American Journal of Epidemiology

DOI: 10.1093/aje/kww172

Original Contribution

Maternal Licorice Consumption During Pregnancy and Pubertal, Cognitive, and Psychiatric Outcomes in Children

Katri Räikkönen*, Silja Martikainen, Anu-Katriina Pesonen, Jari Lahti, Kati Heinonen, Riikka Pyhälä, Marius Lahti, Soile Tuovinen, Karoliina Wehkalampi, Sara Sammallahti, Liisa Kuula, Sture Andersson, Johan G. Eriksson, Alfredo Ortega-Alonso, Rebecca M. Reynolds, Timo E. Strandberg, Jonathan R. Seckl, and Eero Kajantie

* Correspondence to Dr. Katri Räikkönen, Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, PO Box 9 (Siltavuorenpenger 1A), 00014 University of Helsinki, Helsinki, Finland (e-mail: katri.raikkonen@helsinki.fi).

Initially submitted February 17, 2016; accepted for publication September 19, 2016.

Earlier puberty, especially in girls, is associated with physical and mental disorders. Prenatal glucocorticoid exposure influences the timing of puberty in animal models, but the human relevance of those findings is unknown. We studied whether voluntary consumption of licorice, which contains glycyrrhizin (a potent inhibitor of placental 11β-hydroxysteroid dehydrogenase type 2, the "barrier" to maternal glucocorticoids), by pregnant women was associated with pubertal maturation (height, weight, body mass index for age, difference between current and expected adult height, Tanner staging, score on the Pubertal Development Scale), neuroendocrine function (diurnal salivary cortisol, dexamethasone suppression), cognition (neuropsychological tests), and psychiatric problems (as measured by the Child Behavior Checklist) in their offspring. The children were born in 1998 in Helsinki, Finland, and examined during 2009–2011 (mean age = 12.5 (standard deviation (SD), 0.4) years; $n = 378$). Girls exposed to high maternal glycyrrhizin consumption (≥500 mg/week) were taller (mean difference (MD) = 0.4 SD, 95% confidence interval (CI): 0.1, 0.8), were heavier (MD = 0.6 SD, 95% CI: 0.2, 1.9), and had higher body mass index for age (MD = 0.6 SD, 95% CI: 0.2, 0.9). They were also 0.5 standard deviations (95% CI: 0.2, 0.8) closer to adult height and reported more advanced pubertal development ($P < 0.04$). Girls and boys exposed to high maternal glycyrrhizin consumption scored 7 (95% CI: 3.1, 11.2) points lower on tests of intelligence quotient, had poorer memory ($P < 0.04$), and had 3.3-fold (95% CI: 1.4, 7.7) higher odds of attention deficit/hyperactivity disorder problems compared with children whose mothers consumed little to no glycyrrhizin (≤249 mg/week). No differences in cortisol levels were found. Licorice consumption during pregnancy may be associated with harm for the developing offspring.

ADHD; 11β-hydroxysteroid dehydrogenase type 2; cognition; glucocorticoids; glycyrrhizin; puberty

Abbreviations: 11β-HSD2, 11β-hydroxysteroid dehydrogenase type 2; BMI, body mass index; CI, confidence interval; HPA, hypothalamic-pituitary-adrenal; IPCW, inverse-probability-of-censoring weight; IQ, intelligence quotient; MD, mean difference; SD, standard deviation.

Editor's note: *An invited commentary on this article appears on page 000, and the authors' response appears on page 000.*

Earlier puberty, especially in girls, is associated with physical and mental disorders, including hormone-sensitive cancers, cardiometabolic disorders, and depression (1–3). According to the "developmental origins hypothesis" of health and disease, prenatal exposure to environmental adversity "programs" long-lasting structural and functional changes in key organs and homeostatic systems, including the brain and its control of behaviors and reproductive function (4, 5).

Overexposure of the fetus to maternal glucocorticoids is probably a key mechanism underlying these "programming" events. While maternal and fetal cortisol levels are correlated (6), fetal

levels are up to 10 times lower. This gradient is due to the placental "barrier" enzyme 11β-hydroxysteroid dehydrogenase type 2 (11β-HSD2) (7–9). This enzyme catalyzes metabolism of 80%–90% of active cortisol in the maternal bloodstream to inactive cortisone during passage to the fetal circulation (8, 9).

Animal models of fetal glucocorticoid overexposure include genetic lack of fetoplacental 11β-HSD2, its inhibition by glycyrrhizin (glycyrrhetinic acid and glycyrrhizic acid—natural constituents of licorice) or by the synthetic analog carbenoxolone, maternal stress, maternal high- or low-protein diet, or bypass with poor-substrate glucocorticoids such as dexamethasone or betamethasone (7–13). In these models the offspring have alterations in timing of puberty, reproductive function, and hypothalamic-pituitary-adrenal (HPA) axis function as well as having cardiometabolic risk factors, attenuated learning and memory, and increased anxiety- and depression-like behaviors (7–13).

Animal and human placentation and gestation, however, differ, as does the ontogeny of 11β-HSD2 in the placenta (14), precluding predictions applicable to humans. Using a dual-circuit immediate ex vivo perfusion in fresh, intact, term human placentas, we demonstrated that very low doses (10^{-8} M) of glycyrrhetinic acid, coperfused in the maternal circulation with cortisol, resulted in near complete inhibition of placental cortisone production and free passage of cortisol to the fetal circulation (9). Because glycyrrhizin also inhibits maternal 11β-HSD2, it increases cortisol access to renal mineralocorticoid receptors, sometimes causing maternal hypertension. Although other actions of glycyrrhizin include inhibition of 15-hydroxyprostaglandin dehydrogenase, this probably does not occur systemically with licorice consumption in vivo (15).

We have studied the offspring of pregnant women from a unique longitudinal cohort of Finnish mothers and children, in which nearly 50% of mothers reported licorice consumption during pregnancy (16–18). At the age of 8 years, children of women who consumed high amounts of licorice during pregnancy, when compared with children of women who consumed low amounts of licorice or none, scored lower on tests measuring intelligence and memory (17), had an increased risk of externalizing behavior problems (17), and had higher HPA-axis activity upon awakening and during psychosocial stress (18). We examined these children in adolescence to explore associations with pubertal maturation and to learn whether the cognitive, behavioral, and neuroendocrine changes persist.

METHODS

Study design and participants

The participants were from an urban community-based cohort of 1,049 women and their healthy, singleton infants born in 1998 in Helsinki, Finland (16). Between 2009 and 2011, all 920 cohort members who had given permission to be contacted and whose addresses were traceable were invited to participate in a follow-up study; 692 (75.2%) were contacted by phone, and 451 (65.2% of the women who were contacted) participated.

