Behav Genet (2014) 44:456–467
DOI 10.1007/s10519-014-9668-4

ORIGINAL RESEARCH

Maternal Smoking During Pregnancy and Adverse Outcomes in Offspring: Genetic and Environmental Sources of Covariance

Ralf Kuja-Halkola · Brian M. D'Onofrio · Henrik Larsson · Paul Lichtenstein

Received: 16 December 2013 / Accepted: 18 July 2014 / Published online: 13 August 2014

Abstract Maternal smoking during pregnancy (SDP) has been associated with several psychiatric outcomes in the offspring; studies have questioned whether the associations are causal, however. We analyzed all children born in Sweden between 1983 and 2009 to investigate the effect of SDP on multiple indicators of adverse outcomes in three areas: pregnancy outcomes (birth weight, preterm birth and being born small for gestational age), long-term cognitive abilities (low academic achievement and general cognitive ability) and externalizing behaviors (criminal conviction, violent criminal conviction and drug misuse). SDP was associated with all outcomes. Within-family analyses of the pregnancy outcomes were consistent with a causal interpretation as the associations persisted when siblings discordant for SDP were compared. For the cognitive and externalizing outcomes, the results were not consistent with causal effects; when comparing differentially exposed siblings none of the associations remained significant. In quantitative genetic models genetic factors explained the majority of the associations between SDP and cognitive and externalizing outcomes. The results suggest that the associations between SDP in mothers and cognition and externalizing behaviors in their offspring is primarily due to genetic effects that influence the behaviors in both generations.

Keywords Smoking during pregnancy · Children of siblings · Sibling comparison · Cousin comparison · Extended family model

Background

Maternal smoking during pregnancy (SDP) has been associated with adverse pregnancy outcomes and long-term cognitive and behavioral difficulties in the offspring (e.g., ADHD and low cognitive functioning) in humans, as well as in animals (Cnattingius 2004; Huizink and Mulder 2006; Knopik 2009). Researchers have suggested plausible biological mechanisms, such as fetal restriction of nutrients and oxygen (Huizink and Mulder 2006), alterations in neural development through nicotine binding to the nicotinic acetylcholine receptors in the fetal brain (Huizink and Mulder 2006; Knopik 2009), dysregulation of hypothalamic–pituary–adrenal axis (Huizink and Mulder 2006), and epigenetic effects (Knopik et al. 2012). Although SDP seems to be causally related to pregnancy outcomes, the causality of its effects on long-term difficulties have been questioned (Knopik 2009).

Many carefully designed observational studies on SDP and long-term outcomes have been carried out, e.g. by Paradis et al. (2011) investigating SDP and criminality at approximately 33 years of age while controlling for a number of measured covariates. Most observational studies that have compared unrelated individuals have found that SDP independently predicts offspring traits when controlling for parental characteristics that covary with SDP (reviews in Glantz and Chambers 2006; Wakschlag et al.

Edited by Valerie Knopik.

Electronic supplementary material The online version of this article (doi:10.1007/s10519-014-9668-4) contains supplementary material, which is available to authorized users.

R. Kuja-Halkola (✉) · H. Larsson · P. Lichtenstein
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, P.O. Box 281, 171 77 Stockholm, Sweden
e-mail: ralf.kuja-halkola@ki.se

B. M. D'Onofrio
Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

2002). However, a continuing concern is missing adjustments of important unmeasured confounders, such as maternal and paternal personality traits, intellectual abilities, and psychiatric problems that were either not included in the studies or were measured imprecisely, which has highlighted the need for other types of designs to resolve these issues (D'Onofrio et al. 2013). For example, several twin studies have shown that SDP is a genetically influenced trait (Agrawal et al. 2008; D'Onofrio et al. 2003; Ellingson et al. 2012) with heritability estimates ranging from 34 to 52 %; the genetic factors that influence SDP are shared with criminal convictions and drug use (Ellingson et al. 2012), as well as nicotine dependence (Agrawal et al. 2008). Genetic factors passed down from parents to their offspring could, therefore, account for the statistical associations between SDP and offspring traits. Thus, we believe that a genetically sensitive approach to the analysis of the association between SDP and any heritable outcome is of high importance.

Several quasi-experimental studies (e.g., comparisons of siblings differentially exposed to SDP and in vitro fertilization studies) have suggested that the long-term associations are due to familial confounding, rather than being causal (D'Onofrio et al. 2008, 2010a, b, 2012; Gilman et al. 2008; Kuja-Halkola et al. 2010; Lambe et al. 2006; Langley et al. 2012; Lundberg et al. 2010; Thapar et al. 2009), see review in D'Onofrio et al. (2013). The scientific community is, however, still far from a consensus regarding the question of causality (see e.g. Slotkin (2013)).

