version of The Netherlands Central Bureau of Statistics (CBS) classification:[18] no/ primary education or special education (level 0), preparatory vocational education (level 1), intermediate vocational education or higher general secondary education (level 2), and higher vocational education, pre-university secondary education or university (level 3). For some participants, responses allowed multiple codings for current education. In such cases, best and worst case coding was used, coded by two assessors. Consensus about discrepancies was reached through discussion. Both worst- and best-case classifications were analyzed.

Occupation

Participants also provided details on their current occupation through self-report. Participation was coded as follows: no job or education (severe problem); part-time job <16 hours/week with no education, or part-time education without a job (moderate problem); part-time job 16-32 hours/week, or part-time education with a job <16 hours/ week (mild problem); and full-time education, full-time job >32 hours/week, or part-time education with a job 16-31 hours/week (no problem).

Seventeen subjects did not correctly fill in the questionnaire, and their data were therefore excluded.

Health status

The Health Utilities Index Mark 3 (HUI3) was used to determine health status and health-related quality of life. The HUI3 consists of 8 attributes, focusing on functional capacity: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. All attributes have 5 or 6 levels,[19] which were dichotomized as: levels 1 and 2 = "no problem", and level 3 and higher = "moderate to severe problem".[4] Subsequently, dichotomized attributes were combined as: 0 attributes affected (no problem), 1-2 attributes affected (mild problem), 3-4 attributes affected (moderate problem), or ≥5 attributes affected (severe problem).

Perceived health

The London Handicap Scale (LHS) was used to measure perceived health. It measures disadvantages for six dimensions on a 6-point hierarchical scale: mobility, physical independence (self-care), occupation (daily activities), social integration, orientation, and economic self-sufficiency.[20] Coding of responses on the LHS was identical to the method used for the HUI3.

Statistics

Differences in functional outcomes, activities and participation across the three groups were analyzed by multivariate linear or ordinal regression, depending on the measurement level of the outcome variable. Results were expressed as β (95% confidence interval (CI)) for linear regression, or odds ratio (OR) (95%CI) for ordinal regression. Next, analyses were adjusted for: 1) gender, socio-economic status and ethnicity (model 1); and 2) model 1 plus neonatal morbidities (infants respiratory distress syndrome, intraventricular hemorrhage, and sepsis) (model 2). These confounders were selected based on the literature or on differences in baseline characteristics between the 3 groups (Table 1).

For measures yielding multiple outcomes (MCT, Touwen's examination of minor neurologic dysfunction, YASR, and YABCL), α was adjusted to 0.01 to reduce the risk for type 1 errors. For the other outcomes, a *P*-value of <0.05 was considered significant.

For all analyses, the VP+/VLBW+ group was used as the reference group.

Table 1: Perinatal characteristics of the three groups and the nonresponders

	VP+/ VLBW+ n=354	VP+/ VLBW- n=144	VP-/ VLBW+ n=207	Overall P value	Nonresponders n=254	P value
Male	154 (43.5)	85 (59.0)	89 (43.0)	0.003[a,b]	169 (66.5)	<0.001
Birth weight (grams)	1161±211	1721±196	1275±175	<0.001[a,b,c]	1327±256	0.387
Gestational age (weeks)	29.3±1.5	30.7±1.0	34.0±1.6	<0.001[a,b,c]	31.2±2.7	0.352
PROM	82 (23.2)	37 (25.7)	8 (3.9)	<0.001[b,c]	47 (18.5)	0.862
Born via caesarian section	146 (41.2)	39 (27.1)	164 (79.2)	<0.001[a,b,c]	107 (42.1)	0.044
Apgar score >7 after 5 minutes	279 (78.8)	125 (86.8)	185 (89.4)	0.003[a,b]	209 (82.3)	0.896
Duration of hospital stay (days)	79±31	48±15	59±25	<0.001[a,b,c]	67±30	0.996
Days of ventilation (days)	7.3±10.0	2.8±4.5	1.7±8.9	<0.001[a,b]	4.8±10.3	0.887
IRDS	181 (51.1)	66 (45.8)	29 (14.0)	<0.001[b,c]	97 (38.2)	0.788
Sepsis	141 (40.1)	35 (24.3)	50 (24.2)	<0.001[a,b]	95 (37.4)	0.129
IVH	91 (25.7)	22 (15.3)	13 (6.3)	<0.001[a,b,c]	39 (15.4)	0.362
NEC	24 (6.8)	8 (5.6)	11 (5.3)	0.747	12 (4.7)	0.419
Small-for-gestational-age	36 (10.2)	2 (1.4)	159 (76.8)	<0.001[a,b,c]	75 (29.5)	0.631

Values represent mean±SD or n (%). Continuous variables were compared with the one-way ANOVA test when comparing the three groups, and the independent t-test when comparing two groups. Dichotomous variables were compared with the Chi square test.
[a] *P* value <0.05 for VP+/VLBW+ vs. VP+/VLBW-
[b] *P* value <0.05 for VP+/VLBW+ vs. VP-/VLBW+
[c] *P* value <0.05 for VP+VLBW- vs. VP-/VLBW+
VP: very preterm; VLBW: very low birth weight; PROM: premature rupture of membranes; IRDS: infants respiratory distress syndrome; IVH: intraventricular hemorrhage; NEC: necrotizing enterocolitis; SGA: small-for-gestational-age

RESULTS

Perinatal characteristics

At the age 19 years, 959 of the 1,338 sub-ject were alive, and 705 of them (73.5%) had been successfully followed up. Of the deceased, 96% had died within the first year (Figure 1). The characteristics of responders and nonresponders did not significantly differ, except for there being more males and slightly fewer Caesarian births among the latter (Table 1). The distribution of the subjects across the 3 groups was also not significantly different with regard to responders and nonresponders (*P*=0.93).

Figure 1: Flowchart of the follow-up response of POPS subjects at the age of 19 years.

Perinatal characteristics significantly differed in the 3 groups (Table 1). In general, infants in the VP-/VLBW+ group had fewer neonatal morbidities than the other 2 groups but were more often SGA. The VP+/VLBW+ group had the highest prevalence of neonatal morbidity, along with a longer hospital stay and more days on ventilation. The VP+/VLBW- group had the shortest hospital stay and the least SGA births.

Neurodevelopmental outcomes

Cognitive functioning

A trend towards a higher Total IQ in the VP+/VLBW- group versus the VP+/VLBW+ group was observed (0.05>*P*>0.01) (Table 2). No associations or trends were present in models 1 and 2.

The subscore Mathematical Capacity was significantly higher in the VP+/VLBW- group. This association became nonsignificant (*P*>0.01) in models 1 and 2.

Neuromotor function

Total neuromotor score was comparable in the 3 groups (Table 2). However, a trend (0.05>*P*>0.01) towards a higher Passive muscle tone subscore in the VP+/VLBW- group compared to the VP+/VLBW+ group was present, persisting in both models.

Table 2: Differences in neurodevelopment, activities and participation between subjects born VP+/VLBW+, VP+/VLBW- and VP-/VLBW+

		VP+/VLBW+	VP+/VLBW-	VP-/VLBW+	Crude regression
					VP+/VLBW- vs. VP+/VLBW+
Activities and participation	**Education**[a]	n=347	n=138	n=203	
	Level 0	39 (11.2)	14 (10.1)	21 (10.3)	1.3 (0.9 to 1.9)
	Level 1	54 (15.6)	21 (15.2)	27 (13.3)	
	Level 2	159 (45.8)	52 (37.7)	102 (50.2)	
	Level 3	95 (27.4)	51 (37.0)	53 (26.1)	
	Occupation[a]	n=339	n=131	n=201	
	Severe problem	24 (7.1)	10 (7.6)	15 (7.5)	1.0 (0.6 to 1.8)
	Moderate problem	15 (4.4)	7 (5.3)	9 (4.5)	
	Mild problem	13 (3.8)	2 (1.5)	15 (7.5)	
	No problem	287 (84.7)	112 (85.5)	162 (80.6)	
	HUI[a]	n=319	n=132	n=192	
	Moderate problem	7 (2.2)	3 (2.3)	3 (1.6)	1.0 (0.6 to 1.6)
	Mild problem	82 (25.7)	34 (25.8)	76 (39.6)	
	No problem	230 (72.1)	95 (72.0)	113 (58.9)	
	LHS[a]	n=314	n=134	n=181	
	Severe problem	0 (0.0)	2 (1.5)	0 (0.0)	0.9 (0.5 to 1.7)
	Moderate problem	6 (1.9)	3 (2.2)	5 (2.8)	
	Mild problem	31 (9.9)	9 (6.7)	25 (13.8)	
	No problem	277 (88.2)	120 (89.6)	151 (83.4)	
Neurodevelopment	**IQ total**[b]	99.9±15.8	103.7±13.6	98.6±14.3	3.8 (0.5 to 7.1)*
	Linguistic capacity z score	-0.15±0.75	0.02±0.61	-0.13±0.72	0.17 (0.00 to 0.34)
	Mathematical capacity z score	0.11±0.98	0.40±0.91	-0.02±0.90	0.29 (0.07 to 0.50)**
	Logical reasoning z score	0.09±0.87	0.27±0.72	0.01±0.84	0.19 (-0.01 to 0.38)
	Spatial visualization z score	0.17±0.87	0.38±0.77	0.17±0.81	0.21 (0.02 to 0.40)*
	Neuromotor total[b]	58.4±7.6	59.5±7.1	58.3±8.4	1.1 (-0.6 to 2.7)
	Hand function	5.4±1.0	5.5±0.9	5.4±0.9	0.04 (-0.17 to 0.24)
	Walking	7.5±1.3	7.5±1.1	7.5±1.2	0.07 (-0.19 to 0.33)
	Coordination	28.4±4.5	29.0±4.2	28.5±4.8	0.60 (-0.39 to 1.58)
	Passive muscle tone	6.1±1.6	6.5±1.5	5.9±1.8	0.42 (0.06 to 0.77)*
	Posture	11.1±1.2	11.1±1.0	11.0±1.4	0.09 (-0.18 to 0.35)
	Hearing[b]				
	Loss in best ear	6.9±7.8	4.7±5.3	5.9±5.6	-2.2 (-3.7 to -0.8)**
	Loss in worst ear	11.7±10.7	8.5±7.6	11.0±9.6	-3.3 (-5.4 to -1.2)**

[a] Values represent n (%) or OR (95%CI), analyzed with ordinal regression with the VP+/VLBW+ group as the reference.

[b] Values represent mean±SD or β (95% CI), analyzed with linear regression with the VP+/VLBW+ group as the reference.

[c] Analyses adjusted for gender, socio-economic status and ethnicity

[d] Analyses adjusted for model 1 plus neonatal morbidity (IRDS, ICH and sepsis)

* *P* value <0.05; ** *P* value <0.01

Crude regression	Model 1[c]		Model 2[d]	
VP-/VLBW+ vs. VP+/VLBW+	VP+/VLBW- vs. VP+/VLBW+	VP-/VLBW+ vs. VP+/VLBW+	VP+/VLBW- vs. VP+/VLBW+	VP-/VLBW+ vs. VP+/VLBW+
1.0 (0.8 to 1.4)	1.4 (0.9 to 2.0)	1.0 (0.7 to 1.3)	1.3 (0.9 to 2.0)	1.0 (0.7 to 1.4)
0.8 (0.5 to 1.2)	1.0 (0.6 to 1.8)	0.7 (0.5 to 1.2)	1.0 (0.5 to 1.7)	0.7 (0.4 to 1.1)
1.8 (1.2 to 2.6)**	1.0 (0.6 to 1.5)	1.8 (1.2 to 2.6)**	1.0 (0.6 to 1.6	1.7 (1.2 to 2.6)**
1.5 (0.9 to 2.5)	0.9 (0.5 to 1.8)	1.5 (0.9 to 2.6)	1.0 (0.5 to 2.0)	1.8 (1.01 to 3.2)*
-1.2 (-4.1 to 1.6)	3.0 (-0.2 to 6.3)	-1.2 (-4.0 to 1.6)	2.6 (-0.7 to 5.8)	-1.9 (-4.9 to 1.2)
0.02 (-0.13 to 0.18)	0.14 (-0.03 to 0.30)	0.02 (-0.13 to 0.16)	0.13 (-0.04 to 0.29)	-0.02 (-0.18 to 0.13)
-0.13 (-0.32 to 0.06)	0.25 (0.03 to 0.47)*	-0.14 (-0.33 to 0.05)	0.21 (-0.01 to 0.43)	-0.20 (-0.40 to 0.00)*
-0.08 (-0.25 to 0.10)	0.14 (-0.05 to 0.33)	-0.09 (-0.26 to 0.08)	0.10 (-0.10 to 0.29)	-0.16 (-0.34 to 0.02)
0.00 (-0.16 to 0.17)	0.15 (-0.04 to 0.33)	-0.01 (-0.17 to 0.15)	0.14 (-0.05 to 0.32)	-0.01 (-0.19 to 0.16)
-0.1 (-1.6 to 1.3)	1.2 (-0.5 to 2.9)	-0.2 (-1.7 to 1.2)	0.6 (-1.1 to 2.4)	-1.1 (2.7 to 0.4)
-0.04 (-0.21 to 0.14)	0.05 (-0.15 to 0.26)	-0.05 (-0.22 to 0.13)	0.01 (-0.20 to 0.22)	-0.12 (-0.30 to 0.07)
0.02 (-0.22 to 0.23)	0.08 (-0.19 to 0.35)	0.00 (-0.22 to 0.23)	0.03 (-0.25 to 0.30)	-0.09 (-0.34 to 0.16)
0.12 (-0.74 to 0.97)	0.64 (-0.36 to 1.65)	0.04 (-0.82 to 0.90)	0.35 (-0.66 to 1.36)	-0.41 (-1.33 to 0.50)
-0.16 (-0.47 to 0.15)	0.47 (0.10 to 0.83)*	-0.18 (-0.49 to 0.14)	0.38 (0.01 to 0.75)*	-0.32 (-0.66 to 0.01)
-0.10 (-0.34 to 0.13)	0.06 (-0.22 to 0.34)	-0.12 (-0.36 to 0.12)	-0.03 (-0.31 to 0.25)	-0.24 (-0.49 to 0.01)
-1.1 (-2.3 to 0.2)	-2.1 (-3.5 to -0.6)**	-1.0 (-2.3 to 0.3)	-1.9 (-3.4 to -0.4)*	-0.9 (-2.3 to 0.5)
-0.7 (-2.6 to 1.1)	-3.2 (-5.4 to -1.0)**	-0.7 (-2.6 to 1.2)	-2.9 (-5.1 to -0.7)**	-0.6 (-2.6 to 1.4)

Table 3: Differences in self- and parent-reported behavioral and emotional functioning between subjects born VP+/VLBW+, VP+/VLBW- and VP-/VLBW+

| | VP+/VLBW+ [a] | VP+/VLBW- [a] | VP-/VLBW+ [a] | Crude regression | | | Model 1 [c] | | Model 2 [d] |
				VP+/VLBW- vs. VP+/VLBW+ [b]	VP-/VLBW+ vs. VP+/VLBW+ [b]	VP+/VLBW- vs. VP+/VLBW+ [b]	VP-/VLBW+ vs. VP+/VLBW+ [b]	VP+/VLBW- vs. VP+/VLBW+ [b]	VP-/VLBW+ vs. VP+/VLBW+ [b]
	n=315	n=132	n=188						
YASR									
Anxious	6.9±6.5	4.9±5.0	7.2±6.9	-2.0 (-3.3 to -0.7)**	0.3 (-0.8 to 1.5)	-1.5 (-2.8 to -0.2)*	0.3 (-0.8 to 1.4)	-1.4 (-2.6 to -0.1)*	0.3 (-0.9 to 1.5)
Withdrawn	2.5±2.6	2.2±2.1	3.0±2.5	-0.3 (-0.8 to 0.2)	0.5 (0.03 to 0.9)*	-0.2 (-0.8 to 0.3)	0.5 (0.03 to 0.9)*	-0.2 (-0.7 to 0.3)	0.6 (0.1 to 1.0)*
Somatic	3.3±3.5	2.9±3.1	3.4±3.7	-0.4 (-1.1 to 0.3)	0.1 (-0.5 to 0.8)	-0.1 (-0.8 to 0.6)	0.2 (-0.4 to 0.9)	-0.1 (-0.8 to 0.6)	0.0 (-0.7 to 0.6)
Thought	0.3±0.9	0.3±0.6	0.4±1.2	-0.1 (-0.3 to 0.1)	0.1 (-0.1 to 0.2)	-0.1 (-0.3 to 0.1)	0.1 (-0.1 to 0.2)	-0.1 (-0.3 to 0.1)	0.1 (-0.1 to 0.2)
Attention	2.7±2.4	2.1±1.8	2.9±2.2	-0.6 (-1.0 to -0.1)*	0.2 (-0.2 to 0.6)	-0.6 (-1.0 to -0.1)*	0.2 (-0.2 to 0.6)	-0.5 (-1.0 to -0.1)*	0.2)-0.2 to 0.6)
Intrusive	1.8±2.0	1.9±2.1	1.9±2.1	0.0 (-0.4 to 0.5)	0.1 (-0.3 to 0.5)	0.0 (-0.4 to 0.5)	0.1 (-0.3 to 0.5)	0.0 (-0.4 to 0.5)	0.0 (-0.4 to 0.5)
Aggressive	2.6±2.9	2.1±2.2	2.9±3.2	-0.5 (-1.1 to 0.1)	0.3 (-0.3 to 0.8)	-0.4 (-0.9 to 0.2)	0.3 (-0.2 to 0.8)	-0.4 (-1.0 to 0.2)	0.2 (-0.4 to 0.7)
Delinquent	1.0±1.5	0.9±1.4	1.3±1.9	-0.1 (-0.4 to 0.2)	0.2 (-0.05 to 0.5)	-0.2 (-0.5 to 0.1)	0.2 (-0.1 to 0.5)	-0.2 (-0.6 to 0.1)	0.2 (-0.2 to 0.5)
Internalizing	9.4±8.4	7.1±6.4	10.2±8.8	-2.3 (-4.0 to -0.6)**	0.8 (-0.7 to 2.3)	-1.7 (-3.4 to -0.1)*	0.8 (-0.7 to 2.2)	-1.6 (-3.2 to 0.1)	0.8 (-0.7 to 2.4)
Externalizing	5.5±5.2	5.0±4.4	6.1±5.7	-0.6 (-1.6 to 0.5)	0.6 (-0.4 to 1.5)	-0.6 (-1.6 to 0.5)	0.6 (-0.3 to 1.6)	-0.6 (-1.7 to 0.5)	0.4 (-0.7 to 1.4)
Total	32.1±23.5	26.2±18.5	34.5±25.0	-5.9 (-10.6 to -1.2)*	2.4 (-1.7 to 6.6)	-4.8 (-9.5 to -0.1)*	2.6 (-1.6 to 6.7)	-4.6 (-9.4 to 0.2)	1.8 (-2.7 to 6.2)
	n=272	n=120	n=178						

Table 3: Differences in self- and parent-reported behavioral and emotional functioning between subjects born VP+/VLBW+, VP+/VLBW-, VP-/VLBW- and VP-/VLBW+ (continued)

	VP+/VLBW+[a]	VP+/VLBW-[a]	VP-/VLBW+[a]	Crude regression		Model 1[c]		Model 2[d]	
				VP+/VLBW- vs. VP+/VLBW+[b]	VP-/VLBW+ vs. VP+/VLBW+[b]	VP+/VLBW- vs. VP+/VLBW+[b]	VP-/VLBW+ vs. VP+/VLBW+[b]	VP+/VLBW- vs. VP+/VLBW+[b]	VP-/VLBW+ vs. VP+/VLBW+[b]
Anxious	5.3±5.1	3.8±4.0	5.7±5.3	-1.5 (-2.6 to -0.4)**	0.4 (-0.5 to 1.4)	-1.4 (-2.4 to -0.3)*	0.3 (-0.6 to 1.3)	-1.4 (-2.5 to -0.3)*	0.0 (-1.1 to 1.0)
Withdrawn	1.8±2.1	1.4±1.7	1.7±1.9	-0.4 (-0.8 to 0.01)	-0.1 (-0.5 to 0.3)	-0.4 (-0.8 to 0.1)	-0.1 (-0.5 to 0.3)	-0.4 (-0.8 to 0.1)	-0.1 (-0.5 to 0.3)
Somatic	2.1±2.5	2.0±2.4	2.4±2.4	-0.1 (-0.6 to 0.5)	0.3 (-0.2 to 0.8)	0.1 (-0.4 to 0.7)	0.3 (-0.2 to 0.7)	0.1 (-0.5 to 0.6)	0.0 (-0.5 to 0.5)
Thought	0.7±1.5	0.5±1.3	0.7±1.3	-0.2 (-0.5 to 0.1)	0.0 (-0.2 to 0.3)	-0.2 (-0.5 to 0.1)	0.0 (-0.3 to 0.3)	-0.2 (-0.5 to 0.1)	0.0 (-0.3 to 0.3)
Attention	4.7±4.2	3.5±3.8	4.9±3.9	-1.1 (-2.0 to -0.3)**	0.2 (-0.5 to 1.0)	-1.2 (-2.1 to -0.3)**	0.2 (-0.6 to 1.0)	-1.2 (-2.1 to -0.3)**	0.0 (-0.9 to 0.8)
Intrusive	1.7±2.0	1.8±2.3	2.1±2.5	0.1 (-0.4 to 0.6)	0.4 (-0.05 to 0.8)	0.1 (-0.4 to 0.6)	0.4 (-0.03 to 0.8)	0.2 (-0.3 to 0.7)	0.3 (-0.2 to 0.8)
Aggressive	3.4±4.4	2.9±4.3	4.1±4.8	-0.5 (-1.5 to 0.5)	0.7 (-0.2 to 1.5)	-0.3 (-1.3 to 0.7)	0.7 (-0.2 to 1.5)	-0.3 (-1.3 to 0.7)	0.4 (-0.6 to 1.3)
Delinquent	0.7±1.5	0.9±2.3	0.9±1.6	0.1 (-0.2 to 0.5)	0.2 (-0.2 to 0.5)	0.1 (-0.3 to 0.5)	0.2 (-0.2 to 0.5)	0.1 (-0.3 to 0.5)	0.0 (-0.4 to 0.4)
Internalizing	7.1±6.6	5.2±5.1	7.4±6.5	-1.9 (-3.3 to -0.6)**	0.3 (-0.9 to 1.5)	-1.7 (-3.1 to 0.4)*	0.2 (-1.0 to 1.4)	-1.7 (-3.2 to -0.3)*	-0.1 (-1.4 to 1.2)
Externalizing	5.8±6.8	5.6±8.0	7.1±7.7	-0.2 (-1.8 to 1.3)	1.2 (-0.2 to 2.6)	0.0 (-1.7 to 1.6)	1.2 (-0.2 to 2.6)	0.0 (-1.7 to 1.6)	0.7 (-0.9 to 2.2)
Total	24.3±21.5	19.9±20.5	26.4±21.1	-4.4 (-9.0 to 0.3)	2.1 (-2.0 to 6.2)	-3.8 (-8.6 to 1.0)	1.8 (-2.3 to 6.0)	-4.0 (-8.9 to 0.8)	-0.1 (-4.6 to 4.4)

(Row group label, left margin: YABCL)

[a] Values represent mean±SD

[b] Values represent β (95% CI), analyzed with linear regression with the VP+/VLBW+ group as the reference.