Of the follow-up participants, 327 of the children were prenatally exposed to zero–low maternal consumption of glycyrrhizin (≤249 mg/week; mean = 47 mg/week) and 51 to high maternal consumption (≥500 mg/week; mean = 845 mg/week). They represented 48.0% (of 681) and 56.7% (of 90) of those invited, respectively ($P = 0.12$ for difference in participation rates). The children were assessed at follow-up visits held during 2009–2011. At the follow-up visit, children exposed to zero–low amounts of glycyrrhizin had a mean age of 12.5 (standard deviation (SD), 0.4) years, and children exposed to high amounts of glycyrrhizin had a mean age of 12.5 (SD, 0.4) years ($P = 0.26$ for group difference). For girls, mean difference (MD) for age in years was −0.15 (95% confidence interval (CI): −0.32, 0.01) between the exposure groups; for boys, MD = 0.04 (95% CI: −0.14, 0.21) between the exposure groups. Among those exposed to zero–low and high amounts of glycyrrhizin and who were invited ($n = 771$), nonparticipation ($n = 393$) was related to older maternal age (MD = 0.72 years, 95% CI: 0.03, 1.40), higher level of maternal stress (on a 100-mm visual analog scale, MD = 5.3 mm, 95% CI: 1.7, 9.0), and weekly coffee consumption during pregnancy (MD = 1.3 cups/week, 95% CI: 0.3, 2.4). However, it was not related to the other 25 variables that we tested, including maternal weekly licorice consumption, chocolate consumption, cacao consumption, tea consumption, salt consumption, alcohol consumption, smoking status, blood pressure during pregnancy, maternal or paternal weight, maternal height or body mass index (BMI) at delivery, or delivery mode, nor was it related to the offspring's birth order, body size at birth, length of gestation, or Apgar score at 1 minute ($P > 0.08$ for all).

Due to this slightly differential dropout, we computed stabilized inverse-probability-of-censoring weights (IPCWs) (19) using the aforementioned characteristics associated with nonparticipation. These IPCWs were subsequently included in all statistical models in order to reduce the possibility of selection bias affecting our results. The assumptions of positivity and misspecification of the weights were checked as recommended (20).

The study protocol was approved by the ethical committees of the City of Helsinki and the Uusimaa Hospital District. Written informed consent was obtained from the mothers after delivery and from the parent/guardian and adolescent at the follow-up assessment.

Maternal licorice consumption

While in the maternity ward, the mothers reported the brand(s) and weekly consumption of licorice during pregnancy (glycyrrhizin intake calculated as mg/week), using a list of all brands of licorice-containing confectionery available in Finland in 1998 (16). The list was prepared by the National Food Administration in 1993 and updated with information from manufacturers (16). The zero–low-exposure (0–249 mg/week) and high-exposure (≥500 mg/week) groups comprised 75 percent and 11 percent of births in the initial cohort (16).

Pubertal maturation

Estimation of pubertal maturation was based on 3 measures of growth and development. 1) The difference was taken between the child's height-for-age standard-deviation score (21) (using the current measured height without shoes, measured with a Seca stadiometer) (Model 213; Seca GmbH & Co KG, Hamburg, Germany) and the standard-deviation score of the midparental target height (22); this is a measure of remaining growth potential and, consequently, the timing of the pubertal growth spurt. 2) The Tanner staging questionnaire was administered by a research nurse; using schematic drawings of 2 secondary sex characteristics (pubic hair development in girls and boys and breast development in girls or development of genitalia in boys), the examiner derives two 5-stage scores ranging from prepubertal (stage I) to postpubertal (stage V) (23). 3) The Pubertal Development Scale is a self-report on secondary sex characteristics (growth spurt, body hair that is not specifically pubic hair, and skin changes in girls and boys; menarche and breast development in girls; and facial hair and voice change in boys) and yields one 4-stage score ranging from no development (1) to full completion of development (4) (24).

We also measured weight in light clothing without shoes (Model 8; Seca GmbH & Co KG) and calculated BMI (weight (kg)/height $(m)^2$). We transformed the values into weight-for-age and BMI-for-age standard-deviation scores (21).

Cognition

The cognitive test battery comprised tests of intelligence, memory and learning, social perception, attention, and executive function. The short form of the Wechsler Intelligence Scale for Children III (25) included vocabulary, similarities, block design, and picture arrangement subtests, for which we used age-standardized scores as well as estimated age-standardized total intelligence and verbal and performance intelligence quotients (IQs) (26). For the Developmental Neuropsychological Assessment for children (27), we used age-standardized scores for word generation, design fluency, free and cued narrative memory, memory of names, memory of faces, and theory of mind subtests. Measures of attention and executive function included the Continuous Performance Test II (28), the Trail Making Test (29), and the Wisconsin Card Sorting Test (30).

Psychiatric problems

Mothers completed the Child Behavior Checklist, a standardized and validated rating scale screening for psychiatric problems (31). We calculated scores for the scales oriented to the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, and used the 82nd percentile as the cutoff to identify adolescents with borderline clinically significant problems (31).

HPA-axis activity

Samples of saliva were collected using cotton swabs (Salivette; Sarstedt, Nümbrecht, Germany). On the first of the 2 consecutive days, samples were collected upon awakening and 15, 30, 45, and 60 minutes thereafter, at 12:00 noon, at 5:00 p.m., and at bedtime. Dexamethasone was administered after the bedtime saliva sample, and a sample was collected upon awakening the next day. We used a low dose of dexamethasone (3 μg/kg of total body weight) to attempt to detect individual variation in HPA-axis suppression (32).

Salivary cortisol concentrations were determined by solid-phase, time-resolved fluorescence immunoassay with fluorometric end-point detection (DELFIA; Wallac, Turku, Finland). The intraassay and interassay coefficients of variation varied between 4.0% and 7.7%, and the mean coefficient of variation between duplicate analyses was 5.9%.

Covariates and confounders

Covariates and confounders included the adolescent's age (years), highest educational level of either parent (secondary or less, vocational, university) reported at follow-up, gestational length (weeks) as confirmed by ultrasonography, birth weight (grams) of the adolescent as derived from birth records, maternal age (years) and BMI, calculated from weight and height derived from medical records, maternal smoking status (none, 1–10 cigarettes/day, >10 cigarettes/day), maternal weekly alcohol consumption (no, yes, no answer; g/week for those reporting yes), coffee consumption (cups/week), tea consumption (cups/week), cacao consumption (cups/week), salt consumption (no, yes, low-sodium), chocolate consumption (never, seldom, weekly, daily), and stress (measured using a 100-mm visual analog scale) during pregnancy, which was reported while the mother was on the maternity ward. They also included the adolescent's own licorice consumption (never, less than once a week, once a week, 2–4 days a week, daily, no answer), reported in the follow-up assessment.

In addition, we conducted analyses of pubertal maturation adjusting for maternal self-reported age at menarche (years) as a crude proxy of the genetic component of pubertal development and analyses of HPA-axis for time at awakening and time at dexamethasone intake.

Statistical analyses

Differences between the zero–low- and high-exposure groups in the continuous, categorical, and binary outcomes were tested with and without covariates (described in the Results section), using generalized linear models specifying a Gaussian, ordinal logistic, or binary logistic reference distribution for quantitative, ordinal, or binary responses, respectively. All models were weighted by IPCW and computed using robust estimators, which provides consistent covariance estimates even in the case of misspecification of the variance and link functions. By using this analytical strategy, we also tested whether glycyrrhizin intake as a continuous variable (natural log–transformed, to improve linear model fitting)—among those who reported consumption of at least some licorice—was associated in a dose-response manner with the continuous, categorical, and binary outcomes. All P values were 2-sided.