To date most of the quasi-experimental studies on SDP have not investigated the extent to which the familial factors that confound the associations are due to genetic and/or environmental effects. In vitro fertilization studies (Thapar et al. 2009) have suggested that genetic factors confound the associations, but sibling-comparison studies cannot identify the source of familial confounding (Donovan and Susser 2011; Lahey and D'Onofrio 2010). Furthermore, researchers need to examine the assumptions and limitations in the designs they use because each design has limitations and assumptions. Ultimately, finding the sources responsible for the underlying associations between smoking during pregnancy and offspring outcomes is essential for prevention/intervention efforts, as well as informing subsequent basic research (D'Onofrio et al. 2013).

In the current study we used family-based quasi-experimental methods, such as sibling-comparisons and Children of Siblings/Twins designs (D'Onofrio et al. 2003, 2013; Heath et al. 1985; Silberg et al. 2003), on total population data of 2.75 million Swedes, to test causal inferences and disentangle genetic and environmental effects of the association between SDP and outcomes in offspring. We studied three areas of possible adverse effects in offspring that have been related to SDP: pregnancy outcomes, long-term cognitive outcomes, and long-term externalizing outcomes. We also tested several fundamental assumptions of sibling-comparison designs to examine whether these would unduly influence our conclusions.

Methods

Subjects

We linked several nationwide Swedish registries maintained by government agencies using the unique personal identification number given to all Swedish citizens. These registries cover in principle the entire population (Ludvigsson et al. 2009). The use of these databases has been approved by the ethics committee at Karolinska Institutet, Stockholm, Sweden.

We studied all individuals born in Sweden from January 1st 1983 to December 31st 2009, consisting of 2,754,626 children, because valid data on smoking during pregnancy is available with good coverage from 1983. However, because we studied different outcomes, the investigated associations were made on different cohorts; see Table 1 for number of individuals available for analyses for each outcome. All outcomes were followed until 2009.

Relationships

Using the Multi-Generation Register and the Swedish Twin Register, we constructed extended families of different sizes and relations. We randomly selected up to two sisters who were mothers in each extended family (except we chose all female twin pairs). Then, we randomly drew up to two of each mothers' offspring, constructing up to two nuclear families within each extended family. We included nuclear families of three different types: single offspring, full-siblings and maternal half-siblings; and extended families of six different types: mothers without siblings, monozygotic twin mothers, dizygotic twin mothers, mothers who were full-siblings, and mothers who were maternal or paternal half-siblings.

Measures

Smoking during pregnancy

At the first antenatal visit for pregnant women, generally in the first trimester (before week 15), the nurse asked whether the mothers were smoking at that time. This is registered in the Swedish Medical Birth Register (MBR) (Centre for Epidemiology 2012) and was coded as 0 (No)

Table 1 Number of individuals for each outcome analyzed

	Sub-cohort				Families
	From registers	Birth years cohort[a]	Available	After elimination[b] (% of available)	Total[c] (% of available)
Birth weight, Preterm birth, and Born small for gestational age	Medical Birth Register	1983–2009	2,754,626	2,658,974 (96.5)	1,823,697 (66.2)
Low academic achievement	National School Register	1983–1995	1,409,909	1,124,858 (79.8)	869,553 (61.7)
General cognitive ability[d]	Military Conscription Register	1983–1992	429,335	299,450 (69.7)	257,268 (59.9)
Criminality	National Crime Register	1983–1989	712,484	670,953 (94.2)	546,208 (76.7)
Violent criminality	National Crime Register	1983–1989	712,484	669,973 (94.0)	545,537 (76.6)
Drug misuse	National Crime Register, Patient Register	1983–1987	486,353	481,044 (98.9)	398,705 (82.0)

[a] Years when individuals can be included in sub cohort

[b] Excluding individuals with missing values for gender, birth date and/or maternal age at childbirth, who had no possibility of getting outcome (e.g., was not at conscription), and individuals who died/emigrated

[c] Excluding individuals as in previous footnote ([b]) plus exclusion criteria; non-identifiable parents, twins in offspring generation, only inclusion of up to two siblings per nuclear family, only two mothers per extended family

[d] Only males in sub cohort

or 1 (Yes). A study found the self-reported SDP to be valid, with only 6 % of reported non-smokers having cotinine levels indicating that they were actually smoking (Lindqvist et al. 2002).