[c] Analyses adjusted for gender, socio-economic status and ethnicity

[d] Analyses adjusted for model 1 plus neonatal morbidity (IRDS, ICH and sepsis)

* P value <0.05

** P value <0.01

Hearing loss

Hearing loss was significantly less for both the worst and best ear in the VP+/VLBW- group, in the crude and adjusted analyses (Table 2). No differences were found between the VP+/VLBW+ and VP-/VLBW+ groups.

Behavioral and emotional functioning

In the VP+/VLBW- group, the adolescents themselves and their parents reported lower scores on the Anxious/Depressed syndrome scale, as well as on the Internalizing behavior problem scale compared to the VP+/VLBW+ group (Table 3). The parents also reported fewer Attention problems. A trend (0.05>P>0.01) towards fewer self-reported Attention problems as well as Total Problem behavior was present in the VP+/VLBW- group. Most of these associations and trends were still present in both models. Adolescents in the VP-/VLBW+ group showed a trend towards a higher score on the Withdrawn behavior syndrome scale, both in the crude and adjusted analysis.

Functional outcomes

Educational achievement

No differences were found in worst-case coding education in the 3 groups (Table 2). However, a trend towards higher educational achievement in the VP+/VLBW- group than in the VP+/VLBW+ group appeared to be present. Repeated analyses for best-case coding found similar results (data not shown).

Occupation

No differences were found between the 3 groups (Table 2). Most subjects did not experience a problem with regard to occupation.

Health status

The VP-/VLBW+ group had higher odds of reporting a lower health status than the VP+/VLBW+ group did (Table 2). This association remained significant in both models.

Perceived health

No significant differences were found in the perceived health of the 3 groups (Table 2), although there was a nonsignificant tendency towards a higher odds of reporting a worse perceived health in the VP-/VLBW+ group (Table 2).

DISCUSSION

In our study, we found that the long-term outcomes of VP+/VLBW- subjects were more favorable than those of VP+/VLBW+ subjects. On average, the subjects in the VP+/VLBW- group had a trend towards a higher IQ score, as well as less hearing loss and less self- and parent-reported behavioral problems. Additionally, a trend towards higher educational achievement was found in this group. Compared to the VP+/VLBW+ group, the VP-/VLBW+ group reported worse self-perceived health. None of the observed differences were reflected in participants' occupational achievement.

Some associations became nonsignificant after correction for demographic and/or perinatal morbidity variables. Indeed, these factors have previously been identified as predictors for poor outcomes in preterm infants.[21-24] On the other hand, other associations remained significant after correction for these variables. However, it is unclear whether the loss of statistical significance for some associations was due to (appropriate) correction for confounding variables or (inappropriate) correction for intermediate variables in the causal pathway. Nevertheless, neurodevelopmental and functional outcomes still appeared significantly different in infants born VP+/VLBW+, VP+/VLBW- and VP-/VLBW+ after analyses were adjusted for demographic and neonatal morbidity.

Our findings confirm that the entities VP and VLBW are not interchangeable. Previous research has shown that these two entities are associated with different short-term outcomes,[10] with a higher proportion of neonatal morbidities in the VP+/VLBW+ and VP+/VLBW- groups, but more SGA births in the VP-/VLBW+ group. Moreover, we have recently shown that different growth patterns up until final height are also present, with the best growth in VP+/VLBW- infants, while subjects in the VP-/VLBW+ group remained the shortest and lightest.[11] In this study, we also found differences in neurodevelopmental and functional outcomes between the terms VP and VLBW, contributing to the evidence that these two entities are indeed not the same.

The differences found in our study were statistically significant, but the effect sizes were modest, and the differences in the three groups are also likely smaller than if the groups had been compared to a VP-/VLBW- control group. The clinical implications therefore remain to be determined. Our findings mostly have implications for (clinical) research. For future studies on preterm infants, we recommend using the same inclusion criteria, thereby enabling comparisons between cohorts. Previously, recommendations have been made to base epidemiologic studies on preterm infants on GA rather than on BW [25-28]. However, as far as we are aware, these studies only researched short-term (in-hospital) outcomes. The results of our study on long-term neurodevelopmental differences, as well as our previous study on long-term growth outcomes, have added to the available evidence, showing that the differences between VP and VLBW subjects remain present into adolescence [11]. Therefore, since prematurity is defined by GA and since

pregnancy duration can be measured accurately with current technology,[29,30] we concur with the previous recommendations that GA should be used as an inclusion criterion instead of BW, at least in industrialized countries. Simultaneously, we also recommend adjusting for BW SD scores when analyzing (long-term) outcomes, since BW is also a strong determinant of long-term neurodevelopmental outcomes.[9,31,32]

The results of studies on VLBW infants cannot automatically be applied to a VP population, and vice versa; this should be taken into account when interpreting the results of a study on VP or VLBW infants. Nevertheless, the effect sizes found in our study were small, and VP and VLBW populations often do overlap with regard to clinical care. The substantial established body of literature, on both VP and VLBW subjects, therefore remains extremely valuable. However, especially as infants with increasingly younger GA are now being treated, we recommend that future studies select preterm populations primarily based on GA.

Our study has several strengths and limitations. Its major strengths are its large sample size, long-term follow-up, the analytical approach that adjusted for multiple potential confounders, and the use of a broad range of neurodevelopmental and functional outcomes. It also has its limitations. Although we found several differences in neurodevelopmental and functional outcomes in the 3 groups, the mechanism behind these differences cannot be elucidated with the available data, since the etiology of these outcomes is most likely complex and multifactorial. Additionally, since 1983, improvements in neonatal care have been made, while infants with an increasingly younger GA are being treated, and intrauterine growth is better monitored. A VP and/or VLBW cohort is therefore likely to have a different composition nowadays, and the results of this study, as well as the etiology behind these results, can therefore not necessarily be applied to the current generation of preterm infants. However, while mortality has decreased, morbidity has increased,[12] which could entail a higher risk for adverse neurodevelopmental and functional outcomes. Moreover, using either VP or VLBW as an inclusion criterion will most likely still lead to different outcomes. Additionally, we performed multiple statistical tests, and so it is possible that some of the associations were due to chance, even after adjusting the α to 0.01 for measures that yielded multiple outcomes. Our results should therefore be interpreted with caution. Lastly, the gender distribution differed between responders and nonresponders. However, since none of the other characteristics, as well as the distribution of subjects across groups, were different, it is unlikely that our results were subject to attrition bias, although this cannot be ruled out.

In conclusion, subjects born VP+/VLBW+, VP+/VLBW- and VP-/VLBW+ had significantly different neurodevelopmental and functional outcomes, although effect sizes were small. Moreover, previous research has shown that the terms VP and VLBW also lead to different short- and long-term outcomes,[10,11] indicating that these entities are not

the same. We recommend, at least in industrialized countries, that inclusion for future studies in preterm populations be based on GA instead of BW.

REFERENCES

1. Aarnoudse-Moens CS, Weisglas-Kuperus N, van Goudoever JB, Oosterlaan J. Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. Pediatrics 2009; 124:717-728

2. Nosarti C, Giouroukou E, Micali N, Rifkin L, Morris RG, Murray RM. Impaired executive functioning in young adults born very preterm. J Int Neuropsychol Soc 2007; 13:571-581

3. Hille ET, Dorrepaal C, Perenboom R, Gravenhorst JB, Brand R, Verloove-Vanhorick SP. Social lifestyle, risk-taking behavior, and psychopathology in young adults born very preterm or with a very low birthweight. J Pediatr 2008; 152:793-800, 800

4. Hille ET, Weisglas-Kuperus N, van Goudoever JB, Jacobusse GW, Ens-Dokkum MH, de GL, Wit JM, Geven WB, Kok JH, de Kleine MJ, Kollee LA, Mulder AL, van Straaten HL, de Vries LS, van Weissenbruch MM, Verloove-Vanhorick SP. Functional outcomes and participation in young adulthood for very preterm and very low birth weight infants: the Dutch Project on Preterm and Small for Gestational Age Infants at 19 years of age. Pediatrics 2007; 120:e587-e595

5. Hack M, Youngstrom EA, Cartar L, Schluchter M, Taylor HG, Flannery D, Klein N, Borawski E. Behavioral outcomes and evidence of psychopathology among very low birth weight infants at age 20 years. Pediatrics 2004; 114:932-940

6. Hack M, Cartar L, Schluchter M, Klein N, Forrest CB. Self-perceived health, functioning and well-being of very low birth weight infants at age 20 years. J Pediatr 2007; 151:635-641, 641

7. Darlow BA, Horwood LJ, Pere-Bracken HM, Woodward LJ. Psychosocial outcomes of young adults born very low birth weight. Pediatrics 2013; 132:e1521-e1528

8. Saigal S. Quality of life of former premature infants during adolescence and beyond. Early Hum Dev 2013; 89:209-213

9. Pyhala R, Lahti J, Heinonen K, Pesonen AK, Strang-Karlsson S, Hovi P, Jarvenpaa AL, Eriksson JG, Andersson S, Kajantie E, Raikkonen K. Neurocognitive abilities in young adults with very low birth weight. Neurology 2011; 77:2052-2060

10. Lapeyre D, Klosowski S, Liska A, Zaoui C, Gremillet C, Truffert P. [Very preterm infant (< 32 weeks) vs very low birth weight newborns (1500 grammes): comparison of two cohorts]. Arch Pediatr 2004; 11:412-416

11. Hollanders JJ, van der Pal SM, van Dommelen P, Rotteveel J, Finken MJ. Growth pattern and final height of very preterm versus very low birth weight infants. Pediatr Res 2017;

12. Stoelhorst GM, Rijken M, Martens SE, Brand R, den Ouden AL, Wit JM, Veen S. Changes in neonatology: comparison of two cohorts of very preterm infants (gestational age <32 weeks): the Project On Preterm and Small for Gestational Age Infants 1983 and the Leiden Follow-Up Project on Prematurity 1996-1997. Pediatrics 2005; 115:396-405

13. Serenius F, Sjors G, Blennow M, Fellman V, Holmstrom G, Marsal K, Lindberg E, Olhager E, Stigson L, Westgren M, Kallen K, group Es. EXPRESS study shows significant regional differences in 1-year outcome of extremely preterm infants in Sweden. Acta Paediatr 2014; 103:27-37

14. Bleichrodt N, Berg RH. Multicultural Capacity Test: Intermediate Level (MCT-M) - Manual. Amsterdam: NOA.

15. Samsom JF, de GL, Cranendonk A, Bezemer D, Lafeber HN, Fetter WP. Neuromotor function and school performance in 7-year-old children born as high-risk preterm infants. J Child Neurol 2002; 17:325-332

16. Touwen BC. The Examination of the Child With Minor Neurological Dysfunction: Clinics in Developmental Medicine Series. Vol 71. London, England: Heinemann.

17. Achenbach TM. Manual for the young adult self-report and young adult behavioral checklist. Burlington: University of Vermont Department of Psychiatry.

18. Statistiek CBvd. Standaard Onderwijsindeling 2006, editie 2014/'15. 2015.

19. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, Denton M, Boyle M. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care 2002; 40:113-128

20. Harwood RH, Rogers A, Dickinson E, Ebrahim S. Measuring handicap: the London Handicap Scale, a new outcome measure for chronic disease. Qual Health Care 1994; 3:11-16

21. den Ouden L, Verloove-Vanhorick SP, van Zeben-van der Aa DM, Brand R, Ruys JH. Neonatal neurological dysfunction in a cohort of very preterm and/or very low birthweight infants--relation to other perinatal factors and outcome at 2 years. Neuropediatrics 1990; 21:66-71

22. Doyle LW, Cheong JL, Burnett A, Roberts G, Lee KJ, Anderson PJ, Victorian Infant Collaborative Study G. Biological and Social Influences on Outcomes of Extreme-Preterm/Low-Birth Weight Adolescents. Pediatrics 2015; 136:e1513-1520

23. Mitha A, Foix-L'Helias L, Arnaud C, Marret S, Vieux R, Aujard Y, Thiriez G, Larroque B, Cambonie G, Burguet A, Boileau P, Roze JC, Kaminski M, Truffert P, Ancel PY, Group ES. Neonatal infection and 5-year neurodevelopmental outcome of very preterm infants. Pediatrics 2013; 132:e372-380

24. Wong HS, Edwards P. Nature or nurture: a systematic review of the effect of socio-economic status on the developmental and cognitive outcomes of children born preterm. Matern Child Health J 2013; 17:1689-1700

25. Arnold CC, Kramer MS, Hobbs CA, McLean FH, Usher RH. Very low birth weight: a problematic cohort for epidemiologic studies of very small or immature neonates. Am J Epidemiol 1991; 134:604-613

26. Blair E. The undesirable consequences of controlling for birth weight in perinatal epidemiological studies. J Epidemiol Community Health 1996; 50:559-563

27. Koller-Smith LI, Shah PS, Ye XY, Sjors G, Wang YA, Chow SSW, Darlow BA, Lee SK, Hakanson S, Lui K, Australian, New Zealand Neonatal N, Canadian Neonatal N, Swedish Neonatal Quality R. Comparing very low birth weight versus very low gestation cohort methods for outcome analysis of high risk preterm infants. BMC Pediatr 2017; 17:166

28. Mohangoo AD, Blondel B, Gissler M, Velebil P, Macfarlane A, Zeitlin J, Euro-Peristat Scientific C. International comparisons of fetal and neonatal mortality rates in high-income countries: should exclusion thresholds be based on birth weight or gestational age? PLoS One 2013; 8:e64869

29. Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. Am J Obstet Gynecol 1996; 174:278-281

30. Neilson JP. Ultrasound for fetal assessment in early pregnancy. Cochrane Database Syst Rev 2000:CD000182

31. Hack M, Flannery DJ, Schluchter M, Cartar L, Borawski E, Klein N. Outcomes in young adulthood for very-low-birth-weight infants. The New England journal of medicine 2002; 346:149-157

32. Saigal S. Follow-up of very low birthweight babies to adolescence. Seminars in neonatology : SN 2000; 5:107-118

EVA MOUTON 2019

15

Early-life growth of preterm infants and its impact on neurodevelopment

Jonneke J. Hollanders[*],
Charlotte A. Ruys[*],
Tinka Bröring,
Petra E.M. van Schie,
Sylvia M. van der Pal,
Monique van de Lagemaat,
Harrie N. Lafeber,
Joost Rotteveel,
Martijn J.J. Finken

[*] Authors contributed equally to this manuscript

ABSTRACT

Background

Increasing numbers of preterm-born children survive nowadays, and improving long-term health and neurodevelopment is becoming more important. Early-life growth has been linked to neurodevelopmental outcomes. We aimed to study whether this association has changed with time.

Methods

We studied two cohorts of preterm-born children (gestational age ≤32 weeks and/or birth weight ≤1500 g) from 1983 (n = 708) and 2003-2006 (n = 138), respectively. We distinguished four early-life growth patterns at 3 months corrected age: appropriate for gestational age (AGA) with or without growth restriction (AGA GR+/AGA GR-), and small for gestational age (SGA) with or without catch-up growth (SGA CUG+/SGA CUG-). Intelligence quotient (IQ), neuromotor function, and behavior were assessed at ages 19 and 8 years, respectively, for the cohorts.

Results

In the 2003-2006 cohort, less children had early-life GR. In both cohorts, SGA CUG- subjects had unfavorable growth trajectories and neurodevelopmental outcomes (IQ β -6.5, 95% confidence interval (CI) -9.8; -3.2, P < 0.001; neuromotor score β -1.9%, 95% CI -3.2; -0.6, P = 0.005), while SGA CUG+ subjects were comparable to adequately grown subjects.

Conclusion

Although the incidence of adverse growth patterns decreased between the cohorts, possibly indicating improvements in care over time, the impact of these growth patterns on neurodevelopmental outcomes was not significantly different. Achieving adequate early-life growth may be crucial for improving neurodevelopmental outcomes, especially for preterms born SGA.

INTRODUCTION

Neonatology is a rapidly evolving discipline. Throughout the years many advances have been made in both prenatal and neonatal care, including the widespread use of antenatal glucocorticoids, the introduction of synthetic surfactant, and feeding strategies aimed at the early introduction of (par)enteral feeding with high protein contents.[1,2] As a consequence, increasing numbers of extremely-preterm (<28 weeks gestation) and extremely low birth weight (<1,000g) infants survive, and this may in part explain why only a modest decrease in neonatal morbidities such as sepsis and retinopathy was observed.[2,3]

Furthermore, a large proportion of preterm infants experiences intrauterine growth restriction (IUGR) and is, as a consequence, usually born small-for-gestational-age (SGA). Preterm SGA-infants typically remain smaller and lighter throughout childhood, and have poorer long-term neurodevelopmental outcomes than their appropriate-for-gestational-age (AGA) counterparts.[4-6] In the early postnatal period, extra-uterine growth restriction (EUGR) may occur, resulting in 30-95% of very low birth weight (VLBW; birth weight <1,500 grams) infants being growth restricted around term age or at discharge from the hospital.[7,8] Although IUGR and EUGR are considered two separate entities, both have been associated with impairments in growth and neurodevelopment.[4,9,10] After a period of growth restriction (GR), either prenatal or postnatal, most infants show catch-up growth (CUG) in weight and length during infancy and early childhood,[6] and this has been associated with favorable neurodevelopmental outcomes as compared to GR without CUG.[11] Considering the above, it is important to take both prenatal growth (represented as SGA or AGA at birth) and postnatal growth (either steady growth, GR or CUG) into account when assessing long-term outcomes.

We aimed to study whether 1) the incidence of GR and 2) its association with childhood growth and neurodevelopmental outcomes have changed over time, by using the data of preterm-born children from two cohorts recruited 20 years apart, namely in 1983 and 2003-2006. We hypothesized that early-life GR occurs less frequent in the more recent cohort, but that growth-restricted subjects without CUG are still at a disadvantage with regard to neurodevelopmental outcomes.

METHODS

Population

Project on Preterm and Small-for-gestational-age infants
The Project On Preterm and Small-for-gestational-age infants (POPS), a nationwide cohort, consisted of almost all (94%; n=1,338) of the children born alive in 1983 in The

Netherlands who were born either very preterm (VP; gestational age <32 weeks) and/or with a VLBW,[12] of whom 998 survived after discharge. Infants were recruited at 101 out of 115 level 1 to level 3 hospitals in The Netherlands, and no exclusion criteria were applied to the cohort. Follow-up was scheduled regularly. At age 19 years, the 959 surviving subjects were approached for follow-up, of whom 705 participated.

Study Towards the Effects of Postdischarge nutrition

For the nutritional randomized controlled trial (RCT) called 'Study Towards the Effects of Postdischarge nutrition' (STEP), 152 infants born VP and/or with a VLBW between 2003 and 2006 in VU University Medical Center Amsterdam were included at birth. The Neonatal Therapeutic Intervention Scoring System (NTISS) was used as an indicator of neonatal illness severity and mortality risk. At term age, subjects were randomized to receive either a protein- and mineral-enriched post discharge formula or a standard term formula until 6 months corrected age (CA); a control group of infants fed human milk was also included. For the follow-up study at age 8 years (STEP-2), 21 subjects were excluded and 52 declined to participate or could not be traced, resulting in 79 participants.[13] At age 8 years there were no differences between the feeding groups in growth and body composition,[13] or neurodevelopmental outcomes (unpublished data).

Approval of the medical ethical committees of all participating centers was obtained, as well as written informed consent of (parents of) participants.