Analyses of pubertal development were conducted separately in girls and boys because adolescent girls are, on

average, ahead of boys in pubertal development. Other outcomes were tested in girls and boys combined because sex × exposure-level group interactions were not significant in any analyses (P values > 0.05).

RESULTS

The zero–low- and high-exposure groups did not differ significantly in terms of any maternal, paternal, neonatal, or adolescent background characteristics ($P \geq 0.07$ for all; Table 1).

Pubertal maturation

Table 2 shows that in models that did not adjust for covariates, girls in the high-exposure group had significantly higher height-, weight-, and BMI-for-age standard-deviation scores, and their current measured height-for-age standard-deviation scores were closer to the standard-deviation scores for midparental target height. They also reported a more advanced Tanner stage of pubic hair and breast development and a more advanced pubertal stage on the Pubertal Development Scale than did girls in the zero–low-exposure group. In models with covariate adjustment, the associations remained significant except for height and height-for-age standard-deviation score. Further, 41.4% ($n = 12$) of participants in the zero–low-exposure group and 58.6% ($n = 17$) in the high-exposure group reported the occurrence of menarche ($P < 0.006$ with IPCW, both with and without covariates).

Among boys in the high-exposure group, weight, weight-for-age standard-deviation score, height, and height-for-age standard-deviation score were lower, and the difference of current measured height-for-age standard-deviation scores from the standard-deviation scores for midparental target height was higher than among boys in the zero–low-exposure group. The group difference in weight did not survive covariate adjustments, and the groups did not differ in terms of other markers of pubertal maturation (Table 2).

Cognition

Adolescents in the high-exposure group scored significantly lower on general, verbal, and performance IQ scales and on word generation, design fluency, narrative memory, and memory of faces compared with the zero–low-exposure group (Table 3). Except for design fluency and narrative memory, these differences were significant in models that adjusted for covariates, and the results did not change substantially when we excluded children in special education and/or with general estimated IQs of <70 age-standardized points owing to problems in visual processing ($n = 4$ in the zero–low-exposure group and $n = 1$ in the high-exposure group) (Table 3). There were no significant group differences in the indices of executive function (Table 3).

Psychiatric problems

The odds of scoring above the borderline clinical cutoff on the attention deficit/hyperactivity disorder problem scale were significantly higher among adolescents in the high-exposure group than among those in the zero–low-exposure group in models with and without covariates (Table 4), and the odds remained significant when we excluded adolescents with neurocognitive deficits ($n = 3$ in zero–low-exposure group and $n = 1$ in the high-exposure group) (Table 4). Web Table 1 (available at http://aje.oxfordjournals.org/) shows group differences in the Child Behavior Checklist broadband and narrowband scales.

HPA-axis activity

There were no significant group differences in HPA-axis activity (Web Table 2).

Dose-response associations between glycyrrhizin intake and outcomes

Web Tables 3–5 show dose-response associations between the levels of glycyrrhizin consumed in licorice by the pregnant women and their adolescent children's pubertal maturation, cognition scores, and psychiatric problems. These tests were conducted among those who reported consuming at least some licorice during pregnancy (range of glycyrrhizin content in consumed licorice was 10–2,464 mg/week; $n = 159$). Sample size varied slightly by outcome and is reported in the Web Tables. These findings largely replicate the exposure-level group differences.

DISCUSSION

Girls whose mothers consumed high amounts of licorice during pregnancy showed more advanced pubertal maturation at age 12 years than did girls whose mothers consumed none or low amounts. On average, these girls were more than 3 cm taller (0.4 SD taller for age and 0.5 SD closer to their expected adult target height). This probably reflects an earlier pubertal growth spurt and less remaining growth potential. They were also 8 kg heavier (0.6 SD heavier in weight for age) and had BMIs that were 2.2 higher (0.6 SD higher BMI for age), which could be a cause or a consequence of their more advanced maturational stage. There were no consistent associations between maternal licorice consumption during pregnancy and pubertal maturation in boys at this age.

These findings extend previous human studies showing that low birth weight and preterm birth, proxies of prenatal environmental adversity, predict earlier onset of pubertal maturation in girls and/or boys (33). Our findings also corroborate results derived from animal models of glucocorticoid overexposure showing altered timing of pubertal maturation in the exposed offspring (13).

Collection of blood samples for testing of hypothalamic-pituitary-gonadal axis or adrenal androgen parameters was not feasible in the current study. It remains unclear whether earlier hypothalamic-pituitary-gonadal axis maturation and increase in adrenal androgen production underlie the advanced pubertal stage of the female offspring exposed to high maternal consumption of licorice during pregnancy. We did

Table 1. Characteristics of a Sample of Adolescents Born in 1998, According to Level of Maternal Glycyrrhizin Exposure Due to Licorice Consumption During Pregnancy, Helsinki, Finland, 2009–2011

Characteristic	Zero–Low Exposure (≤249 mg/week) (n = 327)			High Exposure (≥500 mg/week) (n = 51)			P Value[b]
	No.[a]	%	Mean (SD)	No.[a]	%	Mean (SD)	
Parental Factors							
Maternal consumption of glycyrrhizin in licorice during pregnancy, mg/week	327		46.8 (75.4)	51		845.4 (405.1)	<0.001
Maternal age at delivery, years	327		30.3 (4.6)	51		29.4 (4.4)	0.17
Mode of delivery	327			51			1.00
Vaginal	295	90.2		46	90.2		
Cesarean (emergency or elective)	32	9.8		5	9.8		
Placenta weight, g	327		626 (121)	50		606 (143)	0.30
Maternal smoking during pregnancy, cigarettes/day	327			51			0.81
0	290	88.7		44	86.3		
1–10	24	7.3		4	7.8		
>10	13	4.0		3	5.9		
Maternal alcohol consumption during pregnancy	327			51			0.39
No	212	64.8		38	74.5		
Yes	65	19.9		7	13.7		
Did not report yes or no	50	15.3		6	11.8		
Among those reporting consumption, g/week	65		12.3 (9.7)	7		19.7 (12.4)	0.07
Maternal coffee consumption during pregnancy, cups/week	314		6.3 (6.7)	50		6.7 (7.2)	0.73
Maternal tea consumption during pregnancy, cups/week	305		5.4 (6.1)	47		5.4 (7.3)	0.83
Maternal cacao consumption during pregnancy, cups/week	283		1.1 (2.3)	47		1.2 (2.3)	0.86
Maternal salt consumption during pregnancy	327			51			0.17
No	288	88.1		41	80.4		
Yes	25	7.6		8	15.7		
Low-sodium	14	4.3		2	3.9		
Maternal chocolate consumption during pregnancy	327			51			0.11
Never	3	0.9		0	0		
Seldom	100	30.6		8	15.7		
Weekly	210	64.2		39	76.5		
Daily	14	4.3		4	7.8		
Maternal weight at delivery, kg	326		62.6 (9.5)	51		63.4 (11.3)	0.61
Maternal height, cm	327		166.9 (5.4)	51		166.4 (6.0)	0.60
Maternal body mass index at delivery[c]	326		22.5 (3.3)	51		22.9 (3.9)	0.46
Maternal gestational diabetes or hypertensive pregnancy disorder	327			51			0.38
No	316	96.6		48	94.1		
Yes	11	3.4		3	5.9		
Maternal systolic blood pressure at delivery, mm Hg	308		121.8 (14.0)	47		120.4 (13.4)	0.51
Maternal diastolic blood pressure at delivery, mm Hg	307		74.5 (10.7)	46		74.6 (9.5)	0.98
Maternal stress during pregnancy (100-mm visual analog scale), mm	326		34.8 (24.1)	51		32.5 (27.3)	0.54
Maternal age at menarche, years	298			50			
Mothers of girls	150		12.8 (1.3)	29		12.4 (1.0)	0.10
Mothers of boys	148		12.8 (1.3)	21		12.4 (1.5)	0.22
Paternal height, cm	323		180.1 (6.3)	51		179.1 (6.6)	0.31