Outcomes

From the MBR we used birth weight (measured by midwifes at the hospital after delivery), gestational age (in days and calculated using either ultrasound or time since last menstrual period), and preterm birth (defined as being born gestational day 259 or earlier, but not before day 155 (week 22) where the birth was considered to be a late miscarriage, rather than a preterm birth). We defined being born small for gestational age as having birth weight in the lowest 10 % compared within gender and gestational day.

From the School Registry we collected grades in upper school, at approximate age 15. We created a binary variable to capture low academic achievement, using similar data as Lambe et al. (2006), although with a different operationalization; the lowest 10 % were coded as a 1 and the rest as 0. Individual with missing grades, indicating non-completion of the compulsory first nine school years, were included in the poor performance group. We collected a measure of general cognitive ability (Frisell et al. 2012b) from the Swedish Conscription Registry. The measure was recorded on a 9-grade scale, with hypothesized mean of 5 and standard deviation of 2. Military conscription was enforced by law until 2007 and was generally performed by men at 18 years of age. Only males that were between 17 and 20 years at conscription were included.

From the Crime Register we collected convictions of crimes in Swedish lower court. As a measure of criminality we used any conviction registered. For violent criminality we used convictions defined as in Frisell et al. (2011). In Sweden the age of criminal responsibility is fifteen, we limited the sub-cohort to cover individuals with at least 5 years at risk, and recorded convictions within this time period, i.e. between ages 15 and 20. Thus the sub-cohort covers offspring born between 1st January 1983 and 31st December 1989. In line with Kendler et al. (2012) we used a combination of diagnoses of alcohol/drug dependence from the Patient Register and drug-related convictions (including convictions for driving while intoxicated) from the Crime Register to get a measure of drug/alcohol misuse. The Patient Register contains diagnoses from all inpatient care instances that require hospitalization overnight, as well as admissions in outpatient settings since 2001. For these analyses we selected a sub-cohort with individuals at risk until age 22 and recorded drug/alcohol misuse as a one if there was a diagnosis or conviction before 22nd birthday, and as a zero otherwise, thus this sub-cohort consisted of individuals born between 1983 and 1987. For these externalizing behavioral outcomes the probability of observing an offense/diagnosis for any subject is dependent on time of follow up. The maximum ages was chosen to balance the probability of outcome with the number of eligible individuals. As a consequence, rather than analyzing convictions/diagnoses at any time in an individual's life, we are analyzing specific age-limited periods. For example, it is possible that early onset criminality has a higher genetic liability. Thus, this selection can

have implication for generalization, and generalizations outside early onset criminality/drug abuse should be done with caution.

General cognitive ability and birth weight were treated as continuous variables in the analyses, whereas the other outcomes were treated as binary variables.

Covariates

We included three covariates from the MBR which we adjusted for in all analyses where applicable; gender of offspring, maternal age at childbirth, and birth year.

Statistical analyses

Ordinary regression: establishing associations

We estimated the crude and covariate adjusted association between SDP and each outcome using linear (continuous outcomes) or logistic (binary outcomes) regressions. We used the statistical software R (R Development Core Team 2012) base package for analyses.

Within family analyses: investigating familial confounding

To investigate potential confounding by factors shared within families we performed within-family analyses. We studied siblings (where comparisons were made within full-siblings and within maternal half-siblings) and cousins (separately for full-cousins and half-cousins). For practical reasons, we selected a random pair of cousins per extended family in the analyses with continuous outcomes.

To estimate the within-family effects (Neuhaus and Kalbfleisch 1998; Neuhaus and McCulloch 2006), we performed conditional logistic regression for the binary outcomes, where each extended family was treated as a cluster in cousin comparisons, and nuclear families were treated as clusters in sibling comparisons. For the continuous outcomes, we included a pair-specific mean of the exposure as a covariate as well as a shared random intercept in a linear mixed model. Both methods produces a within-pair estimate that can be considered to be closer to the true, causal, parameter under certain assumptions (Frisell et al. 2012a). All siblings and cousins were included in these analyses, regardless if they were concordant or discordant in SDP, since information from other included covariates contributes to the likelihood. We used the package survival (Therneau 2012) and the package lme4 (Bates et al. 2012) in the statistical software R (R Development Core Team 2012) to fit the models. Here, and in the ordinary regression analyses, missing values in SDP and outcome were handled by case-wise deletion.