Growth parameters

Length/height, weight, and head circumference (HC) were measured using standard methods at birth, 3, 6, 12, and 24 months CA and at chronological age 5 and 19 years in the POPS cohort, and at birth, term age, 3, 6, 12, and 24 months CA and at chronological age 8 years in the STEP cohort.[14,15]

Growth data and definitions

Body mass index (BMI) was calculated as (weight (kg) / [length (m)]^2) from 3 months CA onwards. Subsequently, standard deviation scores (SDS) for length/height, weight, HC, and BMI were calculated.[16,17] Using a national growth chart [17] from 3 months CA onwards was deemed most feasible for various reasons: 1) the Dutch population is much taller than other populations, 2) the period of reference data collection was best fitting for both cohorts, and 3) consistency in reporting growth data of both cohorts throughout publications. Early-life growth was defined based on the difference in weight and length SDS between birth and 3 months CA, as a proxy for term age.

To define early-life growth patterns (from birth to 3 months CA), we used the following definitions:[15]

1. AGA: birth weight and birth length >−2 SDS

2. AGA with or without GR (AGA GR+/AGA GR−): weight and/or length ≤−2 SDS or >−2 SDS, respectively, at 3 months CA after being born AGA
3. SGA: birth weight and/or birth length ≤−2 SDS (as recommended by the International SGA Advisory Board in 2003)
4. SGA with or without CUG (SGA CUG+/SGA CUG−): weight and length >−2 SDS or ≤−2 SDS, respectively, at 3 months CA after being born SGA

The proportion of infants with a weight and/or length at birth of >2 SDS (2.2% for POPS and 0.7% for STEP) was considered negligible and therefore those infants were classified as AGA.

Growth trajectories were defined as length/height, weight, HC, and BMI SDS over time (from birth until age 19 or 8 years for POPS and STEP cohorts, respectively).

Neurodevelopment

Cognitive functioning

In the POPS cohort, intelligence quotient (IQ) was measured with the Multicultural Capacity Test (MCT)–intermediate level at age 19 years.[18] The MCT has been validated for individuals from different ethnic backgrounds aged ≥16 years, with an educational level ranging from 5 years of secondary school to university level. It assesses verbal and numerical intelligence, spatial visualization, speech fluency, memory, reasoning, and speed of perception.

In the STEP cohort, cognitive functioning was assessed by a child psychologist and a trained research assistant at age 8 years. Estimated Full Scale Intelligence Quotient (eFSIQ) was assessed using a four-subtest short form of the most recent Dutch version of the Wechsler Intelligence Scales for Children–third version (WISC-III).[19] eFSIQ as measured by this short form has a high reliability ($r = .93$) and correlates strongly with FSIQ ($r = .92$).[20]

The total score of both tests was expressed on a scale with a mean of 100 and an SD of 15.

Neuromotor functioning

Neuromotor functioning was assessed with the revised version of Touwen's examination of minor neurologic dysfunction in both the POPS and STEP cohort.[21] It examines six subscales: hand function, diadochokinesis, coordination, quality of walking, posture, and passive muscle tone. Each subscale, as well as the total score, was expressed as a percentage of the maximum score.

Behavior

In the POPS cohort, parental/caregiver ratings of behavioral problems were assessed using the Young Adult Behavior Checklist (YABCL) as developed by Achenbach, which provides standardized scores on behavior, feelings, and thoughts in people aged 18 to 30 years.[22] Three problem scales were derived from eight syndrome scales: "Internalizing problems" (sum of syndrome scales anxious/depressed and withdrawn behavior), "externalizing problems" (sum of syndrome scales aggressive behavior, delinquent behavior and intrusive behavior) and "total problems" (sum of all syndrome scales).

In the STEP cohort parental/caregiver ratings of behavioral problems were assessed using the Child Behavior Checklist (CBCL) as developed by Achenbach, which provides standardized scores on behavior and competences for children aged 6 to 18 years.[23] Three problem scales were derived from eight syndrome scales: "internalizing problems" (sum of syndrome scales anxious/depressed, withdrawn/depressed, and somatic complaints), "externalizing problems" (sum of syndrome scales rule-breaking behavior and aggressive behavior) and "total problems" (sum of all syndrome scales). Scores on the CBCL were standardized using T-scores with a mean of 50 and an SD of 10, where higher scores indicate higher ratings of behavioral problems.

Additionally, since deficits in attention are a known problem in preterm populations,[24] the syndrome scale "attention problems" was analyzed separately in both cohorts.

To combine behavior data from both cohorts, scores were dichotomized ('normal' vs. '(sub)clinical') based on the borderline clinical cut-off points. For the problem scales this cut-off was the 83rd percentile for the YABCL and a T-score ≥60 on the CBCL, and for the syndrome scales the 95th percentile for the YABCL and a T-score ≥65 on the CBCL.

Statistical analyses

Normally distributed variables were reported as mean ± SD and skewed variables as median (interquartile range). Subjects were compared by early-life growth pattern using analysis of variance, $\chi2$/Fisher's exact, or Kruskal-Wallis as appropriate.

Growth trajectories of length/height, weight, HC, and BMI SDS from birth until 24 months CA were visualized for comparisons of the cohorts as well as the four early-life growth patterns within each cohort. The figures were constructed using longitudinal analyses (i.e., generalized estimating equations (GEEs)), where the different measurement points were used as an interaction term with the four growth pattern groups. The method described by Figueiras et al. was used to calculate 95% confidence intervals (95% CIs) for every parameter at each time point.[25]

We also used GEEs to assess mean differences over time in length/height, weight, HC, and BMI SDS between cohorts and between early-life growth patterns within each cohort. GEE accounts for missing data, provided that these data are 'missing completely at random', therefore we included available growth data on all subjects in these analy-

ses. Furthermore, GEE adjusts for grouped samples, collected from the same subject at different times, by using a correlation structure. We used an exchangeable correlation structure, in which one average within-subject correlation between samples over time is assumed. GEEs were performed with growth parameters (expressed as SDS) as continuous, dependent variables, and early-life growth pattern or cohort as categorical, independent variable. Results are presented as β (with 95% CI), which expresses the mean difference between the groups over time, taking into account the repeated measurements within subjects.

To assess the associations between early-life growth pattern and neurodevelopmental outcomes, we combined the data of the POPS and STEP cohort, and performed linear regression analyses. Interaction between cohort and early-life growth patterns was assessed by adding interaction terms (cohort x early-life growth pattern) to the regression analyses. Next, the analyses were adjusted for cohort. Since differences in various perinatal and demographic characteristics were found between the POPS and STEP cohorts (Supplementary Table 1), analyses were subsequently also adjusted for the following potential confounders: sex, GA, antenatal corticosteroids, the presence of ≥1 comorbidities, and maternal education. Comorbidities (sepsis [diagnosed hematologically and/or through blood culture], intraventricular hemorrhage (diagnosed clinically and/or radiographically) and necrotizing enterocolitis (stage I or higher), and infant respiratory distress syndrome (diagnosed clinically and/or radiographically) for the POPS cohort) were tallied and dichotomized as 0 or ≥1 comorbidities. The AGA GR− group was the reference group in all analyses.

We compared the POPS and STEP cohorts on neurodevelopmental outcomes using linear and logistic regression analyses. Analyses were adjusted for cohort and subsequently for birth weight SDS in addition to the previously mentioned potential confounders.

Statistical analyses were performed with IBM SPSS Statistics version 22, and $P < 0.05$ was considered significant.

RESULTS

Population characteristics

In the POPS cohort, 21 children died between 3 months CA and age 19 years. There were no significant differences in mortality between the early-life growth pattern groups ($P = 0.102$). In the STEP cohort, none of the participants died between 3 months CA and age 8 years.

Early-life growth pattern classification was known for 708 of the 998 (71%) POPS subjects, and 138 of the 152 (91%) STEP subjects.

Early-life growth pattern and neurodevelopmental outcomes were known for 509 out of 705 (72%) participating POPS subjects at age 19 years, of whom 254 were AGA GR−, 84 AGA GR+, 117 SGA CUG−, and 54 SGA CUG+. At age 8 years, early life growth pattern and neurodevelopmental outcomes were known for 76 of 79 (96%) STEP subjects, of whom 54 were AGA GR−, 5 AGA GR+, 9 SGA CUG−, and 8 SGA CUG+ (Supplementary Figure 1).

Baseline characteristics according to early-life growth pattern for both cohorts are shown in Tables 1a and 1b. Baseline characteristics were comparable between participants and nonparticipants at follow-up of both the POPS and STEP cohort, with the exception of a lower target height SDS in the nonparticipant group of both cohorts ($P<0.001$ and 0.048 for POPS and STEP, respectively), and more males (69.1% vs. 45.6%) as well as less mothers with higher education (i.e., higher vocational education or university, 10.3% vs. 28.2%) in the non-participant group of the POPS cohort.

Growth

Differences in early-life growth patterns between POPS and STEP
POPS subjects were more often born SGA compared to STEP subjects (35.2% vs. 24.6%) and these SGA subjects less often showed CUG (30.1% vs. 55.9%, $P = 0.003$). In addition, GR was seen more frequent in POPS subjects compared to STEP subjects (24.8% vs. 5.8%, $P < 0.001$).

Growth trajectories
POPS cohort
Figures 1a-d show length/height, weight, HC, and BMI SDS over time, according to early-life growth pattern group. Using GEE analyses, growth trajectories from birth to age 19 years were significantly different between all growth patterns (Table 2), except for the BMI SDS trajectory between the SGA CUG− and AGA GR+ groups ($P = 0.757$).

STEP cohort
Figures 1e-h show length/height, weight, HC, and BMI SDS over time, according to early-life growth pattern group. Using GEE analyses, the AGA GR− and SGA CUG+ groups showed similar growth trajectories from birth to age 8 years for weight, HC and BMI ($P = 0.093$, $P = 0.639$, and $P = 0.914$, respectively), but not for length/height ($P = 0.028$). The SGA CUG− and AGA GR+ groups were significantly different from the AGA GR− group for all growth parameters (Table 2). SGA CUG− and AGA GR+ groups were comparable for all growth parameters (all $P > 0.05$).

Table 1a: Baseline characteristics of the POPS cohort (1983) according to early-life growth pattern

	AGA GR−	AGA GR+	SGA CUG−	SGA CUG+	*P*
n (%)	345 (49)	114 (16)	174 (25)	75 (11)	
Male gender	172 (50)	70 (61)	94 (54)	30 (40)	0.026
Gestational age, weeks	30.2 ± 1.6	29.5 ± 1.9	33.9 ± 2.4	34.0 ± 1.7	<0.001
Birth weight, grams	1423 ± 265	1216 ± 267	1140 ± 242	1309 ± 158	<0.001
SDS	0.0 ± 0.9	−0.4 ± 0.8	−3.2 ± 1.0	−2.6 ± 0.9	<0.001
Smoking during pregnancy	119 (35)	35 (31)	64 (37)	29 (39)	0.767
PROM	85 (25)	26 (23)	3 (2)	5 (7)	<0.001
Antenatal corticosteroids	68 (19.7)	14 (12.3)	5 (2.9)	5 (6.7)	<0.001
Born via caesarian section	121 (35)	42 (37)	144 (83)	54 (72)	<0.001
Apgar score >7 after 5 min	290 (84)	78 (68)	156 (90)	68 (91)	<0.001
Length of hospital stay, days	59.7 ± 20.8	87.5 ± 42.3	76.1 ± 36.1	52.4 ± 12.4	<0.001
Days of ventilation	3.6 ± 6.6	7.4 ± 9.1	2.2 ± 9.6	0.4 ± 1.5	<0.001
IRDS	158 (46)	72 (63)	22 (13)	9 (12)	<0.001
Sepsis	93 (27)	40 (35)	53 (31)	18 (24)	0.285
IVH	46 (13)	31 (27)	17 (10)	5 (7)	<0.001
NEC	11 (3)	8 (7)	16 (9)	5 (7)	0.036
Target height, SD	0.0 ± 0.8	−0.2 ± 0.9	−0.3 ± 0.8	−0.1 ± 0.8	0.002
Maternal higher education	78 (22.6)	25 (21.9)	32 (18.4)	16 (21.3)	0.707

Table 1b: Baseline characteristics of the STEP cohort (2003-2005) according to early-life growth pattern

	AGA GR−	AGA GR+	SGA CUG−	SGA CUG+	*P*
n (%)	98 (71)	6 (4)	15 (11)	19 (14)	
Male gender	42 (43)	5 (83)	11 (73)	14 (74)	0.007
Gestational age, weeks	30.1 ± 1.4	28.9 ± 1.9	30.8 ± 1.1	31.3 ± 1.1	<0.001
Birth weight, grams SDS	1414 ± 271	931 ± 211	996 ± 208	1272 ± 222	<0.001
	0.0 ± 0.7	−1.2 ± 0.4	−1.9 ± 0.6	−1.2 ± 0.6	<0.001
Smoking during pregnancy	12 (12)	2 (33)	3 (20)	4 (21)	0.263
PROM	14 (14)	1 (17)	-	2 (11)	0.441
Antenatal corticosteroids	50 (51)	5 (83)	10 (67)	13 (68)	0.221
Born via caesarian section	49 (50)	5 (83)	15 (100)	15 (79)	<0.001
Apgar score >7 after 5 min	71 (72)	6 (100)	13 (87)	14 (74)	0.429
Length of hospital stay, days	48.5 ± 14.0	59.5 ± 18.8	58.1 ± 12.0	41.3 ± 13.4	0.002
Days of ventilation	8 [2-18]	20 [10-30]	19 [6-25]	4 [0-10]	0.012[a]
NTISS	21.4 ± 7.7	26.7 ± 8.1	25.1 ± 5.5	18.7 ± 7.0	0.032
IVH	12 (12)	-	3 (20)	3 (16)	0.649
Target height, SD	−0.2 ± 0.6	−0.7 ± 0.7	−0.5 ± 0.7	−0.2 ± 0.9	0.211
Maternal higher education	40 (40.8)	1 (16.7)	6 (40.0)	6 (31.6)	0.616

Values represent mean ± SD, n (%) or median [IQR]. Continuous variables were compared with the one-way ANOVA test, dichotomous variables were compared with the Chi square/Fisher's exact test. [a] Kruskal-Wallis Test. Abbreviations: AGA, appropriate for gestational age; CUG, catch-up growth; GR, growth restriction; IVH, intraventricular hemorrhage; IRDS, infant respiratory distress syndrome; NEC, necrotizing enterocolitis; NTISS, neonatal therapeutic intervention scoring system; PROM, premature rupture of membranes; SDS, standard deviation score; SGA, small for gestational age
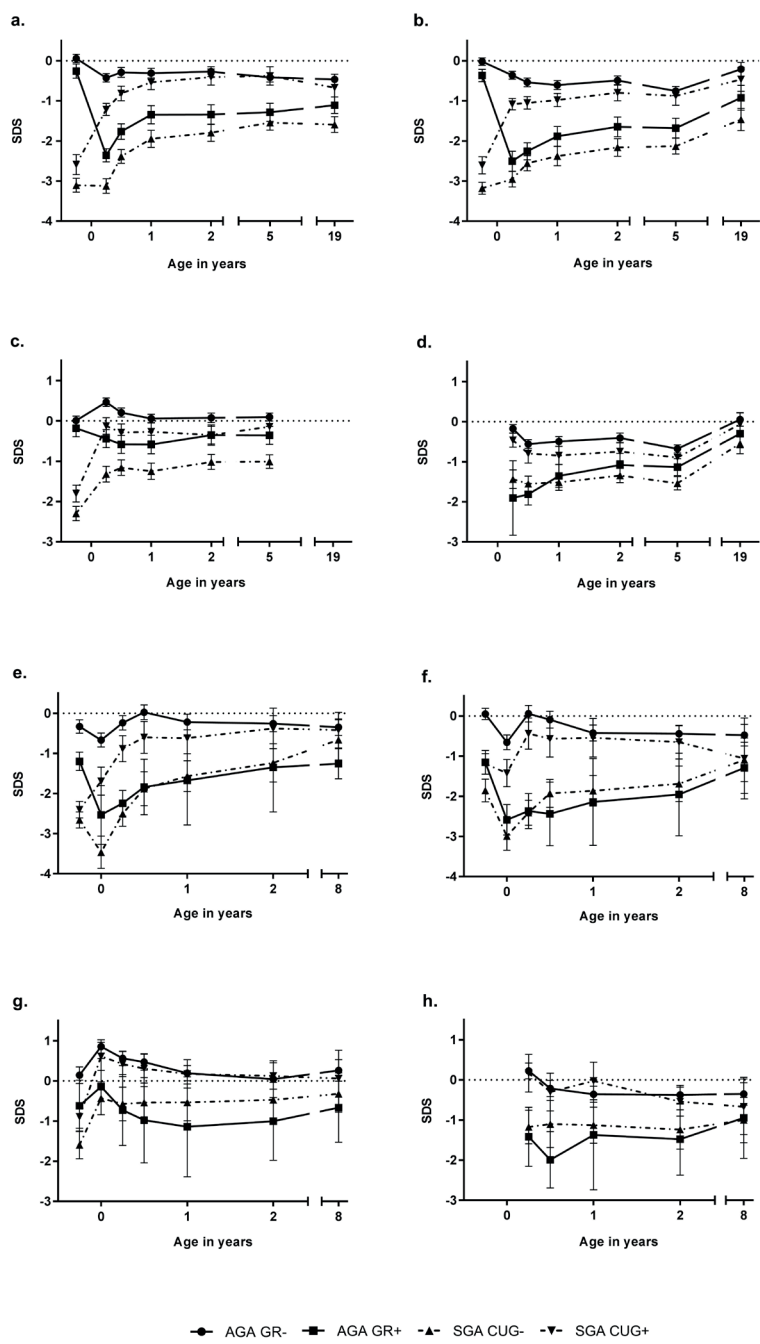
Figure 1: Growth trajectories of a/e) length/height, b/f) weight, c/g) head circumference, and d/h) BMI according to early-life growth pattern for the POPS cohort (a-d) and STEP cohort (e-h). AGA, appropriate for gestational age; CUG, catch-up growth; GR, growth restriction; SGA, small for gestational age; SDS, standard deviation score. Dotted line = reference population mean.

Table 2: Longitudinal analyses (GEE) of mean differences in growth trajectories between early-life growth patterns and between the POPS and STEP cohort.

		Length/height (SDS)		Weight (SDS)		Head circumference (SDS)		BMI (SDS)	
		β (95% CI)	P	β (95% CI)	P	β (95% CI)	P	β (95% CI)	P
POPS	AGA GR– (ref)	–		–		–		–	
	AGA GR+	–1.3 (–1.5; –1.0)	<0.001	–1.4 (–1.6; –1.2)	<0.001	–0.7 (–0.9; –0.5)	<0.001	–0.7 (–1.2; –0.3)	0.001
	SGA CUG–	–1.8 (–2.0; –1.6)	<0.001	–1.8 (–2.0; –1.6)	<0.001	–1.4 (–1.6; –1.2)	<0.001	–0.8 (–1.2; –0.4)	<0.001
	SGA CUG+	–0.4 (–0.6; –0.2)	0.001	–0.4 (–0.6; –0.2)	<0.001	–0.4 (–0.7; –0.2)	<0.001	–0.1 (–0.5; 0.3)	0.700
STEP	AGA GR– (ref)	–		–		–		–	
	AGA GR+	–1.5 (–2.2; –0.8)	<0.001	–1.8 (–2.6; –1.1)	<0.001	–1.2 (–2.3; –0.2)	0.016	–1.3 (–2.1; –0.4)	0.003
	SGA CUG–	–1.5 (–1.9; –1.1)	<0.001	–1.6 (–2.0; –1.3)	<0.001	–0.8 (–1.3; –0.4)	0.001	–1.0 (–1.4; –0.5)	<0.001
	SGA CUG+	–0.4 (–0.8; –0.1)	0.028	–0.4 (–0.8; 0.1)	0.093	–0.1 (–0.4; 0.3)	0.639	0.0 (–0.4; 0.4)	0.914
STEP vs. POPS (ref)		0.4 (0.2; 0.6)	<0.001	0.7 (0.5; 0.8)	<0.001	0.4 (0.2; 0.6)	<0.001	0.5 (0.4; 0.7)	<0.001

Abbreviations: AGA, appropriate for gestational age; CI, confidence interval; CUG, catch-up growth; GR, growth restriction; SDS, standard deviation score; SGA, small for gestational age

Growth pattern and neurodevelopmental outcomes; POPS and STEP combined

No interaction was found between early-life growth pattern and cohort, and analyses were therefore not stratified (all interaction terms *P* > 0.1, data not shown).

Table 3 shows the associations between growth pattern and neurodevelopmental outcomes. Total IQ was lower in the SGA CUG− group as compared to the AGA GR− group. Scores on total neuromotor behavior as well as all subscales, with the exception of hand function, were lower in the SGA CUG− group compared to the AGA GR− group. The AGA GR+ group scored lower on the subscales diadochokinesis and coordination, while the SGA CUG+ scored lower on passive muscle tone. The odds for (sub)clinical behavioral problems were not significantly different between all four early-life growth patterns.

Most of these findings persisted in the SGA CUG− group after adjusting for cohort, while the associations in the AGA GR+ and SGA CUG+ groups disappeared.

Adjustment for potential confounders changed some of the associations as compared to the crude analyses: the AGA GR+ group no longer scored lower on the neuromotor subscales diadochokinesis and coordination (data not shown). The SGA CUG− group no longer scored lower on diadochokinesis, coordination, and walking. However, the association of the SGA CUG− group with a lower IQ became stronger (from β −6.5, 95% CI −9.8; −3.2 to β −8.2, 95% CI −11.9; −4.4), and the association with a higher odds for (sub)clinical attention problems became significant (from odds ratio (OR) 1.5, 95% CI 0.7; 3.3 to OR 3.0 95% CI 1.2; 7.9).