Table continues

Table 1. Continued

Characteristic	Zero–Low Exposure (≤249 mg/week) (n = 327)			High Exposure (≥500 mg/week) (n = 51)			P Value[b]
	No.[a]	%	Mean (SD)	No.[a]	%	Mean (SD)	
Highest educational level of either parent at child's age of 12.5 years	327			51			0.27
Secondary or less	33	10.1		9	17.6		
Vocational	77	23.5		12	23.5		
University degree	217	66.4		30	58.8		
Neonatal and Adolescent Factors							
Sex	327			51			0.31
Male	166	50.8		22	43.1		
Female	161	49.2		29	56.9		
Birth order	327			51			0.46
First	187	57.2		32	62.7		
Second or later	140	42.8		19	37.3		
Length of gestation, weeks	327		40.1 (1.2)	51		39.9 (1.4)	0.19
Birth weight, g	327		3,551 (460)	51		3,478 (410)	0.29
Birth ponderal index, kg/m^3	327		27.9 (2.2)	51		27.5 (2.2)	0.73
Birth length, cm	327		50.3 (1.9)	51		50.0 (1.8)	0.37
Birth head circumference, cm	327		35.6 (1.4)	51		35.5 (1.4)	0.49
Apgar score at 1 minute (1–10)	326		8.7 (1.1)	51		8.8 (0.9)	0.64
Licorice consumption in adolescence	327			51			0.89
Never	58	17.7		7	13.7		
Less than once per week	180	55.0		28	54.9		
Once per week	35	10.7		8	15.7		
2–4 days/week	1	0.3		0	0		
Daily	1	0.3		0	0		
Did not report	52	15.9		8	15.7		

Abbreviation: SD, standard deviation.

[a] The number of participants with available data varied. For continuous variables, we present the number of participants for whom data were available. For categorical variables, we show the total number of participants, under which we present the number of participants in each subgroup.

[b] P values (2-sided) correspond to tests for differences between the zero–low-exposure and high-exposure groups (t test for continuous variables and χ^2 test for categorical and binary outcomes).

[c] Body mass index was calculated as weight (kg)/height (m)^2.

not find any differences in basal diurnal HPA-axis patterns. Although this might be likely to contrast with our previous report in this cohort at an earlier age (18), our prior observations reflected mainly peak HPA-axis responses to psychosocial stress, which we did not test this time. Also, at puberty, the substantial individual differences in maturation often mask more subtle HPA-axis differences.

When compared with the zero–low-exposure group, both adolescent girls and boys in the high-exposure group had IQ test scores more than 7 age-standardized points (more than 0.5 SD in an IQ test with mean = 100 (SD, 15)) lower for estimated general, verbal, and performance IQ, and they fared worse on tests measuring verbal productivity and memory. They also had more than 3-fold greater odds of attention deficit/hyperactivity disorder problems. These findings are in agreement with findings from glucocorticoid overexposure animal models (7–13) and our earlier findings in children from this cohort at a prepubertal age (17). Thus, these detrimental cognitive and neuropsychiatric outcomes following high maternal licorice consumption persist into early adolescence.

The detrimental associations of high maternal licorice consumption with offspring cognition and behavior were widespread, including the domains of general intelligence, language, memory, visuospatial processing, and behavioral and inattention problems, suggesting fairly widespread brain dysfunction. The limbic system may be a key target—particularly the hippocampus, which is involved in regulation of various cognitive processes, emotion, and behavior and is sensitive to early-life glucocorticoid manipulations (7–13). However, executive function was unaffected, suggesting that "higher-level" cognitive functioning resists prenatal glucocorticoid exposure.

Table 2. Growth and Pubertal Development of Adolescent Girls and Boys Born in 1998, According to Level of Maternal Consumption of Glycyrrhizin in Licorice During Pregnancy, Helsinki, Finland, 2009–2011

	With IPCW and Without Covariates										
	Zero–Low Exposure (0–249 mg/week)			High Exposure (≥500 mg/week)			Mean Difference or Odds Ratio	95% CI	Cohen's *d*[a]	Δ[b]	*P* Value[c]	*P* Value With IPCW and With Covariates[d]
	No.	%	EMM (SE)	No.	%	EMM (SE)						
Girls[e]												
Anthropometry	161			29								
Weight, kg			46.9 (0.8)			55.0 (2.4)	8.1	3.1, 13.0	0.80		0.001	<0.001
Weight for age (SD)			0.1 (0.1)			0.7 (0.2)	0.6	0.2, 1.9	0.64		0.002	0.001
BMI[f]			19.3 (0.2)			21.5 (0.8)	2.2	0.6, 3.8	0.69		0.006	0.001
BMI for age (SD)			0.1 (0.1)			0.7 (0.2)	0.6	0.2, 0.9	0.64		0.002	0.002
Height, cm			155.4 (0.6)			159.1 (1.3)	3.7	0.9, 6.9	0.50		0.009	0.09
Height for age (SD)			−0.1 (0.8)			0.3 (0.2)	0.4	0.1, 0.8	0.41		0.03	0.09
Midparental target height (SD)			0.7 (0.1)			0.6 (0.1)	−0.1	−0.3, 0.2	0.16		0.47	0.09
Midparental target height (SD) minus height for age (SD)			0.7 (0.1)			0.2 (0.1)	−0.5	−0.8, −0.2	−0.54		0.001	0.001
Pubic hair development, Tanner stage	142			29			4.2	1.7, 9.9			0.001	0.001
I	22	15.5		1	3.2					−12.1		
II	51	35.9		7	24.1					−11.8		
III	55	38.7		10	34.5					−4.2		
IV	14	9.9		11	37.9					28.0		
Breast development, Tanner stage	142			29			2.1	1.1, 4.1			0.04	0.041
I	10	7.0		0	0					−7.0		
II	53	37.3		8	27.6					−9.7		
III	45	31.7		11	37.9					6.2		
IV	34	23.9		10	34.5					10.6		
Pubertal Development Scale score	144			29			5.5	2.4, 12.8			<0.001	0.001
No development	69	47.9		5	17.2					−30.7		
Development barely begun	60	41.7		13	44.8					3.1		
Development definitely under way	15	10.4		11	37.9					27.5		
Boys[g]												
Anthropometry	152			19								
Weight, kg			48.7 (0.9)			43.5 (2.1)	−5.2	13.0, −3.1	−0.47		0.02	0.052
Weight for age (SD)			0.2 (0.1)			−0.3 (0.2)	−0.5	−0.9, −0.1	−0.47		0.03	0.045
BMI			19.8 (0.3)			18.9 (0.8)	−0.9	−2.5, 0.6	−0.28		0.24	0.18
BMI for age (SD)			0.2 (0.1)			−0.1 (0.2)	−0.3	−0.8, 0.2	−0.33		0.19	0.13
Height, cm			156.1 (0.7)			151.5 (1.1)	−4.6	−7.1, −2.1	−0.57		<0.001	0.031
Height for age (SD)			0.1 (0.1)			−0.5 (0.2)	−0.5	−0.9, −0.2	−0.53		0.002	0.034
Midparental target height (SD)			0.6 (0.1)			0.5 (0.1)	−0.1	−0.4, 0.2	−0.17		0.48	0.95
Midparental target height (SD) minus height for age (SD)			0.5 (0.1)			0.9 (0.1)	0.3	0.1, 0.7	0.41		0.01	0.02
Pubic hair development, Tanner stage	148			19			0.7	0.3, 1.5			0.34	0.21
I	36	24.3		5	26.3					2.0		