Table 2 Parameters in quantitative genetic model; interpretation in each generation

Parameter	Parent generation	Offspring generation
A (Additive genetic)	Additive genetics	Additive genetics
C (Common environment)	Environment unique to one mother	Environment shared between siblings
M (Maternal environment)	Environment shared between sisters who are mothers	Environment shared between cousins
P (Paternal environment)	Spouse effect	Paternal effect
E (Non-shared environment)	Environment unique to each pregnancy	Unique individual environment

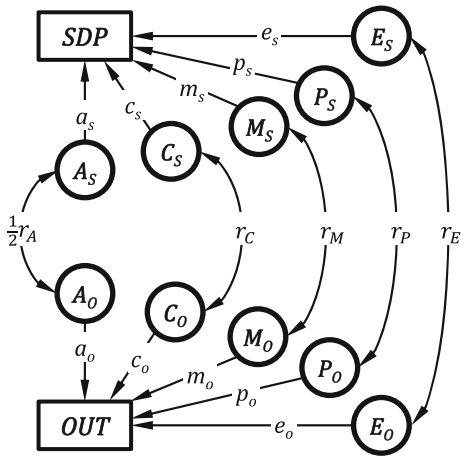

Fig. 1 A representation of the *ACMPE* model. (*Note* The figure represents the covariance between maternal smoking during pregnancy (SDP; indicated by sub-index "*S*") and an outcome (OUT; indicated by sub-index "*O*") within an individual.)

Structural equation models: estimating magnitude of genetic and environmental confounding

We then performed structural equation modeling (SEM) to investigate the source of potential familial confounding in a model we call the *ACMPE* model. We extended the standard models used in twin research, which decompose variance into genetic (*A*) and shared (*C*) and non-shared (*E*) environmental influences (Neale and Cardon 1992), to include five variance sources for each phenotype (Table 2; Fig. 1). The main difference between the *ACMPE* model and the standard twin model is (1) that we estimate an

intergenerational bivariate association (i.e., one phenotype in parental generation and one in offspring generation) and (2) that the shared environment is decomposed into environments shared by mothers and cousins (*M*), fathers and full-siblings (*P*), and all children of one mother (*C*). The model is more extensively explained in Appendix A in Supplementary Material.

To exemplify the variance parameters aimed at capturing different environmental sources of variance and covariance we here present a hypothetical example of SDP, childrearing regimes and academic achievement. Note that all associations are made up.

The C parameter

A woman behaves affectively towards her children, a behavior she learned from experiences not shared with her sister. Affective parenting is correlated with lower levels of SDP and makes all siblings in the nuclear family less liable to act out in school, which influences the offspring's higher academic achievements.

The M parameter

Two sisters were raised by parents who were very goal-oriented, pushing them to aim for good grades, a behavior correlated with higher levels of SDP. Both sisters act similarly as their parents, both by having higher levels of SDP, and pushing their own children. Thus, all cousins in the extended family tend to have high academic achievement.

The P parameter

A mother changes spouse between pregnancies, and the new spouse convinces her to quit smoking in her second pregnancy. The first child does not receive as much reinforcing feedback on her academic development by the father of the second child, making her perform worse academically.

The E parameter

A mother starts to smoke between first and second pregnancy, and continues while pregnant with her second child. The worse environment in utero restricts the neurological development of the second child, making her less mentally capable and therefore she performs academically worse (a scenario where SDP is causally affecting academic achievement).

We fitted this bivariate model separately for SDP and each outcome, and estimated all the covariance parameters. Thus, we partitioned the association between SDP and each offspring outcome into *A*, *C*, *M*, *P* and *E* factors and estimated how much of the correlation between SDP and outcome that is due to each factor.

To find a parsimonious model, where the number of parameters that fit the data is minimized without significant loss in explanatory power, we performed a series of likelihood ratio tests, where we excluded non-significant parameters in the outcomes (at α-level 0.05). We started with the parameter for the outcome with lowest value and the corresponding cross-phenotype parameter (e.g., paternal effect and its correlation with paternal effect in SDP), after elimination of pairs of parameters we continued with just the cross-phenotypic parameters in a similar fashion until all parameters in the model were significantly different from zero. The resulting model is called "best-fitting model". The within-phenotype *E*-parameters were not subject to significance testing since they contain random error. Because of potentially causal interpretation (D'Onofrio et al. 2013; Kendler et al. 1993, Turkheimer and Harden, Submitted) the cross-phenotypic *E*-parameter was included in the final model regardless of statistical significance. For the binary outcomes we used the liability-threshold model, where an underlying normal distribution is assumed for the liability of having the outcome. If an individual has a liability higher than an estimated threshold the variable is observed as 1, otherwise as 0. For each variable we allowed different types of families to have different means/prevalences by letting the mean in the assumed underlying normal distribution be different both for nuclear families and extended families (because the mean/prevalence of exposure and outcomes is different in, for example, full- and half-sibling families). Rather than investigating potential mediating or moderating effects of birth year, gender, and maternal age, we adjusted the means/prevalences of each variable for the covariates, where we included linear and quadratic terms for maternal age and birth year.