POPS vs. STEP

Growth during the first two years of life

Plotting the POPS and STEP growth trajectories showed significant differences for length/height, weight, HC, and BMI, with higher SDS for all growth parameters in the STEP cohort (Figure 2). These results were confirmed by GEE analyses of all growth trajectories (Table 2). At 24 months CA, SDS for weight and length/height were significantly higher in the STEP cohort compared to the POPS cohort, while there were no differences in HC and BMI SDS (data not shown).

Neurodevelopmental outcomes

Total IQ was not significantly different between the two cohorts. Neuromotor scores were significantly higher in the STEP cohort, with the exception of the subscales hand function and walking (Table 4).

For behavior, the percentages of (sub)clinical behavioral problems per cohort (POPS/STEP) were as follows: total problem 25%/20%, internalizing 28%/22%, externalizing 14%/17%, and attention 9%/13%. The odds for (sub)clinical behavioral problems were not significantly different in the STEP vs. POPS cohort (Table 4).

Adjustment for potential confounders did not change these results (data not shown).

Table 3: Association between early-life growth pattern and neurodevelopmental outcomes in the POPS and STEP cohort combined

	AGA GR− (ref) (n = 443)	AGA GR+ (n = 120)				SGA CUG− (n = 189)				SGA CUG+ (n = 94)			
	Mean (95% CI)	β (95% CI)		OR (95% CI)		β (95% CI)		OR (95% CI)		β (95% CI)		OR (95% CI)	
		crude	adjusted [a]	crude	adjusted [a]	crude	adjusted [a]	crude	adjusted [a]	crude	adjusted [a]	crude	adjusted [a]
Total IQ	102.1 (100.3; 103.9)	−3.1 (−7.0; 0.8)	−3.1 (−7.0; 0.9)			−6.5 (−9.8; −3.2)*	−6.5 (−9.8; −3.2)*			−0.9 (−5.3; 3.5)	−0.8 (−5.2; 3.6)		
Neuromotor behavior – % of maximum score													
- Total score	94.4 (93.6; 95.1)	−0.9 (−2.4; 0.6)	−0.3 (−1.8; 1.2)			−1.9 (−3.2; −0.6)**	−1.4 (−2.7; −0.1)**			−1.0 (−2.8; 0.8)	−0.7 (−2.5; 1.0)		
- Hand function	96.8 (96.1; 97.6)	−1.0 (−2.5; 0.6)	−1.5 (−3.0; 0.1)			−1.3 (−2.6; 0.1)	−1.7 (−3.1; −0.4)**			−0.2 (−2.0; 1.6)	−0.5 (−2.2; 1.2)		
- Diadochokinesis	90.1 (88.8; 91.3)	−2.8 (−5.5; −0.2)**	−1.8 (−4.3; 0.8)			−2.6 (−4.9; −0.3)**	−1.7 (−3.9; 0.6)			−1.7 (−4.7; 1.3)	−1.1 (−4.0; 1.8)		
- Coordination	93.4 (92.5; 94.3)	−2.3 (−4.2; −0.4)**	−1.7 (−3.6; 0.1)			−1.9 (−3.5; −0.3)**	−1.4 (−3.0; 0.2)			0.1 (−2.0; 2.2)	0.4 (−1.6; 2.5)		
- Walking	96.9 (95.9; 98.0)	−0.7 (−2.8; 1.5)	−0.4 (−2.6; 1.7)			−2.3 (−4.1; −0.4)**	−2.1 (−4.0; −0.2)**			0.9 (−1.5; 3.3)	1.0 (−1.4; 3.5)		
- Posture	95.6 (94.5; 96.6)	−1.0 (−3.2; 1.2)	−0.1 (−2.2; 2.1)			−3.8 (−5.8; −1.8)*	−3.0 (−5.0; −1.1)**			−2.2 (−4.9; 0.6)	−1.9 (−4.5; 0.8)		
- Passive muscle tone	91.2 (89.8; 92.6)	−0.9 (−3.8; 2.0)	0.4 (−2.4; 3.3)			−5.7 (−8.2; −3.1)*	−4.5 (−7.0; −2.0)**			−3.7 (−7.0; −0.4)**	−2.9 (−6.1; 0.3)		
Behavior – dichotomized [b] (normal/(sub)clinical)				crude	adjusted [a]			crude	adjusted [a]			crude	adjusted [a]
- Total problem	-			1.2 (0.6; 2.2)	1.2 (0.6; 2.2)			1.4 (0.8; 2.5)	1.4 (0.8; 2.4)			1.0 (0.5; 2.0)	1.0 (0.5; 2.0)
- Externalizing	-			0.6 (0.3; 1.4)	0.6 (0.3; 1.5)			1.6 (0.9; 3.0)	1.7 (0.9; 3.1)			1.2 (0.5; 2.6)	1.2 (0.5; 2.6)
- Internalizing	-			1.5 (0.8; 2.6)	1.4 (0.8; 2.5)			1.4 (0.9; 2.4)	1.4 (0.8; 2.3)			1.0 (0.5; 1.9)	0.9 (0.5; 1.9)
- Attention	-			0.8 (0.3; 2.3)	0.9 (0.3; 2.5)			1.5 (0.7; 3.3)	1.6 (0.8; 3.5)			1.4 (0.5; 3.7)	1.4 (0.6; 3.8)

Results of linear and logistic regression analyses, represented as β (95% CI) and OR (95% CI), respectively. The AGA GR− group was used as reference group (ref). Abbreviations: AGA, appropriate-for-gestational-age; CI, confidence interval; CUG, catch-up growth; GR, growth restriction; OR, odds ratio; SGA, small-for-gestational-age. [a] linear/ logistic regression analyses adjusted for cohort, [b] Behavioral scores were dichotomized using the 'borderline clinical cut-off points.' * $P<0.001$, ** $P<0.05$
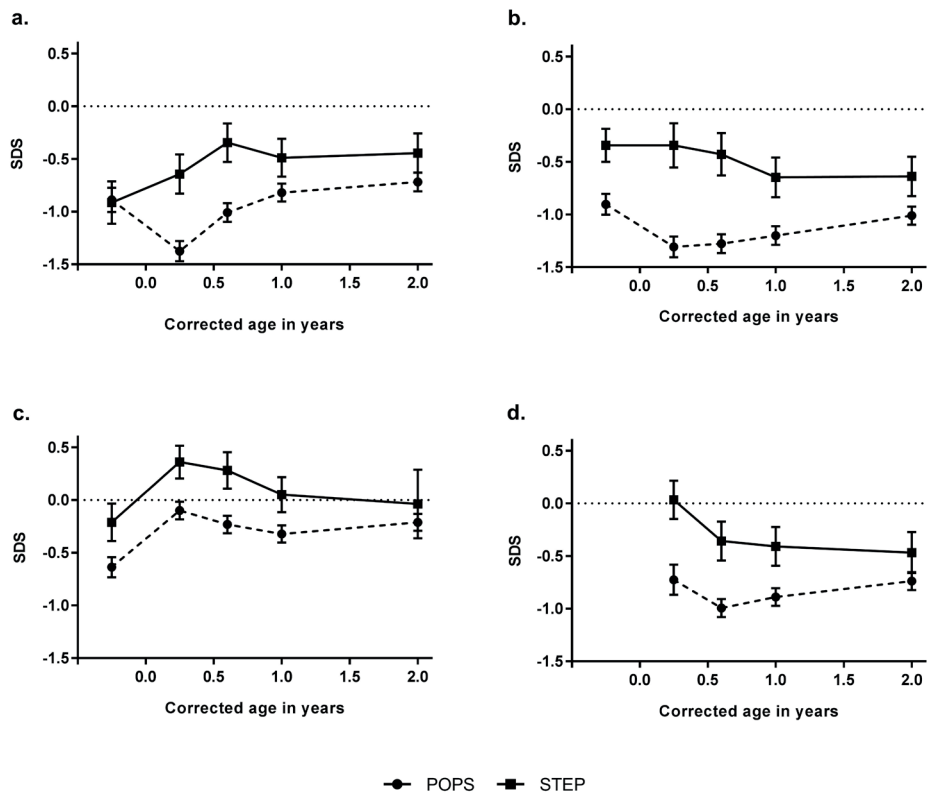
Figure 2: Growth trajectories of a) length/height, b) weight, c) head circumference, and d) BMI until 2 years corrected age, compared between the POPS and STEP cohorts. SDS, standard deviation score. Dotted line = reference population mean.

DISCUSSION

In this historical comparison of two well-described cohorts of preterm infants, we found a decrease in prenatal and postnatal GR in the more recent cohort. This could possibly be attributed to improvements in both prenatal and postnatal care over time. However, the adverse impact of early-life GR on childhood growth and neurodevelopment was not significantly different between the cohorts.

In line with other studies,[26,27] we found that early-life GR was associated with unfavorable outcomes in both cohorts. More specifically, we found that children born SGA without subsequent CUG had the greatest risk of unfavorable long-term neurodevelopmental outcomes, and that appropriate prenatal and postnatal growth were associated with favorable neurodevelopmental outcomes. Furthermore, we found similar IQ and behavioral problems in both cohorts, and better neuromotor outcomes in our more recent cohort.

Table 4: Comparison of neurodevelopmental outcomes between the POPS and STEP cohort

Cohort	POPS (ref) (n=509)	STEP (n=76)	
	Mean (95% CI)	β (95% CI) [a]	*P*
Total IQ	100.2 (99.0; 101.5)	0.3 (−3.4; 4.1)	0.873
Neuromotor function, %			
- Total score	92.7 (92.2; 93.2)	3.4 (2.1; 4.7)	<0.001
- Hand function	93.8 (96.2; 97.3)	−3.1 (−4.6; −1.6)	<0.001
- Diadochokinesis	87.5 (86.6; 88.3)	7.9 (5.4; 10.3)	<0.001
- Coordination	91.8 (91.2; 92.4)	3.6 (1.9; 5.3)	<0.001
- Walking	96.1 (95.5; 96.8)	1.3 (−0.7; 3.3)	0.188
- Posture	93.1 (92.4; 93.9)	6.1 (4.2; 8.0)	<0.001
- Passive muscle tone	87.2 (86.3; 88.1)	9.8 (7.1; 12.5)	<0.001
Behavior – dichotomized (normal/(sub)clinical) [b]	STEP vs. POPS (ref) OR (95% CI) [c]		*P*
- Total problem	0.8 (0.4; 1.5)		0.446
- Externalizing	1.2 (0.6; 2.5)		0.545
- Internalizing	0.7 (0.4; 1.3)		0.285
- Attention	1.5 (0.7; 3.3)		0.327

Abbreviations: CI, confidence interval; OR, odds ratio; ref, reference group
[a] unadjusted linear regression analyses, [b] Behavioral scores were dichotomized using the 'borderline clinical cut-off points', [c] unadjusted logistic regression analyses

As far as we know, no other study has assessed whether the association between early-life growth pattern and (neurodevelopmental) outcomes were different between cohorts recruited 20 years apart, as a proxy for changes in time. Between 1983 and 2003 perinatal care has changed dramatically. With respect to antenatal care, advances in ultrasound evaluation of fetal growth improved the possibility to identify severe IUGR and this identification may support the choice to actively induce (preterm) labor to prevent an infant from being born severely SGA.[28] In addition, the use of antenatal glucocorticoids for the induction of fetal lung maturation has become common practice, and this has been associated with a protective effect against neurodevelopmental impairment.[29] With respect to postnatal care for preterm infants, tremendous progress has been made by the use of surfactant and less invasive ventilation as well as by optimizing early nutrition.

Nevertheless, severe GR in the early postnatal period has been and still is a major concern in the care for preterm infants.[30,31] A significant part of this GR may be contributed to cumulative nutritional deficiencies acquired in the first postnatal weeks, as Embleton et al. showed that recommended daily nutritional intakes were rarely achieved.[32] This could at least in part be prevented by ensuring adequate protein and energy intake according to current guidelines, as soon as possible after birth.[1] A positive effect of

providing adequate early nutritional care on later neurodevelopmental outcomes has been suggested.[31]

However, during the last two decades, the improvements in neonatal care and nutrition have not led to a clear decrease in short-term and long-term morbidities.[3,33] This may in part be attributed to the survival of children born at a lower gestational age, who generally have a more complicated neonatal course. When comparing the two cohorts, we found a decrease in the percentage of children with early-life GR. We might have expected this improvement in early-life growth to be accompanied by equally improved neurodevelopmental outcomes; however, only a modest advantage for the recent cohort in some of the neurodevelopmental outcomes was observed. This may be partly explained by the smaller sample size and the lower gestational age of the more recent cohort, or because the mean scores on neurodevelopmental outcomes in the 1983 cohort were already within the reference range. The small benefits on neuromotor functioning might partly be attributed to improvements in perinatal care and the selection of healthier preterm infants for the STEP RCT.

In the more recent cohort, unfavorable early-life growth patterns were still related to unfavorable neurodevelopmental outcomes, and therefore achieving adequate early-life growth by further optimizing early nutrition and minimizing disease burden appears to be essential to improve outcomes of preterm-born infants.[34] Although acknowledging the importance of CUG for childhood growth and neurodevelopmental outcomes, it is important to recognize the possible downside of (excessive) CUG in weight, that is, the association with an increased risk of cardiometabolic diseases at later age.[35] In addition to adequate early-life growth, interventions such as increasing parent-infant interaction, sensory stimulation, and physiotherapy during and after the neonatal intensive care unit (NICU)-period, could positively influence both motor and cognitive development as well, although long-term effects have to be established.[36,37]

The association between adverse early-life growth and unfavorable neurodevelopmental outcomes might either be causal or might be explained by perinatal characteristics clustering with prenatal and postnatal GR. For example, in our study, subjects with poor early-life growth had a lower birth weight and their length of hospital stay and total days of ventilation were longer, so they already seemed to be at a disadvantage compared to the subjects with appropriate early-life growth. Clustering is particularly evident in case of bronchopulmonary dysplasia, which, besides being a risk factor for postnatal GR, has been strongly associated with poor academic achievement.[38] Furthermore, the effect of severity of illness on later outcomes may be mediated by the early nutritional management.[39] The association between early-life growth and neurodevelopment might at least partly be explained by disease severity, which could actually be the underlying cause for both the adverse early-life growth and the unfavorable neurodevelopmental outcomes.

Strengths and limitations

Our study has several strengths and limitations. The major strength of our study is that we had the unique opportunity to compare extensive data from two cohorts with a long-term follow-up. Our study has several limitations, which could offer an alternative explanation for the results of our study. A selection bias might be present due to the original RCT design of the STEP cohort compared to the observational design of the POPS cohort, as well as due to recruiting from a single center vs. multiple centers. The change in the incidence of adverse early-life growth patterns could also be attributed to these factors. Additionally, the sample size of the STEP cohort at follow-up was very small, with <10 subjects for three out of four early-life growth pattern groups, and the power to detect changes in neurodevelopmental outcomes between the two cohorts might therefore be too limited. Moreover, comparing IQ measured at age 8 and 19 years may ignore the possibility of changes in neurodevelopmental outcomes over time within a preterm-born population. Current literature is inconclusive as to how cognitive performance develops from school age onwards in a preterm population, with studies showing stable,[40] improved,[41] and worsened [42] results within the same cohort. It has also been suggested that executive functioning and academic performance have worsened with time when comparing sequential cohorts.[43] Furthermore, although predischarge and postdischarge growth appears to impact neurodevelopment differently,[44] the lack of anthropometric data at term age for the POPS cohort forced us to use the CA of 3 months to define early-life growth patterns. Methodologically our study was limited as well: both the age at testing and the instruments used to assess neurodevelopment differed between our cohorts. However, the tests were both conceptually and method-ologically similar enough for a pooled data analysis using standardized scores (i.e., by us-ing predefined cut-offs instead of raw scores). The use of parental reports on behavioral problems might lead to an overestimation of behavioral problems, although parents of preterm subjects appear to appraise their child's health quite accurately.[45] We decided to use the parental report because almost 82% of the POPS subjects were still living at home at the time of follow-up, suggestive of a reliable parental report. Moreover, parent reports were available for both cohorts, enabling comparison. Lastly, we acknowledge that, as with any historical cohort, generalizability to current NICU populations may be hampered because of ongoing changes in perinatal care policies.

CONCLUSION

The incidence of adverse early-life growth patterns is significantly lower in our cohort from 2003 compared to our cohort from 1983, possibly indicating improvements in care over time. However, the impact of adverse early-life growth on neurodevelopment was

not significantly different between the cohorts. Children born SGA without CUG remain vulnerable and should be followed closely with regard to the timely detection of neuro-developmental problems during childhood. Children born SGA with CUG had outcomes similar to adequately grown children. Ongoing attention for adequate early-life growth is needed, and interventions to support neurodevelopment, specifically in infants with early-life GR, should be considered.

REFERENCES

1. Senterre T, Rigo J. Reduction in postnatal cumulative nutritional deficit and improvement of growth in extremely preterm infants. Acta Paediatr 2012; 101:e64-70

2. Stoll BJ, Hansen NI, Bell EF, Walsh MC, Carlo WA, Shankaran S, Laptook AR, Sanchez PJ, Van Meurs KP, Wyckoff M, Das A, Hale EC, Ball MB, Newman NS, Schibler K, Poindexter BB, Kennedy KA, Cotten CM, Watterberg KL, D'Angio CT, DeMauro SB, Truog WE, Devaskar U, Higgins RD. Trends in Care Practices, Morbidity, and Mortality of Extremely Preterm Neonates, 1993-2012. Jama 2015; 314:1039-1051

3. Stoelhorst GM, Rijken M, Martens SE, Brand R, den Ouden AL, Wit JM, Veen S. Changes in neonatology: comparison of two cohorts of very preterm infants (gestational age <32 weeks): the Project On Preterm and Small for Gestational Age Infants 1983 and the Leiden Follow-Up Project on Prematurity 1996-1997. Pediatrics 2005; 115:396-405

4. Guellec I, Lapillonne A, Marret S, Picaud JC, Mitanchez D, Charkaluk ML, Fresson J, Arnaud C, Flamand C, Cambonie G, Kaminski M, Roze JC, Ancel PY. Effect of Intra- and Extrauterine Growth on Long-Term Neurologic Outcomes of Very Preterm Infants. J Pediatr 2016; 175:93-99.e91

5. Murray E, Fernandes M, Fazel M, Kennedy SH, Villar J, Stein A. Differential effect of intrauterine growth restriction on childhood neurodevelopment: a systematic review. BJOG : an international journal of obstetrics and gynaecology 2015; 122:1062-1072

6. Euser AM, de Wit CC, Finken MJ, Rijken M, Wit JM. Growth of preterm born children. Horm Res 2008; 70:319-328

7. Horbar JD, Ehrenkranz RA, Badger GJ, Edwards EM, Morrow KA, Soll RF, Buzas JS, Bertino E, Gagliardi L, Bellu R. Weight Growth Velocity and Postnatal Growth Failure in Infants 501 to 1500 Grams: 2000-2013. Pediatrics 2015; 136:e84-92

8. Griffin IJ, Tancredi DJ, Bertino E, Lee HC, Profit J. Postnatal growth failure in very low birthweight infants born between 2005 and 2012. Arch Dis Child Fetal Neonatal Ed 2016; 101:F50-55

9. Levine TA, Grunau RE, McAuliffe FM, Pinnamaneni R, Foran A, Alderdice FA. Early childhood neurodevelopment after intrauterine growth restriction: a systematic review. Pediatrics 2015; 135:126-141

10. Pampanini V, Boiani A, De Marchis C, Giacomozzi C, Navas R, Agostino R, Dini F, Ghirri P, Cianfarani S. Preterm infants with severe extrauterine growth retardation (EUGR) are at high risk of growth impairment during childhood. European journal of pediatrics 2015; 174:33-41

11. Belfort MB, Rifas-Shiman SL, Sullivan T, Collins CT, McPhee AJ, Ryan P, Kleinman KP, Gillman MW, Gibson RA, Makrides M. Infant growth before and after term: effects on neurodevelopment in preterm infants. Pediatrics 2011; 128:e899-906

12. Verloove-Vanhorick SP, Verwey RA, Brand R, Gravenhorst JB, Keirse MJ, Ruys JH. Neonatal mortality risk in relation to gestational age and birthweight. Results of a national survey of preterm and very-low-birthweight infants in the Netherlands. Lancet 1986; 1:55-57

13. Ruys CA, van de Lagemaat M, Finken MJ, Lafeber HN. Follow-up of a randomized trial on postdischarge nutrition in preterm-born children at age 8 y. Am J Clin Nutr 2017; 106:549-558

14. Hollanders JJ, van der Pal SM, van Dommelen P, Rotteveel J, Finken MJJ. Growth pattern and final height of very preterm vs. very low birth weight infants. Pediatr Res 2017; 82:317-323

15. Ruys CA, van der Voorn B, Lafeber HN, van de Lagemaat M, Rotteveel J, Finken MJJ. Birth weight and postnatal growth in preterm born children are associated with cortisol in early infancy, but not at age 8 years. Psychoneuroendocrinology 2017; 82:75-82

16. Niklasson A, Albertsson-Wikland K. Continuous growth reference from 24th week of gestation to 24 months by gender. BMC Pediatr 2008; 8:8

17. Fredriks AM, van Buuren S, Burgmeijer RJ, Meulmeester JF, Beuker RJ, Brugman E, Roede MJ, Verloove-Vanhorick SP, Wit JM. Continuing positive secular growth change in The Netherlands 1955-1997. Pediatr Res 2000; 47:316-323

18. Bleichrodt N, Berg RH. Multicultural Capacity Test: Intermediate Level (MCT-M) - Manual. Amsterdam: NOA.

19. Wechsler D. Wechsler Intelligence Scale for Children-III-NL. 3rd ed. Amsterdam, The Netherlands: Harcourt Test Publishers.