Table continues

Table 2. Continued

	With IPCW and Without Covariates											
	Zero–Low Exposure (0–249 mg/week)			High Exposure (≥500 mg/week)			Mean Difference or Odds Ratio	95% CI	Cohen's *d*[a]	Δ[b]	*P* Value[c]	*P* Value With IPCW and With Covariates[d]
	No.	%	EMM (SE)	No.	%	EMM (SE)						
II	68	45.9		11	57.9					12.0		
III	29	19.6		2	10.5					−9.1		
IV	15	10.1		1	5.3					−4.8		
Development of genitalia, Tanner stage	148			18			1.0	0.5, 2.2			0.98	0.50
I	10	6.8		2	11.1					4.3		
II	54	36.5		4	22.2					−14.3		
III	60	40.5		11	61.1					20.6		
IV	24	16.2		1	5.6					−10.6		
Pubertal Development Scale score	150			19			0.4	0.9, 2.0			0.43	0.35
No development	120	80		17	89.5					9.5		
Development barely begun	30	20		2	10.5					−9.5		
Development definitely under way	0	0		0	0					0.0		

Abbreviations: BMI, body mass index; CI, confidence interval; EMM, estimated marginal mean age; IPCW, inverse-probability-of-censoring weight; SD, standard deviation; SE, standard error.

[a] Effect size calculated as the difference between means divided by the pooled SD weighted by sample size.

[b] Δ for the arithmetic difference in percentage between the zero–low- and high-exposure groups.

[c] *P* values (2-sided) correspond to tests for differences between zero–low- and high-exposure groups (generalized linear models weighted by IPCW specifying a Gaussian distribution for quantitative (EMM, SE, mean difference) and ordinal logistic reference (no., %, odds ratio) distribution for ordinal outcomes).

[d] *P* values (2-sided) correspond to tests for differences between zero–low- and high-exposure groups (generalized linear models weighted by IPCW specifying a Gaussian distribution for quantitative (EMM, SE, mean difference) and ordinal logistic reference (no., %, odds ratio) for ordinal outcomes), adjusted for age at 12.5 years (all outcomes except for weight, height, and BMI for age), maternal age at menarche, age and BMI at delivery, smoking status, alcohol consumption, coffee consumption, tea consumption, cacao consumption, salt consumption, chocolate consumption, and stress during pregnancy and for the adolescent's birth weight by sex, gestational length, highest educational level of either parent, and adolescent's licorice consumption at age 12.5 years.

[e] Anthropometric, Tanner staging, and Pubertal Development Scale data were missing for 0, 19, and 17 girls, respectively, in the zero–low-exposure group ($n = 161$) and none of the girls in the high-exposure group ($n = 29$).

[f] BMI was calculated as weight (kg)/height $(m)^2$.

[g] Anthropometric, Tanner staging, and Pubertal Development Scale data were missing for 14, 18, and 16 boys, respectively, in the zero–low-exposure group ($n = 166$) and for 3 boys in the high-exposure group ($n = 22$; 1 boy reported development of pubic hair only on the Tanner staging questionnaire).

Limitations of our study include, first, that we measured average weekly consumption levels of glycyrrhizin during pregnancy and consumption of licorice confectionery only. Hence, we could not determine the level of glycyrrhizin consumed per dose of licorice, how much the consumed doses varied during gestation, or whether glycyrrhizin exposure had a "critical window" for adverse associations. One double-blind, randomized, controlled study in 39 healthy female volunteers showed that the "no observed adverse effect level" for the effect of glycyrrhizin on plasma volume expansion (suppression of plasma renin-angiotensin-aldosterone system and increase in body weight) and electrolyte concentrations (decline in plasma potassium and rise in bicarbonate concentrations) is 2 mg/kg body weight (34). In our sample, the average weekly glycyrrhizin content (per kg body weight at delivery) in licorice products consumed was 2.3 mg/kg body weight (range, 0–4.5 mg/week) in the zero–low-exposure group and 13.7 mg/kg body weight (range, 6.4–41.4 mg/week) in the high-exposure group. While we cannot draw conclusions about placental 11β-HSD2 function from these findings, and we are not aware of studies that have tested whether there is a similar "no observed adverse effect level" of glycyrrhizin on placental 11β-HSD2, our ex vivo dual-perfusion-method study of fresh, intact, term placentas showed that even very low doses of glycyrrhizin potently inhibit the placental glucocorticoid barrier function (9). Because the associations between maternal glycyrrhizin intake and pubertal maturation in girls and cognition and attention deficit/hyperactivity disorder problems in both girls and boys were linear, it appears that no safe exposure during human pregnancy exists.