20. Kaufman AS, Kaufman JC, Balgopal R, McLean JE. Comparison of three WISC-III short forms: Weighing psychometric, clinical and practical factors. Journal of Clinical Child Psychology 1996; 25:97-105

21. Samsom JF, de GL, Cranendonk A, Bezemer D, Lafeber HN, Fetter WP. Neuromotor function and school performance in 7-year-old children born as high-risk preterm infants. J Child Neurol 2002; 17:325-332

22. Achenbach TM. Manual for the young adult self-report and young adult behavioral checklist. Burlington: University of Vermont Department of Psychiatry.

23. Verhulst FC, van der Ende J, Koot JM. Handleiding voor de CBCL/4-18. Afdeling Kinder- en Jeugdpsychiatrie, Sophia Kinderziekenhuis/Academisch Ziekenhuis Rotterdam/Erasmus Universiteit Rotterdam 1996;

24. Eryigit-Madzwamuse S, Wolke D. Attention problems in relation to gestational age at birth and smallness for gestational age. Early Hum Dev 2015; 91:131-138

25. Figueiras A, Domenech-Massons JM, Cadarso C. Regression models: calculating the confidence interval of effects in the presence of interactions. Stat Med 1998; 17:2099-2105

26. Varella MH, Moss WJ. Early growth patterns are associated with intelligence quotient scores in children born small-for-gestational age. Early Hum Dev 2015; 91:491-497

27. Jensen RB, Juul A, Larsen T, Mortensen EL, Greisen G. Cognitive ability in adolescents born small for gestational age: Associations with fetal growth velocity, head circumference and postnatal growth. Early Hum Dev 2015; 91:755-760

28. Monier I, Ancel PY, Ego A, Guellec I, Jarreau PH, Kaminski M, Goffinet F, Zeitlin J, Group ES. Gestational age at diagnosis of early-onset fetal growth restriction and impact on management and survival: a population-based cohort study. BJOG : an international journal of obstetrics and gynaecology 2017; 124:1899-1906

29. Chawla S, Natarajan G, Shankaran S, Pappas A, Stoll BJ, Carlo WA, Saha S, Das A, Laptook AR, Higgins RD. Association of Neurodevelopmental Outcomes and Neonatal Morbidities of Extremely Premature Infants With Differential Exposure to Antenatal Steroids. JAMA Pediatr 2016; 170:1164-1172

30. Ehrenkranz RA, Younes N, Lemons JA, Fanaroff AA, Donovan EF, Wright LL, Katsikiotis V, Tyson JE, Oh W, Shankaran S, Bauer CR, Korones SB, Stoll BJ, Stevenson DK, Papile LA. Longitudinal growth of hospitalized very low birth weight infants. Pediatrics 1999; 104:280-289

31. Belfort MB, Ehrenkranz RA. Neurodevelopmental outcomes and nutritional strategies in very low birth weight infants. Semin Fetal Neonatal Med 2017; 22:42-48

32. Embleton NE, Pang N, Cooke RJ. Postnatal malnutrition and growth retardation: an inevitable consequence of current recommendations in preterm infants? Pediatrics 2001; 107:270-273

33. Costeloe KL, Hennessy EM, Haider S, Stacey F, Marlow N, Draper ES. Short term outcomes after extreme preterm birth in England: comparison of two birth cohorts in 1995 and 2006 (the EPICure studies). BMJ (Clinical research ed) 2012; 345:e7976

34. Chan SH, Johnson MJ, Leaf AA, Vollmer B. Nutrition and neurodevelopmental outcomes in preterm infants: a systematic review. Acta Paediatr 2016; 105:587-599

35. Okada T, Takahashi S, Nagano N, Yoshikawa K, Usukura Y, Hosono S. Early postnatal alteration of body composition in preterm and small-for-gestational-age infants: implications of catch-up fat. Pediatr Res 2015; 77:136-142

36. Hughes AJ, Redsell SA, Glazebrook C. Motor Development Interventions for Preterm Infants: A Systematic Review and Meta-analysis. Pediatrics 2016; 138

37. Pineda R, Guth R, Herring A, Reynolds L, Oberle S, Smith J. Enhancing sensory experiences for very preterm infants in the NICU: an integrative review. Journal of perinatology : official journal of the California Perinatal Association 2017; 37:323-332

38. Twilhaar ES, de Kieviet JF, Aarnoudse-Moens CS, van Elburg RM, Oosterlaan J. Academic performance of children born preterm: a meta-analysis and meta-regression. Arch Dis Child Fetal Neonatal Ed 2017;

39. Ehrenkranz RA, Das A, Wrage LA, Poindexter BB, Higgins RD, Stoll BJ, Oh W. Early nutrition mediates the influence of severity of illness on extremely LBW infants. Pediatr Res 2011; 69:522-529

40. Linsell L, Johnson S, Wolke D, O'Reilly H, Morris JK, Kurinczuk JJ, Marlow N. Cognitive trajectories from infancy to early adulthood following birth before 26 weeks of gestation: a prospective, population-based cohort study. Arch Dis Child 2017;

41. Luu TM, Vohr BR, Allan W, Schneider KC, Ment LR. Evidence for catch-up in cognition and receptive vocabulary among adolescents born very preterm. Pediatrics 2011; 128:313-322

42. O'Brien F, Roth S, Stewart A, Rifkin L, Rushe T, Wyatt J. The neurodevelopmental progress of infants less than 33 weeks into adolescence. Arch Dis Child 2004; 89:207-211

43. Burnett AC, Anderson PJ, Lee KJ, Roberts G, Doyle LW, Cheong JLY. Trends in Executive Functioning in Extremely Preterm Children Across 3 Birth Eras. Pediatrics 2018; 141

44. Belfort MB, Gillman MW. Healthy infant growth: what are the trade-offs in the developed world? Nestle Nutr Inst Workshop Ser 2013; 71:171-184

45. Saigal S, Rosenbaum PL, Feeny D, Burrows E, Furlong W, Stoskopf BL, Hoult L. Parental perspectives of the health status and health-related quality of life of teen-aged children who were extremely low birth weight and term controls. Pediatrics 2000; 105:569-574
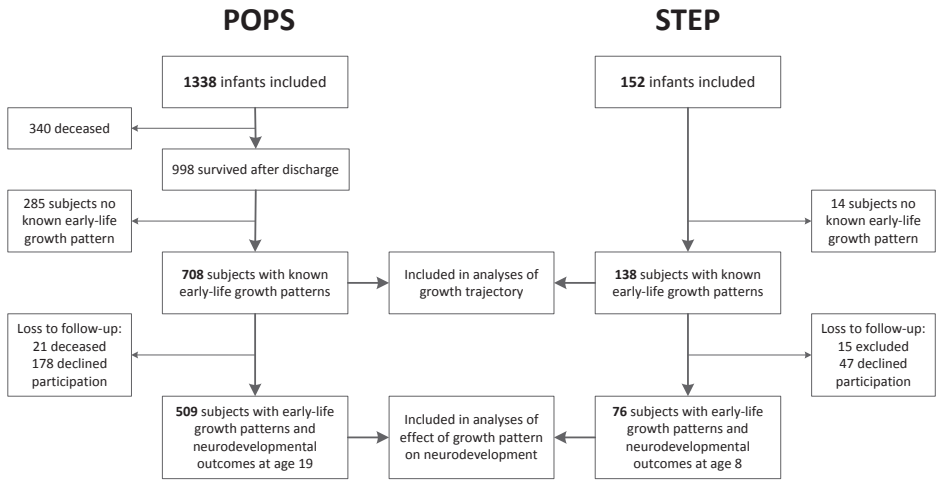
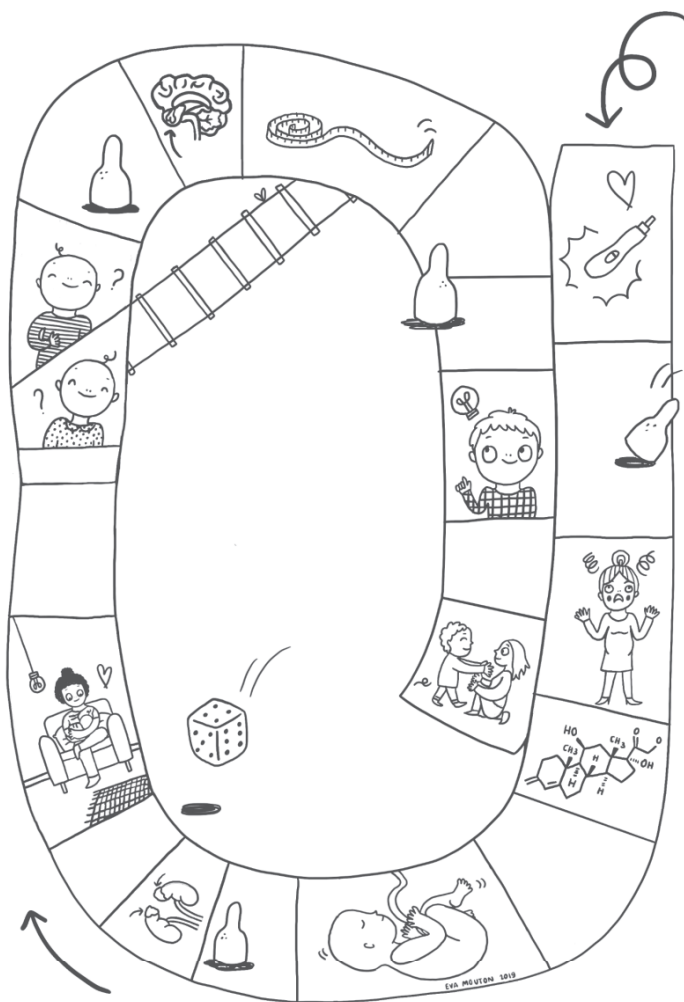**Supplementary Figure 1:** Flowchart of subjects included in the different analyses for both cohorts.

Supplementary Table 1: Baseline characteristics of participants in the STEP vs. POPS follow-up

	STEP	POPS	*P*
n	63	509	
Male sex	29 (46.0)	232 (45.6)	0.946
Gestational age, weeks	30.1 ± 1.7	31.3 ± 2.6	<0.001
Birth weight, grams	1292 ± 310	1302 ± 282	0.797
SDS	−0.4 ± 1.0	−1.1 ± 1.6	<0.001
Smoking during pregnancy	6 (9.5)	165 (32.4)	<0.001
PROM	5 (7.9)	88 (17.3)	0.058
Antenatal corticosteroids	37 (58.7)	66 (13.0)	<0.001
Born via caesarian section	41 (65.1)	267 (52.5)	0.058
Apgar score >7 after 5 min	47 (74.6)	435 (85.5)	<0.001
Length of hospital stay, days	51.1 ± 16.1	65.7 ± 25.7	<0.001
Days of ventilation	10 [2-22]	0 [0-5]	<0.001
≥1 comorbidity	33 (52.4)	293 (57.7)	0.423
Target height, SD	0.5 ± 0.7	−0.0 ± 0.8	<0.001
Early-life growth pattern			
AGA GR−	44 (69.8)	254 (49.9)	0.010
AGA GR+	3 (4.8)	84 (16.5)	
SGA CUG−	9 (14.3)	117 (23.0)	
SGA CUG+	7 (11.1)	54 (10.6)	
Maternal higher education	26 (41.3)	136 (28.7)	0.041

Values represent n (%) or mean ± sd. Continuous variables were compared with the independent samples t-test and dichotomous variables with the Chi square test.

Discussion

16

General discussion

GENERAL DISCUSSION

In this chapter, the main findings of this thesis are summarized, some strengths and limitations are reflected on, and implications for future research are discussed.

METHODOLOGICAL CONSIDERATIONS

Composition of study population

In order to measure the correct determinants, composing an appropriate study population is crucial. When studying preterm infants, many inclusion criteria have previously been used, ranging from a gestational age (GA) of <37 weeks to <28 weeks. Additionally, infants have also been included based on birth weight (BW), from <2,500 grams to <1,000 grams. In **chapters 13 and 14** we have compared the differences in outcomes when infants were included based on being born very preterm (VP; i.e., GA<32 weeks) and/or with a very low birth weight (VLBW; i.e., BW<1,500 grams) by using a cohort which was established in 1983, the Project On Preterm and Small-for-gestational-age infants (POPS cohort). Not surprisingly, VP infants significantly differed from VLBW infants in growth pattern, final height and neurodevelopmental outcomes. For future studies it is desirable to use inclusion criteria that result in representative samples of preterm infants. Since the definition of prematurity is based on pregnancy duration and not because of a low birth weight, we recommend that future studies in preterm infants should include subjects based on gestational age rather than birth weight, at least in countries where gestational age can be reliably assessed in utero. This is in line with previous recommendations which were based on short-term outcomes.[1-4] In addition, the inclusion criteria of the study population should be taken into account when interpreting results from previous studies, since results in populations included based on BW cannot necessarily be applied to preterm populations, and vice versa. In meta-analyses it is especially important to include a uniform group of infants, as only in this way valid conclusions can be reached.

Influence of sex on (endocrine) outcomes

Aside from including the appropriate study population, statistical analyses should also be performed correctly, including adjustment for potential confounders and/or effect modifiers. With regard to hypothalamus-pituitary-adrenal (HPA) axis activity, it was previously thought that sex differences arose during puberty, under the influence of sex hormones.[5,6] However, as shown in **chapters 9 and 10**, basal HPA-axis activity as well as HPA-axis reactivity already appear to be different between sexes in early childhood.

Consequently, it is important to take sex into account when analyzing HPA-axis (re)activity, regardless of age.

Standardization of protocols

Lastly, comparing sex differences in HPA-axis reactivity in **chapter 10** was hampered by the lack of standardization, both with regard to the testing protocol as well as when presenting results. We recommend that future studies assessing HPA-axis reactivity take the following considerations into account: 1) the type of stress that is to be assessed and which level of the HPA-axis is to be tested, 2) whether a standardized stress protocol is available for the desired type of stress to be tested, and 3) the use of a standardized presentation of results, by reporting both absolute glucocorticoid (GC) concentrations as well as derived variables, which will allow for a full overview of HPA-axis reactivity. This would enable comparisons between studies, which consequently will lead to improved conclusions with regard to stress reactivity in general and sex differences in HPA-axis reactivity specifically.

EARLY-LIFE CORTISOL REGULATION

Intra-uterine glucocorticoid regulation

By using GC concentrations measured in hair cut directly postpartum, we attempted to describe intra-uterine GC regulation in **chapters 2 and 3**.

In **chapter 2**, neonatal hair GCs were found to be >5 times higher compared to maternal GC levels directly postpartum, with a sharp decrease within ±6 weeks after birth. However, GC concentrations remained significantly higher than maternal GC levels, and it therefore appears that hair GC levels in the infant reflect both the intra- and extrauterine period at that age. Additionally, a strong positive association was found between neonatal hair GC concentrations and gestational age. Weaker, but nonetheless significant, associations were found with other perinatal factors, such as birth weight (but only in kgs, not when expressed as SDS), perinatal infection, and delivery via caesarian section. The association with gestational age is thought to reflect the positive feedback loop, a placenta-driven phenomenon, which causes an increase in GCs at the end of the third trimester.[7] This positive feedback loop might be a part of the mechanism behind the induction of labor, and appears to promote fetal organ maturation.[8,9]

Next, the associations between maternal distress during pregnancy and neonatal hair GC levels were assessed in **chapter 3**. Elevated stress scores pre- and perinatally were associated with decreased neonatal hair GC levels, with the largest decrease seen in infants who were exposed to persistent maternal stress. In contrast, maternal hair GCs were found to be increased when mothers experienced excessive anxiety symptoms

around birth. We speculate that exposure to increased maternal stress and/or maternal GCs in utero is associated with suppression of fetal HPA-axis activity. This in turn might underlie the associations between fetal exposure to maternal stress and/or excessive maternal GCs in utero and neurodevelopmental problems as well as altered HPA-axis settings in the offspring.[10] Lastly, maternal use of selective antidepressants did not affect neonatal hair GC levels, and antidepressants are therefore unlikely to explain the associations between maternal stress and neonatal hair GC levels.

Extra-uterine glucocorticoid regulation

After birth, infants are still exposed to small amounts of maternal GCs through breastmilk. Previously, our group has shown that breastmilk GCs follow the maternal HPA-axis activity, but the effect of this diurnal rhythm on the infant is unknown. In **chapter 4** we have summarized the existing knowledge on the effects of breastmilk GCs on offspring. Both in vitro and in vivo studies have shown that breastmilk GCs could promote intestinal maturation. Additionally, systemic effects have also been found, and breastmilk GCs might also affect the intestinal microbiome. However, both the laboratory analysis of GCs, as well as the sample collection in most studies leave a lot to be desired. Many studies use immunoassays, which show cross-reactivity with other hormones,[11] and none of the studies took the diurnal rhythm of breastmilk GCs into account.

We therefore conducted the Cortisol in Mother's Milk (CosMos) study, of which the results are presented in **chapters 5 to 8**. In order to ascertain that any associations found between breastmilk GCs and outcomes in the offspring were not due to other components in breastmilk, the association between breastmilk GC levels and macronutrient concentrations was studied in **chapter 5.** No association between breastmilk GCs and macronutrients were found, probably due to differences in excretion mechanisms: whereas breastmilk GCs are presumably excreted into breastmilk via passive diffusion,[12] macronutrients are subject to active secretion.[13]

Next, the effect of the diurnal rhythm of breastmilk GCs was assessed. In **chapter 6**, the diurnal rhythm of GCs in infants' saliva was determined, and several possible influencing factors were analyzed. At the age of one month, infants displayed a diurnal GC rhythm at a group level, which was strikingly found to be biphasic, possibly reflecting HPA-axis development. In utero, fetuses appear to already display a diurnal HPA-axis activity, but with a peak in the afternoon.[14] Our results might reflect the transitional period between fetal-type and adult-type HPA-axis activity. Of the studied influences, only breastmilk GC exposure and variability were weakly associated with the time of peak in the infants. These associations might be due to a signaling effect of breastmilk GCs, or due to the increased mother-infant synchrony seen in breastfed compared to formula-fed infants. Lastly, the associations between breastmilk GC rhythmicity and the infants' behavior, sleep and body composition were assessed in **chapters 7 and 8**. No associations were

found, although some results seemed to suggest that outcomes with regard to infant behavior and sleep could be modified by infant sex and/or experienced maternal stress. The lack of associations might be due to methodological shortcomings, such as a small sample size and a short follow-up, but it could also mean that breastmilk GC rhythmicity does not significantly influence the infants.

Sex differences in basal cortisol concentrations and HPA-axis reactivity in children aged 0-18 years were assessed in **chapters 9 and 10**. The meta-analysis performed in **chapter 9** showed that basal cortisol levels were already different between boys and girls prepubertally, with higher serum, saliva and urine cortisol levels in boys <8 years. During and after puberty, cortisol levels also showed sex differences, but now cortisol levels in serum and saliva were found to be lower in boys compared to girls >8 years, while urine cortisol levels remained higher in boys. However, the sex differences in serum cortisol levels were not very robust. In **chapter 10**, sex differences in HPA-axis reactivity were analyzed in a systematic review. Due to the heterogeneity of methods and presentation of results of all the included studies, age-stratified analyses could not be performed. Additionally, due to this methodological heterogeneity, although it appeared that girls had a more variable diurnal rhythm, a higher cortisol awakening response and a stronger response to social stress tests than boys, definitive conclusions could not be drawn. However, although the sex differences found in these studies were of a small magnitude and not very conclusive, they might contribute to the sex-specific origins of health and disease in the long-term.

NEURODEVELOPMENTAL OUTCOMES IN PRETERMS

Preterm thyroid regulation

Due to the immature hypothalamus-pituitary-thyroid (HPT) axis after preterm birth, preterm infants are at a risk of developing transient hypothyroxinaemia of prematurity (THoP). THoP is characterized by low T4 but normal TSH levels, and has previously been associated with adverse neurodevelopmental outcomes in childhood.[15-18] We used the POPS cohort to assess the effects of THoP on neurodevelopment at age 19 years in **chapters 11 and 12**. No associations were found between THoP and IQ or neuromotor development. THoP was associated with more internalizing and total problem behavior. However, it is unclear whether these associations are due to causality, especially since problem behavior has a multifactorial etiology. Since no effects were found on IQ and neuromotor development, and only small effects were found on behavior, our results do not support screening preterm infants for THoP. As a recent review stresses,[19] it is important to distinguish THoP from congenital hypothyroidism, so repeated screening for this purpose seems necessary. Recent findings also point towards a role of the placenta

in the availability of thyroid hormones in preterm infants,[20,21] suggesting that thyroid regulation is complex, especially in preterm infants.

Intra- and extrauterine growth

Significant medical advances have been made with regard to the treatment of preterm infants in the past decades. This has shifted attention towards long-term outcomes in these infants, especially since only a modest decrease in morbidity has been observed, despite decreasing mortality. Growth has previously been shown to be a significant contributor to neurodevelopmental outcomes.[22-25] In **chapter 15**, growth patterns and the impact of these growth patterns on neurodevelopment are compared between cohorts 20 years apart. While the occurrence of adverse growth patterns (i.e., small-for-gestational-age (SGA) without catch-up growth or appropriate-for-gestational-age with postnatal growth retardation) has decreased, which might be a reflection of improved care over time, the associations between adverse growth and neurodevelopmental outcomes did not significantly differ. Therefore, it is important to remain focused on achieving optimal growth in preterm infants.

This conclusion is strengthened by the results of **chapters 13 and 14**, in which the effects of being born VP and/or with a VLBW were compared. While these studies were not strictly performed to assess the effect of intra-uterine growth, the presence of SGA infants was highest in the VP-/VLBW+ group, and lowest in the VP+/VLBW- group. Concomitantly, growth trajectories and final height, as well as neurodevelopmental and functional outcomes were worst in the VP-/VLBW+ group, and best in VP+/VLBW- infants. Whether growth directly impacts neurodevelopment, or whether a common denominator is associated with both worse growth and worse neurodevelopmental outcomes is unclear. Additionally, the POPS cohort, in which these analyses were performed, was established in 1983. It is therefore likely that the distribution of infants between the VP and/or VLBW groups has changed. Moreover, due to improved care, which has resulted in better growth as seen in **chapter 15**, the magnitude of associations might have changed. However, the results of **chapters 13 and 14** do support the notion that continued attention should be paid to growth, both in utero and after birth.

STRENGTHS AND LIMITATIONS

One of the strengths of this thesis is the use of novel laboratory techniques to measure GCs. Our group previously developed an LC-MS/MS method to measure GCs in breast-milk, which enabled us to assess the associations between the GC diurnal rhythm in milk and (neurodevelopmental) outcomes in offspring. We also used GCs measured in hair for some of our analyses, which offers a retrospective view of HPA-axis activity,[26] thereby

enabling us to assess intra-uterine HPA-axis regulation. Additionally, a wide variety of statistical methods were used, among them the use of specialized rhythm software, as well as the comparison of two cohorts. Lastly, several of these studies were performed using the POPS cohort, which encompasses 94% of the children born VP and/or VLBW in 1983. It is therefore a large and comprehensive cohort, with a long follow-up period.

This thesis also has its limitations. All of these studies are observational, and it is therefore not possible to assess the causality of any associations found. Observational studies are also subject to several sources of bias and uncontrollable confounding factors. For instance, in this thesis, research was hampered by inclusion of a selected population (**Chapters 2, 3, 5-8 and 15**), losses to follow-up (**Chapters 3, 11-15**), unstandardized timing of collection of samples (**Chapters 5-9**), use of non-specific laboratory methods (**chapters 9 and 10**), selection by survival (**Chapters 11 and 12**), and lack of standardized protocols and data presentation (**Chapter 10**). Additionally, the studies which used the POPS cohort (**Chapters 11-14**) might not be externally valid, since this cohort was established in 1983, and the associations found could be subject to changes in time. The most obvious change between the POPS cohort and the current NICU population, is the distribution of infants among the three VP and/or VLBW groups, as studied in **chapters 13 and 14**. Due to increased surveillance during pregnancy, severe growth retardation is less common, and it is therefore likely that the number of children born VP-/VLBW+ has decreased in comparison with the VP+ groups.

Lastly, newly gained knowledge cannot retroactively be applied to past research. The effect of THoP on neurodevelopment and behavior was therefore studied in the entire POPS cohort (**chapters 11 and 12**), rather than only the VP infants, as recommended in **chapter 13 and 14**. However, stratified analyses according to gestational age (< and ≥ 29 weeks) were performed, which did not change the results.

FUTURE PERSPECTIVES AND IMPLICATIONS

Future research

As was seen in several of the chapters in this thesis, research in children and in general would benefit from an increase in standardization. When studying preterm infants it is best to include infants based on gestational age rather than birth weight, at least in countries where gestational age can be reliably assessed in utero. Simultaneously, the application of results to clinical practice should be limited to the researched population in the study, and not extrapolated to infants with different characteristics. Moreover, systematic reviews and meta-analyses should only use studies with uniform inclusion criteria, as only then valid conclusions can be drawn. Methodology, especially for stress tests, should also be standardized as best as possible, which will facilitate the comparison

of results between studies. Lastly, correcting for sex is recommended when analyzing results of studies into HPA axis activity.

Glucocorticoid regulation

Although this thesis has already shed some light on intra-uterine HPA-axis activity, future research should further elucidate this. Additionally, HPA-axis development in early life should be further investigated through a longitudinal study. This way, the origins of the biphasic diurnal rhythm found in this thesis can be further researched, while other aspects of HPA-axis development, such as the possible influence of breastmilk GCs, can also be studied in more detail.

Knowledge of physiological HPA-axis development and activity in healthy newborns enables the subsequent study of pathological HPA-axis development and activity. In preterm infants, HPA-axis activity is often inadequate, manifested as adrenal insufficiently.[27] Future research should investigate HPA-axis development in preterm infants, as well as the associations between HPA-axis activity and clinical correlates including refractive hypotension and bronchopulmonary dysplasia (BPD). Additionally, prematurity has been associated with an upregulated HPA-axis in childhood,[28] and it should be further investigated whether this is associated with early-life HPA-axis activity.

Neurodevelopment in preterms

Neurodevelopment in preterm infants is a difficult and multifactorial issue. This thesis once again confirms the importance of adequate growth, both in utero as well as after birth, on neurodevelopmental outcomes. It is therefore important to pay continued attention to promoting healthy growth in preterm infants. THoP does not appear to influence neurodevelopment at a later age, although selection by survival cannot be ruled out, since T4 concentrations were significantly lower in the deceased group. Additionally, it is well known that thyroid hormones are crucial for the developing brain,[29] and the etiology of THoP appears to be more complex than previously thought.[20,21] It is therefore recommendable to further study the effect of THoP on neurodevelopment in a more recent cohort, in which the prevalence of morbidity and mortality is different compared to the POPS cohort. However, the influence of hormones might not be noticeable due to the bigger influences of general preterm illnesses. For example, a recent study showed that BPD explained 65% of the variance in intelligence.[30] The prevention of BPD is therefore of paramount importance. The HPA-axis has been implicated as a contributor to the development and/or prevention of BPD, and should therefore also be studied.

CONCLUSIONS

In this thesis, we have explored HPA-axis activity and development. We found that hair GCs reflect intra-uterine HPA-axis activity and were greatly affected by gestational age, while maternal distress during pregnancy also has its influences. After birth, HPA-axis activity in 1-month-old infants showed a biphasic diurnal rhythm, possibly reflecting HPA-axis development. Breastmilk GC rhythmicity might influence the infant's HPA-axis activity at age 1 month, but it was not associated with other neurodevelopmental and growth outcomes in the infants at age 3 months. Neurodevelopment at age 19 was associated with intra- and extra-uterine growth, but not with thyroid functioning in preterm infants. Lastly, in order to improve research and the comparison of studies, we concluded that methodological standardization with regard to inclusion criteria as well as testing protocols should be encouraged. These new insights form a foundation and a framework for future studies, particularly with regard to HPA-axis functioning in preterm infants and its effects on long-term neurodevelopmental outcomes.

REFERENCES

1. Arnold CC, Kramer MS, Hobbs CA, McLean FH, Usher RH. Very low birth weight: a problematic cohort for epidemiologic studies of very small or immature neonates. Am J Epidemiol 1991; 134:604-613

2. Blair E. The undesirable consequences of controlling for birth weight in perinatal epidemiological studies. J Epidemiol Community Health 1996; 50:559-563

3. Koller-Smith LI, Shah PS, Ye XY, Sjors G, Wang YA, Chow SSW, Darlow BA, Lee SK, Hakanson S, Lui K, Australian, New Zealand Neonatal N, Canadian Neonatal N, Swedish Neonatal Quality R. Comparing very low birth weight versus very low gestation cohort methods for outcome analysis of high risk preterm infants. BMC Pediatr 2017; 17:166

4. Mohangoo AD, Blondel B, Gissler M, Velebil P, Macfarlane A, Zeitlin J, Euro-Peristat Scientific C. International comparisons of fetal and neonatal mortality rates in high-income countries: should exclusion thresholds be based on birth weight or gestational age? PLoS One 2013; 8:e64869

5. McCormick CM, Lewis E, Somley B, Kahan TA. Individual differences in cortisol levels and performance on a test of executive function in men and women. Physiol Behav 2007; 91:87-94

6. Wudy SA, Hartmann MF, Remer T. Sexual dimorphism in cortisol secretion starts after age 10 in healthy children: urinary cortisol metabolite excretion rates during growth. Am J Physiol Endocrinol Metab 2007; 293:E970-E976

7. McLean M, Smith R. Corticotrophin-releasing hormone and human parturition. Reproduction 2001; 121:493-501

8. Fencl MD, Stillman RJ, Cohen J, Tulchinsky D. Direct evidence of sudden rise in fetal corticoids late in human gestation. Nature 1980; 287:225-226

9. Fowden AL, Li J, Forhead AJ. Glucocorticoids and the preparation for life after birth: are there long-term consequences of the life insurance? Proc Nutr Soc 1998; 57:113-122

10. Duthie L, Reynolds RM. Changes in the maternal hypothalamic-pituitary-adrenal axis in pregnancy and postpartum: influences on maternal and fetal outcomes. Neuroendocrinology 2013; 98:106-115

11. Ackermans MT, Endert E. LC-MS/MS in endocrinology: what is the profit of the last 5 years? Bioanalysis 2014; 6:43-57

12. Hollanders JJ, Heijboer AC, van der Voorn B, Rotteveel J, Finken MJJ. Nutritional programming by glucocorticoids in breast milk: Targets, mechanisms and possible implications. Best Pract Res Clin Endocrinol Metab 2017; 31:397-408

13. Truchet S, Honvo-Houeto E. Physiology of milk secretion. Best Pract Res Clin Endocrinol Metab 2017; 31:367-384

14. Seron-Ferre M, Riffo R, Valenzuela GJ, Germain AM. Twenty-four-hour pattern of cortisol in the human fetus at term. Am J Obstet Gynecol 2001; 184:1278-1283

15. Delahunty C, Falconer S, Hume R, Jackson L, Midgley P, Mirfield M, Ogston S, Perra O, Simpson J, Watson J, Willatts P, Williams F. Levels of neonatal thyroid hormone in preterm infants and neurodevelopmental outcome at 5 1/2 years: millennium cohort study. J Clin Endocrinol Metab 2010; 95:4898-4908

16. Den Ouden AL, Kok JH, Verkerk PH, Brand R, Verloove-Vanhorick SP. The relation between neonatal thyroxine levels and neurodevelopmental outcome at age 5 and 9 years in a national cohort of very preterm and/or very low birth weight infants. Pediatr Res 1996; 39:142-145

17. Meijer WJ, Verloove-Vanhorick SP, Brand R, van den Brande JL. Transient hypothyroxinaemia associated with developmental delay in very preterm infants. Arch Dis Child 1992; 67:944-947

18. Reuss ML, Paneth N, Pinto-Martin JA, Lorenz JM, Susser M. The relation of transient hypothyroxinemia in preterm infants to neurologic development at two years of age. N Engl J Med 1996; 334:821-827

19. Iijima S. Current knowledge of transient hypothyroxinemia of prematurity: to treat or not to treat? J Matern Fetal Neonatal Med 2019; 32:2591-2597

20. Eerdekens A, Langouche L, Guiza F, Verhaeghe J, Naulaers G, Vanhole C, Van den Berghe G. Maternal and placental responses before preterm birth: adaptations to increase fetal thyroid hormone availability? J Matern Fetal Neonatal Med 2019; 32:2746-2757

21. Eerdekens A, Verhaeghe J, Darras V, Naulaers G, Van den Berghe G, Langouche L, Vanhole C. The placenta in fetal thyroid hormone delivery: from normal physiology to adaptive mechanisms in complicated pregnancies. J Matern Fetal Neonatal Med 2019:1-10

22. Guellec I, Lapillonne A, Marret S, Picaud JC, Mitanchez D, Charkaluk ML, Fresson J, Arnaud C, Flamand C, Cambonie G, Kaminski M, Roze JC, Ancel PY. Effect of Intra- and Extrauterine Growth on Long-Term Neurologic Outcomes of Very Preterm Infants. The Journal of pediatrics 2016; 175:93-99.e91

23. Jensen RB, Juul A, Larsen T, Mortensen EL, Greisen G. Cognitive ability in adolescents born small for gestational age: Associations with fetal growth velocity, head circumference and postnatal growth. Early Hum Dev 2015; 91:755-760

24. Levine TA, Grunau RE, McAuliffe FM, Pinnamaneni R, Foran A, Alderdice FA. Early childhood neurodevelopment after intrauterine growth restriction: a systematic review. Pediatrics 2015; 135:126-141

25. Varella MH, Moss WJ. Early growth patterns are associated with intelligence quotient scores in children born small-for-gestational age. Early Hum Dev 2015; 91:491-497

26. Staufenbiel SM, Penninx BW, Spijker AT, Elzinga BM, van Rossum EF. Hair cortisol, stress exposure, and mental health in humans: a systematic review. Psychoneuroendocrinology 2013; 38:1220-1235

27. Fernandez EF, Watterberg KL. Relative adrenal insufficiency in the preterm and term infant. J Perinatol 2009; 29 Suppl 2:S44-49

28. Ng PC. Effect of stress on the hypothalamic-pituitary-adrenal axis in the fetus and newborn. The Journal of pediatrics 2011; 158:e41-43

29. Zoeller RT, Rovet J. Timing of thyroid hormone action in the developing brain: clinical observations and experimental findings. J Neuroendocrinol 2004; 16:809-818

30. Twilhaar ES, Wade RM, de Kieviet JF, van Goudoever JB, van Elburg RM, Oosterlaan J. Cognitive Outcomes of Children Born Extremely or Very Preterm Since the 1990s and Associated Risk Factors: A Meta-analysis and Meta-regression. JAMA Pediatr 2018; 172:361-367

EVA MOUTON 2019
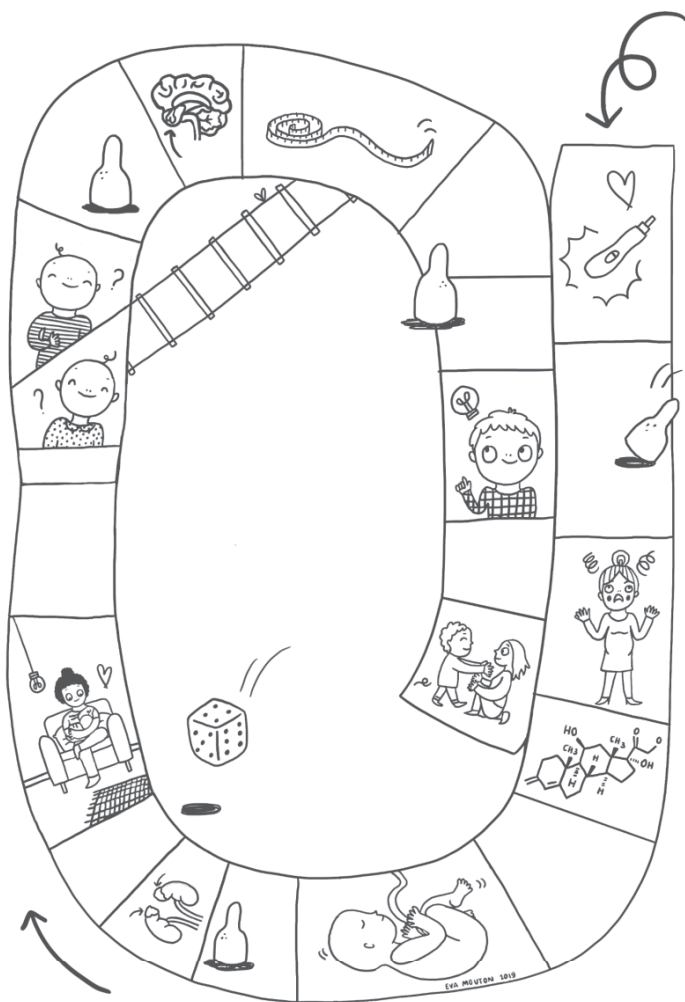
Summary

BACKGROUND

This thesis is centered around the hypothesis that events in early-life can affect health and disease at an older age.

Infants who are born preterm (with a gestational age of less than 37 weeks), form an interesting group within this hypothesis. Where normally they would still be safe in the womb, they are suddenly exposed to disturbances, like medications, procedural pain related to operations and blood draws, and infections. Meanwhile, they have trouble breathing unassisted and their organs are still immature.

In 2017, 11.978 children were born prematurely in the Netherlands, of which 2.274 were very preterm (gestational age <32 weeks). Due to improved care, an increasing number of preterm born infants survive, and focus is therefore shifting towards the long-term effects of prematurity; previous studies have shown that children who were born preterm are at an increased risk for neurodevelopmental problems, cardiovascular diseases and deviating growth. It is important to research which factors contribute to this increased risk, and it is equally important to study the factors that can improve or prevent the adverse consequences of preterm birth. Treatment of preterm infants is likely to be most successful when normal physiology is pursued, and a comprehensive understanding of these processes is therefore crucial.

The work presented in this thesis aimed to elucidate both normal physiology, as well as some of the factors contributing to or preventing adverse outcomes in preterm infants, with a focus on early-life endocrine regulation.

SUMMARY OF THE RESULTS

Part 1: Early-life glucocorticoid regulation

In **chapter 2** we explored cortisol and cortisone (glucocorticoids; GCs) concentrations measured in hair of neonates. We found that with an increasing gestational age, concentrations of GCs increased as well. This is in line with the hypothesis that a spike in GC concentrations might be a part of the mechanism behind the induction of labor. In **chapter 3** we researched the association between hair GC concentrations and experienced maternal distress during pregnancy. The infants' hair contained lower GC concentrations when their mothers experienced elevated levels of distress during pregnancy, while the mothers themselves had higher hair GC concentrations. The results of these two studies suggest that GC levels measured in hair could aid in the understanding of glucocorticoid regulation in utero.

Infants are exposed to maternal GCs in the womb, but after birth they can still be exposed to small amounts of maternal GCs through breastmilk. In **chapter 4** we have

reviewed all the available research on the effects of breastmilk GCs on the offspring thus far. However, we concluded that most studies were not performed optimally. Our research group has previously shown that breastmilk GCs follow the maternal GC rhythm that is present in blood: GC concentrations are high in the (early) morning, and are low during the evening and night. None of the studies assessing the effects of breastmilk GCs on offspring took this rhythm into account. Moreover, laboratory analyses to measure GC levels were not optimal either. We therefore recommended that future research should take breastmilk GC rhythmicity into account and determine GC levels with a sensitive method.

We performed research ourselves with regard to the effect of breastmilk rhythmicity on the offspring. First, we assessed whether there was an association between milk macronutrients (carbohydrates, protein and fat) and GCs in breastmilk in **chapter 5**, which we did not find. Next, in **chapter 6**, we studied the association between GC rhythmicity in breastmilk and in the infants' saliva at the age of 1 month. There are some indications that such an association is present, and that breastmilk GC rhythmicity could play a role in the development of a GC rhythm in infants, but the associations are quite weak. We did find that on a group level the infants displayed a different GC rhythm compared to the rhythm seen in adults: the GC rhythm in infants appeared to be biphasic, with a peak in both the morning and evening. This rhythm might be a reflection of the development towards an adult-like glucocorticoid regulation out of fetal-type glucocorticoid regulation. Subsequently, in **chapters 7 and 8**, we studied the associations between breastmilk GC rhythmicity and body composition, behavior and sleep of the infants at the age of 3 months. We did not find any associations, which might indicate that GCs in breastmilk do not significantly influence outcomes in the infants. The lack of associations might also be explained by the small sample size and relatively short follow-up period of our study.

Part 2: Glucocorticoid regulation and sex

It is known that sex differences in glucocorticoid regulation are present in adults. However, sex differences in (clinical) outcomes are already present in preterm born infants; for example, mortality risk is higher in boys compared to girls. In **chapters 9 and 10** we assessed whether sex differences with regard to basal GC concentrations as well as after stress tests are already present during childhood. To do this, we reviewed the existing literature systematically, and we also performed a meta-analysis. We discovered that sex differences are already present during childhood, both with regard to basal GC concentrations as well as after stress tests. These differences appeared to change under the influence of puberty, at least for the basal GC concentrations. The presence of sex differences in glucocorticoid regulation might partly explain the observed differences in (clinical) outcomes seen after preterm birth. It could also be a part of the mechanism

behind the different disease risk profiles between adult men and women. However, comparing the study results – especially with regard to GC concentrations after stress tests – was hampered by the many different protocols which were used, as well as the lack of a uniform presentation of the data. We have therefore recommended to use standardized stress tests as well as a standardized presentation of results for future studies on (sex differences in) glucocorticoid regulation.

Part 3: Early-life thyroid regulation in preterm infants

The organs of preterm born infants are still immature. This is also the case for the thyroid gland: it is possible that a temporary dip in thyroid hormone concentrations occurs in preterm born infants. Thyroid hormones are important for the development of the brain; children with congenital hypothyroidism or infants whose mothers had low thyroid hormone concentrations during pregnancy are at an increased risk for neurodevelopmental delays. Previous studies have shown that the temporary dip in thyroid hormones is also associated with adverse neurodevelopmental outcomes in childhood. In **chapters 11 and 12** we studied whether these associations are still present at the age of 19 years. No differences were found in IQ and motor performance between preterm-born children who did and did not have a temporary dip in thyroid hormones. Some differences in behavioral outcomes were found: adolescents were more likely to have internalizing behavior when a dip in thyroid hormones had occurred. However, we did not consider this to be clinically significant enough to recommend standard screening for the temporary thyroid hormone dip in preterm infants.