Table 3. Cognitive Ability of Adolescents Born in 1998, According to Level of Maternal Consumption of Glycyrrhizin in Licorice During Pregnancy, Helsinki, Finland, 2009–2011

Cognitive Ability Test Score	With IPCW and Without Covariates									*P* Value With IPCW and With Covariates[d]
	Zero–Low Exposure (0–249 mg/week)		High Exposure(≥500 mg/week)		Mean Difference	95% CI	Cohen's *d*[a]	*P* Value[b]	*P* Value[c]	
	No.	EMM (SE)	No.	EMM (SE)						
Wechsler Intelligence Scale for Children III[e,f]	302		48							
General IQ		106.0 (0.8)		98.9 (1.9)	−7.1	−11.2, −3.1	−0.50	0.001	<0.001	0.003
Verbal IQ		111.7 (1.0)		103.6 (2.4)	−8.2	−13.2, −3.1	−0.48	0.002	0.001	0.002
Vocabulary		11.9 (0.2)		9.7 (0.4)	−1.4	−2.4, −0.5	−0.49	0.002	0.002	0.01
Similarities		12.3 (0.2)		11.2 (0.4)	−1.2	−2.0, −0.3	−0.40	0.007	0.014	0.003
Performance IQ		101.1 (1.1)		94.1 (2.5)	−7.1	−12.4, −1.8	−0.38	0.008	0.009	0.042
Block design		10.0 (0.2)		9.1 (0.5)	−1.0	−2.0, −0.1	−0.30	0.06	0.07	0.09
Picture Arrangement		10.2 (0.2)		9.2 (0.4)	−0.9	−1.8, −0.2	−0.31	0.03	0.03	0.18
Developmental Neuropsychological Assessment II[e,f]	305		49							
Word generation		9.5 (0.2)		8.1 (0.3)	−1.4	−2.1, −0.7	−0.54	<0.001	<0.001	0.002
Design fluency		11.8 (0.2)		10.9 (0.4)	−0.9	−1.7, −0.1	−0.33	0.03	0.047	0.15
Narrative memory		10.7 (0.2)		9.7 (0.4)	−1.0	−1.8, −0.2	−0.39	0.02	0.019	0.072
Memory of faces		10.2 (0.2)		9.1 (0.5)	−10	−2.0, −0.1	−0.38	0.04	0.051	0.04
Memory of names		9.6 (0.2)		9.3 (0.4)	−0.2	−1.1, 0.5	−0.13	0.41	0.56	0.17
Theory of mind verbal task		15.8 (0.6)		15.5 (0.2)	−0.3	−0.6, 0.1	−0.31	0.06	0.09	0.007
Theory of mind contextual task		7.0 (0.1)		7.0 (0.2)	−0.0	−0.3, 0.3	0.02	0.91	0.96	0.93
Continuous Performance Test II[e]	278		45							
No. of omission errors		9.0 (1.0)		9.7 (1.9)	0.6	−3.5, 4.8	0.04	0.80	0.67	0.49
No. of commission errors		23.9 (0.4)		23.3 (1.1)	0.6	−2.8, 1.6	−0.09	0.28	0.40	0.84
Reaction time, ms		349.0 (2.9)		359.6 (7.7)	10.6	−26.6, 10.5	0.21	0.20	0.19	0.35
D prime score[g]		0.3 (0.0)		0.4 (0.0)			0.14	0.36	0.30	0.63
Wisconsin Card Sorting Task[e]	295		48							
No. of perseverative errors		9.2 (0.3)		9.1 (0.7)	0.1	−1.6, 1.4	−0.02	0.88	0.71	0.81
No. of trials to complete first category		18.7 (0.6)		18.7 (1.4)	−1.3	−6.1, 3.6	0.09	0.61	0.74	0.56
Trail Making Test[e]	305		47							
Trail Making ratio[g]		44.2 (0.2)		44.8 (1.8)	0.6	−3.2, 4.4	0.04	0.78	0.94	0.71

Abbreviations: CI, confidence interval; EMM, estimated marginal mean; IPCW, inverse-probability-of-censoring weight; IQ, intelligence quotient; SE, standard error.

[a] Effect size calculated as the difference between means divided by the pooled standard deviation weighted by sample size.

[b] *P* values (2-sided) correspond to tests for differences between zero–low- and high-exposure groups (generalized linear models weighted by IPCW and specifying a Gaussian distribution for quantitative outcomes).

[c] *P* values (2-sided) correspond to tests for differences between zero–low- and high-exposure groups (generalized linear models weighted by IPCW and specifying a Gaussian distribution for quantitative outcomes) when adolescents in special education and/or with general estimated IQ <70 age-standardized points owing to problems in visual processing are excluded ($n = 4$ and $n = 1$ in zero–low- and high-exposure groups, respectively).

[d] *P* values (2-sided) correspond to tests for differences between zero–low- and high-exposure groups (generalized linear models weighted by IPCW and specifying a Gaussian distribution for quantitative outcomes) adjusted for sex, age at 12.5 years, maternal age and body mass index at delivery, smoking status, alcohol consumption, coffee consumption, tea consumption, cacao consumption, salt consumption, chocolate consumption, and stress during pregnancy and for the adolescent's birth weight by sex, gestational length, highest educational level of either parent, and the adolescent's licorice consumption at age 12.5 years.

[e] Wechsler Intelligence Scale for Children III, Developmental Neuropsychological Assessment II, Continuous Performance Test II, Wisconsin Card Sorting Task, and Trail Making Test data were missing for 25, 22, 49, 32, and 22 children in the zero–low-exposure group ($n = 327$), and 3, 2, 6, 3, and 4 children in the high-exposure group ($n = 51$), respectively.

[f] Scale scores of IQ are age-standardized to a mean of 100 (standard deviation, 15), and Developmental Neuropsychological Assessment scores are age-standardized scores.

[g] D prime score is calculated as the difference between target letter and nontarget letter distributions, and Trail Making ratio is calculated as (time to conduct Trail Making part B/time to conduct Trail Making part A) × 100.

Table 4. Presence of Psychiatric Problems Above the Borderline Clinical Cutoff Level Among Adolescents Born in 1998[a], According to Level of Maternal Consumption of Glycyrrhizin in Licorice During Pregnancy, Helsinki, Finland, 2009–2011

DSM-IV-Oriented Problem Scale	With IPCW and Without Covariates								*P* Value With IPCW and With Covariates[d]
	Zero–Low Exposure (0–249 mg/week) (*n* = 286)		High Exposure (≥500 mg/week) (*n* = 42)		Odds Ratio	95% CI	*P* Value[b]	*P* Value[c]	
	No.	%	No.	%					
Affective problems	66	23.1	9	21.4	0.8	0.4, 1.8	0.61	0.81	0.25
Anxiety problems	31	10.8	8	19.0	1.9	0.7, 4.4	0.16	0.11	0.13
Somatic problems	88	30.8	20	47.6	1.9	1.0, 3.8	0.052	0.07	0.004
ADHD problems	23	8.0	10	23.8	3.3	1.4, 7.7	0.005	0.005	<0.001
Oppositional defiant problems	44	15.4	6	14.3	0.7	0.3, 2.2	0.72	0.99	0.96
Conduct problems	27	9.4	7	16.7	1.9	0.8, 4.8	0.16	0.08	0.15

Abbreviations: ADHD, attention deficit/hyperactivity disorder; CI, confidence interval; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; IPCW, inverse-probability-of-censoring weight.

[a] We calculated scores for the DSM-IV–oriented scales and used the 82nd percentile as the cutoff to identify adolescents with borderline clinically significant psychiatric problems. In the zero–low-exposure group, 41 children had missing data on the problem scales. In the high-exposure group, 9 had missing data on problem scales.

[b] *P* values (2-sided) correspond to tests for differences between zero–low- and high-exposure groups (generalized linear models weighted by IPCW and specifying binary logistic reference distribution for binary outcomes).

[c] *P* values (2-sided) correspond to tests for differences between zero–low- and high-exposure groups (generalized linear models weighted by IPCW and specifying binary logistic reference distribution for dichotomous outcomes) when adolescents in special education and/or with general estimated IQ <70 age-standardized points owing to problems in visual processing are excluded (*n* = 3 and 1 in zero–low- and high-exposure groups, respectively) .

[d] *P* values (2-sided) correspond to tests for differences between zero–low- and high-exposure groups (generalized linear models weighted by IPCW and specifying binary logistic reference distribution for binary outcomes) adjusted for sex, age at 12.5 years, maternal age and body mass index at delivery, smoking status, alcohol consumption, coffee consumption, tea consumption, cacao consumption, salt consumption, chocolate consumption, and stress during pregnancy and for the adolescent's birth weight by sex, gestational length, highest educational level of either parent, and adolescent's licorice consumption at age 12.5 years.

These linear associations should, however, be interpreted with caution and confirmed in future studies specifically designed to test dose-response associations. Our observational study was designed primarily to test differences in pubertal, cognitive, and psychiatric outcomes between the zero–low-exposure and high-exposure groups.

Second, glycyrrhizin is also added to other food products, including candy, chewing gum, cookies, ice creams, syrups, herbal tea, alcoholic and nonalcoholic beverages, and traditional and herbal medicine. We were unable to determine consumption of other glycyrrhizin-containing products. However, according to the US Food and Drug Administration (35), the maximum allowable level of glycyrrhizin in foods is 16% for hard candy and 3.1% for soft candy, whereas the maximum allowable levels in other foods vary from 0.05% to 0.15%. This suggests that the main source of glycyrrhizin is licorice.

Third, we measured pubertal maturation in cross-section and thus could not determine pubertal timing. In addition, stage of puberty was self-reported. While self-ratings have been challenged for being inaccurate (36), there is also evidence suggesting good concordance of the Tanner staging questionnaire and the Pubertal Development Scale with physicians' ratings (37).

Fourth, our cohort comprised singleton, healthy babies, thus resulting in a lower prevalence of common pregnancy disorders in our sample than in the general population. This may limit generalizability from the findings to populations with higher postnatal morbidity. We cannot determine the extent to which our findings generalize to countries where licorice is not as commonly consumed. Yet, even in the United States, where licorice consumption is not common, the average daily consumption of glycyrrhizin can be up to 215.2 mg (38), suggesting widespread exposure beyond ostensible licorice products.

Fifth, we were able to adjust for a range of potential confounders, including maternal education as a proxy for maternal intelligence; this did not substantially affect our main findings. Neither did we find differences between the groups in any maternal, paternal, neonatal, or adolescent background characteristics. However, we cannot exclude the possibility that some unmeasured dietary and/or behavioral factor associated with maternal licorice consumption may in part explain our findings. Finally, we cannot address the possibility of selection bias resulting from sample attrition beyond the method of IPCW (19).

Our findings, which varied in effect size from small to large, are comparable to the effects of self-reported maternal binge drinking during pregnancy on poorer cognition (Cohen's $d = -0.13$) and greater behavioral problems (Cohen's $d = -0.15$) in the offspring (39). Prevention of alcohol use during pregnancy is a public health priority (40).

We are limited in drawing causal conclusions; however, as with ethanol, there are clear mechanistic grounds for suspecting that glycyrrhizin—a potent 11β-HSD2 inhibitor (the affinity the inhibitor has for the enzyme is high)—may be harmful in pregnancy (8, 9). Nutritional recommendations of various expert organizations do not mention glycyrrhizin use during pregnancy (35, 41, 42). The present findings suggest that pregnant women should be informed that consumption of licorice and other food products containing glycyrrhizin may be associated with harm for their developing offspring.

ACKNOWLEDGMENTS

Author affiliations: Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland (Katri Räikkönen, Silja Martikainen, Anu-Katriina Pesonen, Jari Lahti, Kati Heinonen, Riikka Pyhälä, Marius Lahti, Soile Tuovinen, Sara Sammallahti, Liisa Kuula, Alfredo Ortega-Alonso); Helsinki Collegium of Advanced Studies, University of Helsinki, Helsinki, Finland (Jari Lahti); Children's Hospital, Helsinki University Hospital and University of Helsinki, Helsinki, Finland (Karoliina Wehkalampi, Sture Andersson); Department of General Practice and Primary Health Care, Helsinki University Hospital and University of Helsinki, Helsinki, Finland (Johan G. Eriksson); Folkhälsan Research Center, Helsinki, Finland (Johan G. Eriksson); Endocrinology Unit, British Heart Foundation/University of Edinburgh Centre for Cardiovascular Science, Queen's Medical Research Institute, College of Medicine and Veterinary Medicine, University of Edinburgh, Edinburgh, United Kingdom (Rebecca M. Reynolds, Jonathan R. Seckl); Department of Medicine, Helsinki University Hospital and University of Helsinki, Helsinki, Finland (Timo E. Strandberg); Center for Life-Course Health Research, University of Oulu, Oulu, Finland (Timo E. Strandberg); Chronic Disease Prevention Unit, National Institute for Health and Welfare, Helsinki, Finland (Eero Kajantie); and Department of Obstetrics and Gynecology, MRC Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland (Eero Kajantie).

This work was supported by the Academy of Finland, the Signe and Ane Gyllenberg Foundation, the National Doctoral Programme of Psychology, the Finnish Ministry of Education and Culture, the Emil Aaltonen Foundation, the Foundation for Pediatric Research, the Foundation for Cardiovascular Research, the Juho Vainio Foundation, the Novo Nordisk Foundation, the Sigrid Jusélius Foundation, the Yrjö Jahnsson Foundation, and the University of Helsinki Research Funds.

The study sponsors had no role in study design; in the collection, analysis, or interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

Conflict of interest: none declared.