Part 4: Early-life growth and neurodevelopment

Infants can be admitted to the neonatal intensive care unit (NICU) for a multitude of reasons: for example, they were born prematurely, or their birth weight was too low. Many studies on prematurity include study subjects based on gestational age. However, some studies (also) include infants based on birth weight. Two often-used terms in studies concerning prematurity are Very Preterm (VP; gestational age <32 weeks) and Very Low Birth Weight (VLBW; birth weight <1500 grams). It is already known from previous research that the short-term outcomes are different between these two terms. We have studied whether these two terms also lead to different long-term outcomes in **chapters 13 and 14**. We discovered that infants born VP had a different growth pattern and final height compared to infants born with a VLBW. Additionally, these infants also differed with regard to neurodevelopmental outcomes: they had a different IQ-score as well as different behavioral outcomes, although no differences were found regarding education level and occupation. We have therefore recommended that future studies in preterm infants should include subjects based on gestational age rather than birth weight, at least in countries where gestational age can be reliably assessed in utero. This way, a
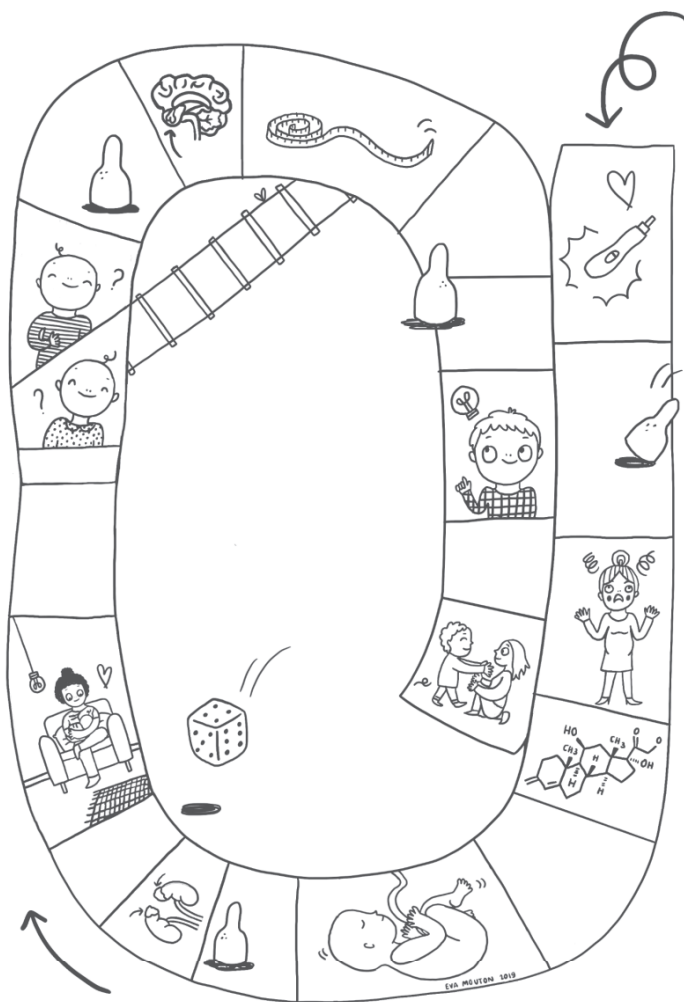
representative study population can be created and the true impact of preterm birth can be studied.

In the past decades, treatment strategies in the NICU have been improved significantly. We studied whether this has also led to different growth and neurodevelopmental outcomes in **chapter 15**. We found that an adverse growth pattern (appropriate-for-gestational-age with postnatal growth retardation, or small-for-gestational-age without catch-up growth) occurred less frequently in a cohort established in 2003 compared to a cohort established in 1983. However, the associations between adverse growth patterns and neurodevelopmental outcomes did not change in those 20 years. Therefore, it is important to remain focused on achieving optimal growth in preterm infants.

CONCLUSION

In this thesis, we explored glucocorticoid regulation and its development. We found that GCs measured in hair of neonates reflects glucocorticoid regulation in utero. Hair GC levels are mostly affected by gestational age, while maternal distress during pregnancy also has its influences. After birth, GC levels in 1-month-old infants showed a biphasic diurnal rhythm, possibly reflecting the development of an adult-like glucocorticoid regulation. Breastmilk GC rhythmicity might influence the infant's glucocorticoid regulation at age 1 month, but it was not associated with other neurodevelopmental and growth outcomes in the infants at age 3 months. Neurodevelopment at age 19 was associated with intra- and extra-uterine growth, but not with thyroid functioning in preterm infants. Lastly, in order to improve research and the comparison of studies, we concluded that methodological standardization with regard to inclusion criteria as well as testing protocols should be encouraged.

These new insights form a foundation and a framework for future studies, particularly with regard to glucocorticoid regulation in preterm infants and its effects on long-term neurodevelopmental outcomes.

Nederlandse samenvatting

ACHTERGROND

Dit proefschrift is gebaseerd rond de hypothese dat gebeurtenissen vroeg in het leven effecten kunnen hebben op de gezondheid (en het risico op aandoeningen) op latere leeftijd.

Te vroeg geboren kinderen (prematuren, geboren met een zwangerschap van minder dan 37 weken) vormen binnen deze hypothese een interessante groep. Daar waar zij namelijk normaal gesproken nog veilig in de baarmoeder zouden zitten, worden ze nu blootgesteld aan medicatie, operaties, injecties en infecties, terwijl ze moeite hebben met zelfstandig ademhalen en al hun organen nog onrijp zijn.

In Nederland werden er in 2017 11.978 kinderen te vroeg geboren, waarvan er 2.274 extreem prematuur (zwangerschapsduur <32 weken) geboren werden. Dankzij een verbeterde zorg overleven er steeds meer kinderen hun vroeggeboorte, en er is daardoor steeds meer aandacht voor de langetermijneffecten van vroeggeboorte; het is bekend dat kinderen die prematuur geboren werden een hogere kans hebben op gedrags- en groeiproblemen, en een hoger risico hebben op hart- en vaatziekten. Het is belangrijk om uit te zoeken welke factoren rondom vroeggeboorte bijdragen aan dit verhoogde risico, en het is evengoed belangrijk welke factoren hiertegen beschermen. De uitkomsten op lange termijn bij prematuren zijn waarschijnlijk het optimaalst op het moment dat de normale omstandigheden (fysiologie) zoveel als mogelijk kunnen worden nagebootst, en het is daarom belangrijk om de normale fysiologie goed te begrijpen.

In dit proefschrift hebben we een deel van deze fysiologie bij pasgeborenen onderzocht, samen met andere factoren die kunnen bijdragen aan of helpen met het voorkomen van de nadelige gevolgen van prematuriteit, waarbij de focus lag op de hormoonhuishouding (endocriene regulatie) in het vroege leven.

SAMENVATTING VAN DE RESULTATEN

Deel 1: Stresshormoonhuishouding in het vroege leven

In **hoofdstuk 2** hebben we gekeken naar de hoeveelheid stresshormonen cortisol en cortison gemeten in haar van pasgeboren baby's. We vonden dat naarmate de zwangerschapsduur langer is, de hoeveelheid stresshormoon in het haar van baby's ook toeneemt. Dit past bij een hypothese dat stijgende stresshormoonconcentraties bij de baby de bevalling in gang kan zetten. In **hoofdstuk 3** hebben we gekeken naar het verband tussen stresshormoon gemeten in haar en ervaren stress van de moeder tijdens de zwangerschap. Bij baby's zaten er minder stresshormonen in de haren als de moeder een verhoogde hoeveelheid stress had ervaren tijdens de zwangerschap, terwijl bij moeders de stresshormoonconcentratie juist hoger was. De resultaten van deze twee

onderzoeken suggereren dat stresshormoon gemeten in haar bij pasgeboren baby's kunnen helpen bij het vormen van een beeld van de stresshormoonhuishouding tijdens de zwangerschap.

Baby's worden tijdens de zwangerschap blootgesteld aan stresshormonen van moeder, maar ook na de bevalling kunnen ze in aanraking komen met de stresshormonen van moeder door middel van borstvoeding. In **hoofdstuk 4** hebben we alle tot nog toe bekende onderzoeken naar het effect van stresshormoon in borstvoeding op het kind samengevat. We kwamen echter ook tot de conclusie dat de meeste van deze onderzoeken niet optimaal zijn uitgevoerd. Onze onderzoeksgroep heeft eerder aangetoond dat de stresshormonen in moedermelk het ritme volgen dat stresshormonen ook hebben in bloed: een hoge concentratie in de (vroege) ochtend, en lage concentraties in de avond en nacht. Geen enkele van de studies die zijn gedaan naar de effecten van borstvoeding op de nakomelingen hield rekening met dit ritme. Daarnaast zijn de stresshormonen niet op de meest optimale manier bepaald. We hebben er dan ook voor gepleit om in vervolgonderzoeken rekening te houden met het ritme van stresshormonen in borstvoeding, en om de stresshormonen met een sensitieve meetmethode te bepalen.

Wij hebben zelf een onderzoek gedaan naar het effect van het ritme van stresshormonen in borstvoeding op kinderen. We hebben eerst gekeken of er een verband bestaat tussen voedingsbestanddelen (koolhydraten, eiwit en vet) en stresshormonen in melk in **hoofdstuk 5**, welke we niet gevonden hebben. Daarna hebben we in **hoofdstuk 6** gekeken naar de relatie tussen het ritme van stresshormonen in borstvoeding en in het speeksel van kinderen op de leeftijd van 1 maand. Er zijn enkele aanwijzingen dat er een verband bestaat, en dat stresshormonen in moedermelk dus mogelijk een rol spelen in de ontwikkeling van het ritme in kinderen, maar heel sterk zijn deze verbanden niet. We vonden wel dat er op groepsniveau in de kinderen een ander ritme aanwezig is dan bij volwassenen. Er werd namelijk niet alleen een piek in de ochtend gezien, maar ook in de avond. Dit ritme is mogelijk een uiting van de ontwikkeling van een volwassen stresshormoonritme vanuit een foetaal stresshormoonritme. In **hoofdstuk 7 en 8** hebben we vervolgens gekeken naar het verband tussen het ritme van stresshormoon in moedermelk en de lichaamssamenstelling, gedrag en slaap van de kinderen op de leeftijd van 3 maanden. Wij vonden geen relaties, wat erop kan wijzen dat stresshormonen in moedermelk geen significante invloed uitoefenen op kinderen. Het kan ook liggen aan het feit dat we weinig moeder-kindparen in ons onderzoek hadden en we de uitkomstmaten (lichaamssamenstelling, gedrag en slaap) op vrij jonge leeftijd hebben bepaald.

Deel 2: Stresshormoonhuishouding en sekse

Het is bekend dat de stresshormoonhuishouding verschilt tussen volwassen mannen en vrouwen. Echter, bij te vroeg geboren kinderen zijn er ook al verschillen in uitkomsten tussen jongens en meisjes; jongens hebben bijvoorbeeld een grotere kans op

overlijden. In **hoofdstuk 9 en 10** hebben we gekeken naar sekseverschillen in stres-shormoonconcentraties zonder stimulatie en bij stresstesten. Dit hebben wij gedaan door een overzicht te maken van de tot nog toe bekende literatuur (een systematisch review) en door met al die data nieuwe analyses te doen (een meta-analyse). We vonden dat er op de kinderleeftijd al verschillen zijn in zowel ongestimuleerde stresshormoon-concentraties als in concentraties tijdens en na stresstesten. Deze verschillen lijken – in ieder geval voor de ongestimuleerde concentraties – onder invloed van de puberteit te veranderen. De aanwezigheid van deze verschillen zou een deel van de sekseverschillen in uitkomsten bij vroeggeboorte kunnen verklaren. Het is ook bekend dat mannen en vrouwen een ander risicoprofiel hebben voor bepaalde aandoeningen, en de aanwezig-heid van sekseverschillen in stresshormoonhuishouding kan ook daar een (deel van een) verklaring voor zijn. Echter, het vergelijken van de studies, met name wat betreft de stresshormoonconcentraties tijdens en na stresstesten, werd bemoeilijkt door het feit dat er veel verschillende protocollen werden gehanteerd en ook de rapportage van de resultaten niet uniform was. We hebben er dan ook voor gepleit dat het belangrijk is om de stressprotocollen en rapportage van resultaten te standaardiseren.

Deel 3: Schildklierhormoonhuishouding in het vroege leven bij te vroeg geboren kinderen

Bij te vroeg geboren kinderen zijn hun organen nog onrijp. Zo ook hun schildklier: bij te vroeg geboren kinderen kan het voorkomen dat er een tijdelijke dip is in de concentratie van schildklierhormonen. Schildklierhormonen zijn belangrijk voor de hersenontwikkeling, zoals is gebleken bij kinderen die geen schildklierhormoon kunnen aanmaken, of wiens moeders tijdens de zwangerschap een lage concentratie schild-klierhormoon hebben. Het is uit eerdere onderzoeken bekend dat de tijdelijke dip in schildklierhormoon bij te vroeg geboren kinderen op de kinderleeftijd kan leiden tot problemen in de ontwikkeling. Wij hebben in **hoofdstuk 11 en 12** onderzocht of deze ontwikkelingsproblemen ook nog aanwezig zijn op 19-jarige leeftijd. Er werden geen verschillen gevonden wat betreft IQ en motorische ontwikkeling. Er werden wel wat meer gedragsproblemen gevonden: als er een dip in schildklierhormoon had plaatsge-vonden, dan hadden deze kinderen een grotere kans om zich wat meer teruggetrokken (internaliserend) te gedragen. Wij vonden dit echter geen reden om te pleiten voor screening op deze tijdelijke dip bij te vroeg geboren kinderen.

Deel 4: Groei in het vroege leven en (neurologische) ontwikkeling

Kinderen kunnen om verschillende redenen worden opgenomen op de neonatale intensive-careafdeling: ze zijn bijvoorbeeld te vroeg geboren, of hun geboortegewicht is te laag. Veel onderzoeken naar vroeggeboorte bekijken of kinderen geschikt zijn om mee te doen aan de studie op basis van zwangerschapsduur. Echter, sommige studies

laten kinderen ook meedoen op basis van geboortegewicht. Twee veelgebruikte termen bij studies naar te vroeg geboren kinderen zijn zeer prematuur (zwangerschapsduur < 32 weken) en zeer laag geboortegewicht (geboortegewicht < 1500 gram). Het is al bekend dat de uitkomsten op vroege termijn verschillen tussen deze twee termen. Wij hebben in **hoofdstuk 13 en 14** onderzocht of deze twee termen ook leiden tot verschillende uitkomsten op lange termijn. We vonden dat kinderen die zeer prematuur waren verschilden qua groeipatroon en uiteindelijke lengte met kinderen die geboren waren met een zeer laag geboortegewicht. Daarnaast werden er verschillen gevonden in de neurologische ontwikkeling: IQ en gedrag verschilden tussen deze twee termen, alhoewel er geen verschil werd gevonden qua opleidingsniveau en maatschappelijk functioneren. Wij hebben er dan ook voor gepleit om in de toekomst onderzoeken naar vroeggeboorten alleen te doen in kinderen die zijn geselecteerd op basis van zwangerschapsduur en niet op geboortegewicht. Op deze manier kan de echte impact van vroeggeboorte veel beter onderzocht worden.

Gelukkig is de zorg op de neonatale intensive-careafdelingen enorm verbeterd. Of dit ook gevolgen heeft gehad voor de groei van kinderen en hun neurologische ontwikkeling hebben we onderzocht in **hoofdstuk 15**. We vonden dat ongunstige groeipatronen (normaal geboortegewicht met groeivertraging na de geboorte, of een laag geboortegewicht zonder groeiversnelling) minder vaak voorkomen in een groep te vroeg geboren kinderen uit 2003 in vergelijking met te vroeg geboren kinderen in 1983. Echter, het verband tussen deze groeipatronen en neurologische uitkomsten was niet veranderd in die 20 jaar. Het is dus erg belangrijk om gunstige groeipatronen na te streven bij te vroeg geboren kinderen.

CONCLUSIE

In dit proefschrift hebben we gekeken naar stresshormoonhuishouding. We concludeerden dat stresshormoon gemeten in haar een goed beeld geeft van de stresshormoonhuishouding van het kind tijdens de zwangerschap. Op dat moment wordt de stresshormoonconcentratie met name beïnvloed door de zwangerschapsduur, terwijl stress van de moeder tijdens de zwangerschap ook kan worden teruggezien. Na de geboorte hebben kinderen van 1 maand oud een stresshormoonritme met 2 pieken, wat mogelijk een uiting is van de ontwikkeling van een volwassen stresshormoonritme vanuit een foetaal stresshormoonritme. Er is mogelijk een relatie tussen het ritme van stresshormonen in moedermelk en het ritme van stresshormonen in het kind, maar het stresshormoonritme in moedermelk leek geen verband te hebben met groei en gedrag van de kinderen op de leeftijd van 3 maanden. De neurologische ontwikkeling van te vroeg geboren kinderen op 19-jarige leeftijd had een verband met groei in het vroege

leven (zowel tijdens de zwangerschap als de periode erna), maar er was geen verband met de schildklierhormoonhuishouding vlak na de geboorte. Tenslotte hebben we aangeraden om onderzoeksmethoden (zoals de selectie van proefpersonen en protocollen van uit te voeren testen) te standaardiseren.

De resultaten in dit proefschrift vormen de basis voor toekomstige studies, met name wat betreft (de uitkomsten op lange termijn van) stresshormoonhuishouding bij te vroeg geboren kinderen.
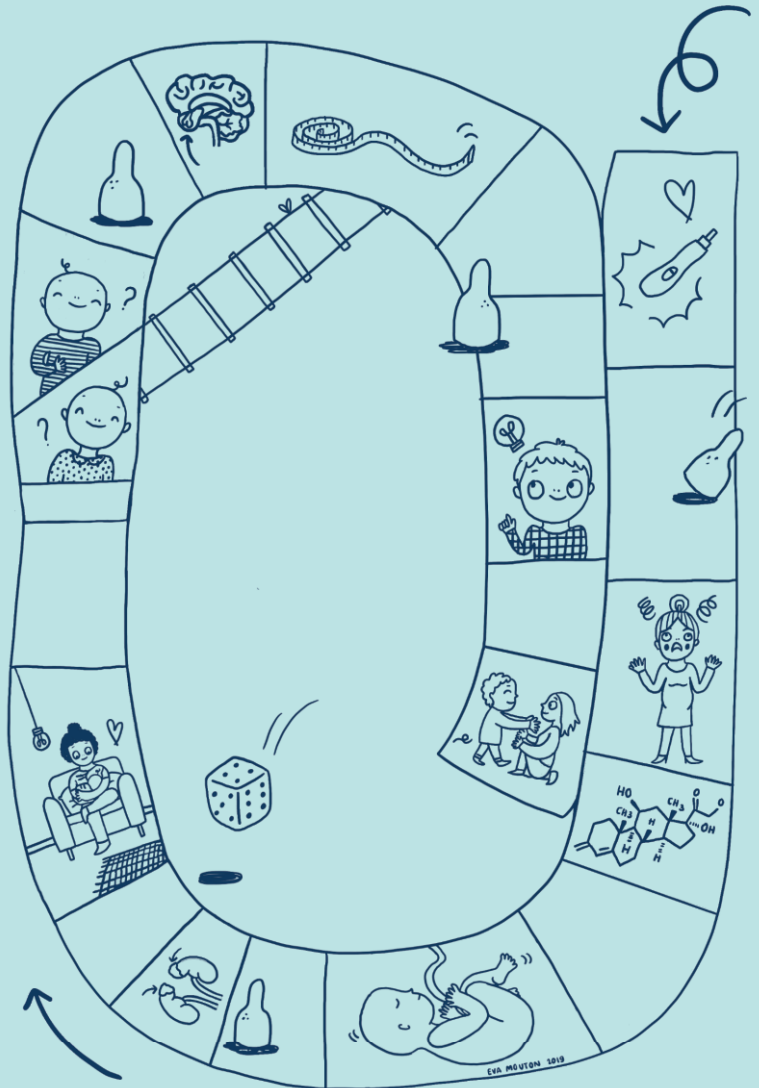
Appendices

LIST OF CO-AUTHORS

van den Akker, Erica L.T.

Department of Pediatric Endocrinology, Erasmus MC Sophia Children's Hospital, Rotterdam, the Netherlands

Bröring, Tinka

Department of Medical Psychology, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

Dijkstra, Lisette R.

Department of Pediatric Endocrinology, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

Dolman, Koert M.

Department of Psychiatry Obstetrics and Pediatrics (POP), OLVG, Amsterdam, The Netherlands

van Dommelen, Paula

TNO Child Health, Leiden, The Netherlands

Finken, Martijn J.J.

Department of Pediatric Endocrinology, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

de Goede, Paul

Laboratory of Endocrinology, Amsterdam UMC, Amsterdam Gastroenterology & Metabolism, Amsterdam, The Netherlands
Netherlands Institute for Neuroscience (NIN), Royal Dutch Academy of Arts and Sciences (KNAW), Amsterdam, Netherlands

van Goudoever, Johannes B.