REFERENCES

1. World Cancer Research Fund; American Institute for Cancer Research. *Food Nutrition, Physical Activity, and the Prevention of Cancer: A Global Perspective*. Washington, DC: American Institute for Cancer Research; 2007.
2. Prentice P, Viner RM. Pubertal timing and adult obesity and cardiometabolic risk in women and men: a systematic review and meta-analysis. *Int J Obes (Lond)*. 2013;37(8):1036–1043.
3. Joinson C, Heron J, Araya R, et al. Association between pubertal development and depressive symptoms in girls from a UK cohort. *Psychol Med*. 2012;42(12):2579–2589.
4. Barker DJ. Developmental origins of adult health and disease. *J Epidemiol Community Health*. 2004;58(2): 114–115.
5. Räikkönen K, Forsén T, Henriksson M, et al. Growth trajectories and intellectual abilities in young adulthood: the Helsinki Birth Cohort study. *Am J Epidemiol*. 2009;170(4): 447–455.
6. Gitau R, Cameron A, Fisk NM, et al. Fetal exposure to maternal cortisol. *Lancet*. 1998;352(9129):707–708.
7. Seckl JR, Meaney MJ. Glucocorticoid programming. *Ann NY Acad Sci*. 2004;1032:63–84.
8. Edwards CR, Benediktsson R, Lindsay RS, et al. Dysfunction of placental glucocorticoid barrier: link between fetal environment and adult hypertension? *Lancet*. 1993; 341(8841):355–357.
9. Benediktsson R, Calder AA, Edwards CR, et al. Placental 11 beta-hydroxysteroid dehydrogenase: a key regulator of fetal glucocorticoid exposure. *Clin Endocrinol (Oxf)*. 1997;46(2): 161–166.
10. Holmes MC, Abrahamsen CT, French KL, et al. The mother or the fetus? 11β-hydroxysteroid dehydrogenase type 2 null mice provide evidence for direct fetal programming of behavior by endogenous glucocorticoids. *J Neurosci*. 2006; 26(14):3840–3844.
11. Welberg LA, Seckl JR, Holmes MC. Inhibition of 11β-hydroxysteroid dehydrogenase, the foeto-placental barrier to maternal glucocorticoids, permanently programs amygdala GR mRNA expression and anxiety-like behavior in the offspring. *Eur J Neurosci*. 2000;12(3):1047–1054.
12. Wyrwoll CS, Holmes MC. Prenatal excess glucocorticoid exposure and adult affective disorders: a role for serotonergic and catecholamine pathways. *Neuroendocrinology*. 2012; 95(1):47–55.
13. Zambrano E, Guzmán C, Rodríquez-González GL, et al. Fetal programming of sexual development and reproductive function. *Mol Cell Endocrinol*. 2014;382(1):538–549.
14. Brown RW, Diaz R, Robson AC, et al. The ontogeny of 11 beta-hydroxysteroid dehydrogenase type 2 and mineralocorticoid receptor gene expression reveal intricate control of glucocorticoid action in development. *Endocrinology*. 1996;137(2):794–797.
15. Teelucksingh S, Benediktsson R, Lindsay RS, et al. Liquorice. *Lancet*. 1991;337(8756):1549.
16. Strandberg TE, Järvenpää AL, Vanhanen H, et al. Birth outcome in relation to licorice consumption during pregnancy. *Am J Epidemiol*. 2001;153(11):1085–1088.
17. Räikkönen K, Pesonen AK, Heinonen K, et al. Maternal licorice consumption and detrimental cognitive and psychiatric outcomes in children. *Am J Epidemiol*. 2009: 170(9):1137–1146.
18. Räikkönen K, Seckl JR, Heinonen K, et al. Maternal prenatal licorice consumption alters hypothalamic-pituitary-adrenocortical

axis function in children. *Psychoneuroendocrinology*. 2010; 35(10):1587–1593.
19. Robins JM, Rotnitzky A, Zhao LP. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J Am Stat Assoc*. 1995;90(429):106–121.
20. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med*. 2015;34(28):3661–3679.
21. Saari A, Sankilampi U, Hannila ML, et al. New Finnish growth references for children and adolescents aged 0 to 20 years: length/height-for age, weight-for-length/height, and body mass index-for-age. *Ann Med*. 2011;43(3):235–248.
22. Sorva R, Tolppanen EM, Lankinen S, et al. Growth evaluation: parent and child specific height standards. *Arch Dis Child*. 1989;64(10):1483–1487.
23. Morris NM, Udry JR. Validation of a self-administered instrument to assess stage of adolescent development. *J Youth Adolesc*. 1980;9(3):271–280.
24. Petersen AC, Crockett L, Richards M, et al. A self-report measure of pubertal status: reliability, validity, and initial norms. *J Youth Adolesc*. 1980;17(2):117–133.
25. Wechsler D. *WISC-III: Wechsler Intelligence Scale for Children: Manual*. 3rd ed. San Antonio, TX: The Psychological Corporation; 1991.
26. Kaufman AS, Kaufman JC, Ramaswamy B, et al. Comparison of three WISC-III short forms: weighing psychometric, clinical, and practical factors. *J Clin Child Psychol*. 1996;25(1):97–105.
27. Korkman M, Kirk U, Kemp S. *NEPSY-II: Clinical and Interpretative Manual*. 2nd ed. San Antonio, TX: Harcourt Assessment; 2007.
28. Conners CK. *Conners' Continuous Performance Test (CPT II)*. Toronto, ON, Canada: Multi Health Systems; 2004.
29. Lezak MD. *Neuropsychological Assessment*. 4th ed. New York, NY: Oxford University Press; 2004.
30. Heaton RK. *WCST-64: Computer Version for Windows—Research edition*. Odessa, FL: Psychological Assessment Resources; 2000.
31. Achenbach TM, Rescorla LA. *Manual for the ASEBA School-Age Forms and Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families; 2001.
32. Martikainen S, Pesonen A-K, Lahti J, et al. Physical activity and hypothalamic-pituitary-adrenal axis function in adolescents. *Psychoneuroendocrinology*. 2014;49:96–105.
33. Wehkalampi K, Hovi P, Dunkel L, et al. Advanced pubertal growth spurt in subjects born preterm: the Helsinki Study of Very Low Birth Weight Adults. *J Clin Endocrinol Metab*. 2011;96(2):525–533.
34. Bijlsma J, Van Vloten P, Van Gelderen C, et al. *An Assessment of the Effects of Various Dosages of Glycyrrhizin in Healthy Female Volunteers* [in Dutch]. Bilthoven, The Netherlands: RIVM; 1996 (RIVM report no. 348801004).
35. US Food and Drug Administration. *Listing of specific substances affirmed as generally recognized as safe. United States Food and Drug Administration: licorice and licorice derivatives*. Code of Federal Regulations, 21CFR184.1408. (April 1, 2016).
36. Rasmussen A, Wohlfahrt-Veje C, Tefre de Renzy-Martin K, et al. Validity of self-assessment of pubertal maturation. *Pediatrics*. 2015;135(1):86–93.
37. Coleman L, Coleman J. The measurement of puberty: a review. *J Adolesc*. 2002;25(5):535–550.
38. Isbrucker RA, Burdock GA. Risk and safety assessment on the consumption of licorice root (*Glycyrrhiza* sp.), its extract, powder as a food ingredient with emphasis on the pharmacology and toxicology of glycyrrhizin. *Regul Toxicol Pharmacol*. 2006; 46(3):167–192.
39. Flak AL, Su S, Bertrand J, et al. The association of mild, moderate, and binge prenatal alcohol exposure and child neuropsychological outcomes: a meta-analysis. *Alcohol Clin Exp Res*. 2014;38(1):214–226.
40. World Health Organization. *Global Status Report on Alcohol and Health 2014*. Geneva, Switzerland: World Health Organization; 2014.
41. European Commission. *Opinion of the Scientific Committee on Food on Glycyrrhizinic Acid and Its Ammonium Salt*. (European Commission Scientific Committee on Food report SCF/CS/ADD/EDUL/255, Final, April 10, 2003). Brussels, Belgium: European Commission; 2003.
42. World Health Organization. *Evaluation of Certain Food Additives: Sixty-Third Report of the Joint FAO/WHO Expert Committee on Food Additives*. Geneva, Switzerland: World Health Organization; 2005. (WHO Technical Report Series no. 928).