Department of Pediatrics, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

Heijboer, Annemieke C.

Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC, Amsterdam, the Netherlands

Honig, Adriaan

Department of Psychiatry Obstetrics and Pediatrics (POP), OLVG, Amsterdam, The Netherlands

Israëls, Joël

Department of Pediatrics, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

Kalsbeek, Andries

Laboratory of Endocrinology, Amsterdam UMC, Amsterdam Gastroenterology & Metabolism, Amsterdam, The Netherlands
Netherlands Institute for Neuroscience (NIN), Royal Dutch Academy of Arts and Sciences (KNAW), Amsterdam, Netherlands

Ket, Johannes C.F.

Medical Library, Vrije Universiteit, Amsterdam, The Netherlands

Kieviet, Noera

Department of Psychiatry Obstetrics and Pediatrics (POP), OLVG, Amsterdam, The Netherlands

Kouwenhoven, Stefanie M.P.

Department of Pediatrics/Neonatology, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

Lafeber, Harrie N.

Department of Pediatrics/Neonatology, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

van de Lagemaat, Monique

Department of Pediatrics/Neonatology, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

Oosterlaan, Jaap

VU University Amsterdam, Department of Clinical Neuropsychology, Amsterdam, The Netherlands.

van der Pal, Sylvia M.

TNO Child Health, Leiden, The Netherlands

de Rijke, Yolanda B.

Department of Clinical Chemistry, Erasmus MC University Medical Center, Rotterdam, The Netherlands.

van Rossum, Elisabeth F.C.

Department of Internal Medicine, Division of Endocrinology, Erasmus MC University Medical Center, Rotterdam, the Netherlands.

Rotteveel, Joost

Department of Pediatric Endocrinology, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

Ruys, Charlotte A.

Department of Pediatrics/Neonatology, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

Schaëfer, Nina

Department of Pediatric Endocrinology, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

van Schie, Petra E.M.

Department of Rehabilitation Medicine, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

Toorop, Alyssa A.

Department of Pediatric Endocrinology, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands

Verkerk, Paul H.

TNO Child Health, Leiden, The Netherlands

van der Voorn, Bibian

Department of Paediatric Endocrinology, Obesity Center CGG, Sophia Children's Hospital, Rotterdam, The Netherlands

ABBREVIATIONS

11β-HSD1/2	11β-hydroxysteroid dehydrogenase type 1/2
95% CI	95% confidence interval
ACTH	adrenocorticotropin hormone
ADHD	attention deficit/hyperactivity disorder
ADP	air-displacement plethysmography
AGA	appropriate-for-gestational-age
AUC i/g	Area Under the Curve increase/ground
BMI	body mass index
BW	birth weight
CA	corrected age
CAR	cortisol awakening response
CBCL	Child Behavior Checklist
CBG	corticosteroid binding globulin
CBS	The Netherlands Central Bureau of Statistics
CHT	congenital hypothyroidism
CRH	corticotropin-releasing hormone
CUG	catch-up growth
CV	coefficients of variation
CosMos	Cortisol in Mother's Milk
DHEAS	dehydroandrostenedione
DOHaD	Developmental Origins of Health and Disease
eFSIQ	estimated Full Scale Intelligence Quotient
ER	estrogen receptor
EUGR	extrauterine growth restriction
FFM	fat free mass
FFMI	fat free mass index
FM	fat mass
FMI	fat mass index
GA	gestational age
GC	glucocorticoid
GEE	Generalized Estimating Equation
GR	glucocorticoid receptor
GR	growth restriction
HADS	Hospital Anxiety and Depression Scale
HAS	Hospital Anxiety Subscale
HC	head circumference
HDS	Hospital Depression Subscale

HELLP	Hemolysis, Elevated Liver enzymes and Low Platelet count
HPA	hypothalamus-pituitary-adrenal
HPG	hypothalamus-pituitary-gonadal
HPLC	high-performance liquid chromatography
HPT	hypothalamus-pituitary-thyroid
HUI3	Health Utilities Index Mark 3
IBQ	Infant Behavior Questionnaire
IQ	intelligence quotient
IQR	interquartile range
IUGR	intrauterine growth restriction
LC-MS/MS	liquid chromatography – tandem mass spectrometry
LHS	London Handicap Scale
LLoQ	Lower Limit of Quantification
LMM	Linear Mixed Model
MC	mineralocorticoid
MCT(-IL)	Multicultural Capacity Test(–Intermediate Level)
MR	mineralocorticoid receptor
NaSSAs	noradrenergic or specific serotonin antidepressants
NICU	neonatal intensive care unit
NTISS	Neonatal Therapeutic Intervention Scoring System
OPV	outpatient visit
OR	odds ratio
POP	psychiatric obstetric pediatric
POPS	Project On Preterm and Small-for-gestational-age infants
pp	postpartum
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
RCT	randomized controlled trial
SADs	selective antidepressants
SCN	suprachiasmatic nuclei
SD(S)	standard deviation (score)
SEM	standard error of the mean
SES	socio-economic status
SGA	small-for-gestational-age
SNRIs	serotonin-norepinephrine reuptake inhibitors
SSRIs	selective serotonin reuptake inhibitors
STEP	Study Towards the Effects of Postdischarge nutrition'
T4	thyroxine
TSH	thyroid stimulating hormone
TSST(-c)	Trier Social Stress Test (for children)

THoP	transient hypothyroxinemia of prematurity
VLWB	very low birth weight
VP	very preterm
YABCL	Young Adult Behavior Checklist
YASR	Young Adult Self Report

LIST OF PUBLICATIONS

Long-term neurodevelopmental and functional outcomes of infants born very preterm versus with a very low birth weight. Hollanders JJ, Schaëfer N, van der Pal SM, Oosterlaan J, Rotteveel J, Finken MJJ; Dutch POPS-19 Collaborative Study Group. Neonatology. 2019;115(4):310-319.

Early-life growth of preterm infants and its impact on neurodevelopment. Ruys CA, *Hollanders JJ*, Bröring T, van Schie PEM, van der Pal SM, van de Lagemaat M, Lafeber HN, Rotteveel J, Finken MJJ. Pediatr Res. 2019 Feb;85(3):283-292.

The association between breastmilk glucocorticoid concentrations and macronutrient contents throughout the day. *Hollanders JJ*, Kouwenhoven SMP, van der Voorn B, van Goudoever JB, Rotteveel J, Finken MJJ. Nutrients. 2019 Jan 24;11(2).

Maternal stress during pregnancy is associated with decreased cortisol and cortisone levels in neonatal hair. van der Voorn B, *Hollanders JJ*, Kieviet N, Dolman KM, de Rijke YB, van Rossum EFC, Rotteveel J, Honig A, Finken MJJ. Horm Res Paediatr. 2018 Dec 12:1-9.

Interpretation of glucocorticoids in neonatal hair: a reflection of intrauterine glucocorticoid regulation? *Hollanders JJ*, van der Voorn B, Kieviet N, Dolman KM, de Rijke YB, van den Akker ELT, Rotteveel J, Honig A, Finken MJJ. Endocr Connect. 2017 Nov;6(8):692-699.

Nutritional programming by glucocorticoids in breast milk: Targets, mechanisms and possible implications. *Hollanders JJ*, Heijboer AC, van der Voorn B, Rotteveel J, Finken MJJ. Best Pract Res Clin Endocrinol Metab. 2017 Aug;31(4):397-408.

Growth pattern and final height of very preterm vs. very low birth weight infants. *Hollanders JJ*, van der Pal SM, van Dommelen P, Rotteveel J, Finken MJJ. Pediatr Res. 2017 Aug;82(2):317-323.

Is HPA axis reactivity in childhood gender-specific? A systematic **review.** *Hollanders JJ*, van der Voorn B, Rotteveel J, Finken MJJ. Biol Sex Differ. 2017 Jul 11;8(1):23.

Gender-specific differences in hypothalamus-pituitary-adrenal axis activity during childhood: a systematic review and meta-analysis. van der Voorn B, *Hollanders JJ*, Ket JCF, Rotteveel J, Finken MJJ. Biol Sex Differ. 2017 Jan 19;8:3.

Transient hypothyroxinemia of prematurity and problem behavior in young adult-hood. *Hollanders JJ*, van der Pal SM, Verkerk PH, Rotteveel J, Finken MJ; Dutch POPS-19 Collaborative Study Group. Psychoneuroendocrinology. 2016 Oct;72:40-6.

No association between transient hypothyroxinemia of prematurity and neurode-velopmental outcome in young adulthood. *Hollanders JJ*, Israëls J, van der Pal SM, Verkerk PH, Rotteveel J, Finken MJ; Dutch POPS-19 Collaborative Study Group. J Clin Endocrinol Metab. 2015 Dec;100(12):4648-53.

Not in this thesis

Programming of the hypothalamus-pituitary-adrenal axis by very preterm birth. Finken MJJ, van der Voorn B, *Hollanders JJ*, Ruys CA, de Waard M, van Goudoever JB, Rotteveel J. Ann Nutr Metab. 2017;70(3):170-174.

Cortisol in human milk: The good, the bad, or the ugly? Finken MJJ, van der Voorn B, *Hollanders JJ*, Dijkstra LR, Toorop AA, Rotteveel J. Obesity (Silver Spring). 2017 Jul;25(7):1153.

CURRICULUM VITAE

Jonneke Hollanders was born on 7 October 1988 in Delft, The Netherlands. She attended secondary school at the Vincent van Gogh school in Assen from 2001 to 2005, after which she moved to Houston, USA. There she obtained her high school and (bilingual) International Baccalaureate diplomas at the Awty International School. In 2007, she returned to the Netherlands, where she started medical school at the VU University in Amsterdam. She completed both her bachelor and master degree with honors (*cum laude*). During her internships, she became particularly interested in pediatrics. She therefore did her final clinical internship and her research internship within this field. She also did a tropical pediatric internship in the St. Francis Referral Hospital in Ifakara, Tanzania. After finishing her studies in 2014, she worked as a pediatric resident not in training (*ANIOS*) at the Medisch Centrum Haaglanden (currently Haaglanden Medisch Centrum) for a year. After this, she had the opportunity to start as a PhD candidate in pediatric endocrinology at the VU University Medical Center (currently Amsterdam UMC, location VUMC) under the supervision of Prof. Dr. Hans van Goudoever, dr. Martijn Finken en dr. Joost Rotteveel. She worked on her thesis from 2015 to 2019, during which period she gave several poster and oral presentations at both national and international conferences, for some of which she received awards and/or travel grants. She also helped organize the "Amsterdam Kindersymposium" in 2017, together with 8 other PhD students. She started her training to become a child healthcare doctor (*jeugdarts*) in September 2019. She lives in Leiden with her husband and daughter.

DANKWOORD

Hier is 'ie dan! Heel wat passie, liefde en domweg jaren van hard werken zijn er gestopt in dit proefschrift. Het voelt heel fijn om deze eindstreep gehaald te hebben! Dat was natuurlijk niet mogelijk zonder alle steun die ik van heel veel mensen gehad heb. (Mocht ik iemand vergeten zijn, mijn oprechte excuses; ik wijd het aan de euforie van dit moment!)

Mijn promotor, prof.dr. Hans van Goudoever. Hans, dank voor je steun in de afgelopen jaren en de vrijheid die ik van je kreeg om mijn boekje tot dit resultaat te kunnen brengen.

Mijn copromotoren, Martijn Finken en Joost Rotteveel. Martijn, jouw wetenschappelijk inzichten, passie voor het onderzoek en het vak, en alle tijd die je in mij en dit boekje gestoken hebt: dank je wel daarvoor! Joost, enorm bedankt voor alle fijne gesprekken die we gehad hebben en jouw relativeringsvermogen. Jullie waren een fijn team om mee samen te werken!

Dank ook aan de leescommissie voor hun tijd om dit boekje te lezen en te beoordelen: prof.dr. de Groot, prof.dr. Gemke, prof.dr. Oosterlaan, prof.dr. van Rossum, prof.dr. van Trotsenburg, prof.dr. van Kaam, en prof.dr. de Vries. Het merendeel van hen zit ook in de oppositie, bedankt voor het nogmaals vrijmaken van jullie tijd.

Bibian, endo-maatje, zonder jou zou mijn boekje een stuk dunner zijn. Dank je wel voor het toevertrouwen van CosMos3 aan mij; ik hoop dat het resultaat iets is waar ook jij trots op bent. Dank je wel voor je luisterende oor, je wetenschappelijke inzichten, de gezellige congressen, en je vriendschap.

Ik zou dit boekje niet kunnen hebben geschreven zonder alle co-auteurs. Ten eerste mijn (ex-)kamergenootjes Charlotte en Stefanie. Dank jullie wel voor het fijne samenwerken! Ook aan mijn mede-promovendi en collega's Noera, Paul en Joël, super bedankt voor jullie bijdrages aan dit boekje! Prof. Oosterlaan, prof. van Rossum, prof. Honig en prof. Kalsbeek: dank voor jullie wijze inzichten en expertise. Dr. de Rijke en dr. van den Akker, bedankt voor het meedenken met en meeschrijven aan de haarcortisol stukken. Dr. Dolman, dank voor het meedenken en praktisch ondersteunen van CosMos3, alsmede voor het brainstormen over en meeschrijven aan de haarcortisol stukken. Alle POPS-artikelen zouden niet zo mooi geworden zijn zonder hulp vanuit TNO. Als eerste natuurlijk Sylvia van der Pal: vanaf mijn wetenschappelijke stage al betrokken bij mijn onderzoek. Dank voor de fijne inzichten en de tijd die je voor mijn stukken maakte, ondanks het feit dat je die tijd niet echt had. Verder natuurlijk ook nog dank aan dr. Verkerk en dr. van Dommelen voor hun inzichten. Het POPS-STEP stuk had niet kunnen worden geschreven zonder de inzichten en bijdrages van Monique van de Lagemaat, Tinka Bröring en Petra van Schie, dank jullie wel!

Voor CosMos heb ik heel wat uren doorgebracht in het endo-lab. Dank aan alle medewerkers daar voor de warme ontvangst, de interesse in ons onderzoek en alle hulp. In het bijzonder wil ik Anneke Frans, Frans Martens en Annemieke Heijboer bedanken voor hun hulp en het meedenken met de analyses. Ook wil ik Annemieke bedanken voor haar bijdrages aan meerdere van mijn manuscripten, en in het bijzonder het moedermelk review.

Ook al is het werk dat ik heb verricht aan de GBA analyses niet in dit proefschrift verschenen, alsnog wil ik het metabool lab bedanken voor hun behulpzaamheid, geduld en bijdrages aan het werk dat ik bij hen gedaan heb. Met name aan Warsha en Gajja een dikke dank je wel, voor alle tijd en alle hulp die ik gekregen heb van jullie.

Met plezier naar werk gaan, daar zijn ook je collega's heel belangrijk voor. Het grootste gedeelte van mijn promotieonderzoek heb ik doorgebracht op PK4X, en mijn collega's daar wil ik bedanken voor alle gezellige lunches, verjaardagen, het sparren en de gezelligheid. In het bijzonder wil ik Sandra, Lucie, Dominique, Anne, Laura, Fatma, Stijn, Sophie, Francis, Marloes, Katie, Diane, en Eline H. bedanken. Uiteraard ook een hele dikke knuffel aan iedereen op kamer 33; we hebben samen heel wat ups en downs beleefd: Dana, Kim, Lindsay, Marita, Charlotte en Stefanie. Mirjam, we hebben heel wat uren weg gekletst; jouw aanwezigheid was een geweldige steun voor me.

Kamer 9D36 was mijn laatste toevluchtsoord op de VU, en ik wil graag mijn kamergenootjes heel erg bedanken voor de enorm fijne tijd en de gigantische goede afsluiting van mijn promotietraject. Sofia en Thomas: dank je wel voor de gezelligheid, succes met alle poepjes! Michelle, we zijn maar kort collega's geweest, maar het was in die tijd wel heel leuk. Succes met PredictBPD, ik blijf graag op de hoogte! Britt, we hebben heel wat afgekletst. Het was heerlijk om jou als directe collega te hebben, om te sparren en te klagen en te lachen. 't Gaat helemaal goedkomen met jou! Dank je wel voor alle gezelligheid en het luisterende oor, en natuurlijk heel veel succes met alles!

Dan, dank je wel aan Nina en Daisy, de studenten die mijn onderzoek ondersteund hebben. En natuurlijk de honors program studenten: Alyssa en Lisette. Wat was het leuk om jullie te zien groeien in jullie wetenschappelijke rol. Het was enorm fijn samenwerken met jullie, met een heel mooi resultaat!

Eva en Bert, dank jullie wel voor de prachtige omslag, en het fijne samenwerken. Hij is precies zoals ik had gehoopt dat 'ie zou zijn!

Het organiseren van het AKS was een welkome afwisseling van het onderzoek doen. Dank jullie wel voor de fijne tijd: Bart, Annemarie, Kim, Mirjam, Lindsay, Ilse, Bas, en Irene.

Dank ook aan GGD Hollands Midden en TNO voor jullie warme welkom. Jullie interesse in mijn promotie wordt erg gewaardeerd!

Lieve paranimfen, dank jullie wel dat jullie me willen begeleiden tijdens dit (belangrijke) staartje van mijn promotie. Jojanneke, studiemaatje van het eerste uur! Zo zie je maar dat eerste indrukken absoluut niet hoeven te kloppen. We bewandelen hele

andere, maar toch best op elkaar lijkende wegen, en het is geweldig om jou er in deze rol bij te hebben. Elise, jij snapt me helemaal! Wat heerlijk om een vriendin te hebben die niet alleen over heel veel dingen hetzelfde denkt, maar ook heel veel hetzelfde doet. Onze levens hebben heel wat jaren eng veel op elkaar geleken, en ik ben heel blij dat jij nu net als ik iets doet waar je helemaal je ei in kwijt kan!

Ik heb echt een heleboel fijne vrienden om me heen, die me gedurende deze jaren enorm veel ontspanning en afleiding hebben gegeven. Lieve Mirella, jij bent en blijft de beste! Vriendschap is zeker niet onvoorwaardelijk, maar stiekem ben je zo langzamerhand gewoon m'n zus(je). Heel fijn om jou in mijn leven te hebben. Alle lieve mensen in Gent, en in het bijzonder Kate, Dennis, Teresa, en Nicolas: dank jullie wel! De Rode Dobbelsteners, en met name Anne, Mathijs, Inge, Abel, Maaike, Aldo, en Daniël: dank jullie wel voor de heerlijke ontspannen avonden, weekends en zelfs vakanties. Mijn lieve vriendinnetjes van BWFNL, en Tia natuurlijk ook: best wishes and DFTBA! En dan natuurlijk nog alle andere vrienden die mijn leven opfleuren, onder andere Joost, Lara, Anja, en nog een hoop anderen: dank jullie wel!

Bas, Betty, en Hanneke, ook jullie moeten natuurlijk bedankt worden, want zonder jullie zou ik niet staan waar ik nu sta. Dank jullie wel voor alle lieve en goede zorgen toen ik jullie weekenddochter mocht zijn. Griekse salades, roeren in potten, films, spelletjes en gewoon het feit dat ik bij jullie terecht kon: jullie waren echt mijn thuisbasis tijdens mijn studie. En natuurlijk, ook bedankt aan mijn weekendbroertjes en -zusjes: Helen, Luka, Siebe, Annemarije en Daan!

Oma Zwaan en oma Hollanders, dank voor jullie eeuwige geloof in mij!

Lieve schoonfamilie: ik heb toch maar geluk met jullie! Dank voor de eeuwige gastvrijheid, de vakanties, de lol en de steun van de afgelopen jaren. Lieve Nadine, zo fijn dat ik altijd even met je kan sparren! Chris, jouw interesse in alles wat ik en wij doen is geweldig. Sanne en Kristof, dank voor alle gezellige dagen samen! Femke en Dieter, gesprekken met jullie zijn altijd of gezellig, of inzichtgevend, of inspirerend, dank jullie wel. En last but not least, Suus en Nand, jullie zijn de geweldigste neefjes die ik me kan wensen!

Lieve Elsje: dank je wel voor je eeuwige luisterende oor. We zijn nogal verschillend, jij en ik, maar ik ben zo blij dat jij er altijd voor me bent. Hopelijk vice versa ook. En ik ben zo trots op jou en waar je nu staat! Thijs, lief broertje, ik denk dat je zo langzamerhand doorhebt hoe het is om een PhD te doen. Onthoud maar: promoveren is wél leuk! Dank je wel voor je steun tijdens het werken aan de mijne.

Lieve papa en mama. Er is niet iets wat ik kan opschrijven dat alles kan verwoorden wat ik tegen jullie zou willen zeggen. Dank jullie wel voor het altijd prikkelen, voor het altijd aanwakkeren van de nieuwsgierigheid, voor alle kansen die jullie me hebben gegeven en voor het steunen van mijn keuzes. Zonder jullie had ik hier niet gestaan. Ik hou van jullie!

En als laatste, lieve lieve lieve Bouke en Liene. Er zijn maar een paar constanten in deze wereld; voor mij zijn jullie er daar een van. Thuis komen is en blijft een feestje. Lieve Bouke, dank je wel voor alle steun de afgelopen jaren; jij hebt me overeind gehouden, en zonder jou had ik het echt niet tot hier kunnen brengen! Lieve Liene, jij plaatst alles in perspectief. Blijf net zo ondeugend, lief, aandachtig en nieuwsgierig als dat je nu bent, dan komt sowieso alles goed. Ik hou hou hou hou HOU van jullie!