We used the package OpenMx (Boker et al. 2011, 2012) in the software R (R Development Core Team 2012) to fit SEMs. All code is available on request from the corresponding author. Missing values in exposure and outcome were handled using full information maximum likelihood.

Sensitivity analyses

As in any statistical analysis, the within-sibling analyses (from which we aim at drawing the strongest causal inferences) rest on a number of assumptions. In Appendix B in Supplementary Material we investigate three of the assumptions which potentially may affect our analyses (generalizability of mothers who change their smoking pattern between pregnancies, carry-over effects, and

Table 3 Means and prevalences for outcomes in pregnancies where the mother was not smoking, was smoking, and where smoking status is missing

	Mean/prevalence (95 % confidence interval) per smoking status		
Outcome	SDP = 0	SDP = 1	SDP = missing
Birth weight (g)	3588 (3587–3589)	3388 (3386–3390)	3496 (3493–3500)
Preterm birth (%)	4.6 (4.6–4.7)	6.1 (6.0–6.2)	7.7 (7.6–7.9)
Born small for gestational age (%)	4.9 (4.8–4.9)	11.4 (11.3–11.5)	6.3 (6.1–6.4)
Low academic achievement (%)	6.9 (6.9–7.0)	17.5 (17.4–17.7)	11.2 (10.9–11.4)
General cognitive ability (9-point score)	5.29 (5.28–5.30)	4.65 (4.64–4.67)	5.07 (5.03–5.10)
Criminality (%)	8.0 (7.9–8.1)	14.7 (14.5–14.9)	11.2 (10.9–11.5)
Violent criminality (%)	1.3 (1.3–1.3)	3.8 (3.7–3.9)	2.5 (2.4–2.6)
Drug misuse (%)	4.2 (4.1–4.3)	8.6 (8.4–8.7)	6.4 (6.1–6.7)

Note Values are from the analytic samples (column "After elimination" in Table 1)

sibling contagion effects), see D'Onofrio et al. (2013) for an in-depth description of issues in sibling comparisons.

Results

Descriptive

In Table 3 we present the means and prevalences of the outcomes for pregnancies where the mother did and did not smoke while pregnant, as well as when smoking status was missing. For all of the outcomes the values in the group where mothers smoked while pregnant compared were worse than when the mother did not smoke. The group missing SDP status has values in between the observed smokers and non-smokers, except for the preterm birth outcome, where the prevalence of preterm births is greater in the group missing SDP than in either of the non-missing groups. For further investigations of different patterns of smoking between pregnancies in the same mother see Appendix B and Appendix Table 6a–h in Supplementary Material.

Ordinary regression

Results from the ordinary cohort analyses for each outcome can be found in Table 4. Results from previous studies were confirmed; SDP was associated with each of the outcomes. For example, offspring to mothers who smoked during pregnancy weighed 201 g less than offspring to mothers not smoking, and the association remained after controlling for the potential confounders (181 g less).

Within analyses: investigating familial confounding

We proceeded to perform within-family analyses, where we compared half-cousins, full-cousins, maternal half-siblings, and full-siblings respectively (Table 4). Our results were in line with previous research; the effect of SDP on pregnancy outcomes persisted in the within family analyses, although the estimates were somewhat attenuated. For example, even in the most controlled analyses (within full-siblings), a child where the mother smoked while pregnant was on average 92 g lighter than his/her full-sibling born in a pregnancy where the mother did not smoke. In contrast, SDP seemed to have no direct effect on cognitive and externalizing outcomes when siblings discordant for smoking during pregnancy were compared. These results imply familial confounding for all of the long-term associations.

It should be noted that the within-sibling analyses are adjusting for unmeasured confounders assumed stable between pregnancies. To be a confounder a variable has to be related to the exposure as well as the outcome; such potential confounders may thus be viewed as stable *in the mother* (having an effect on the exposure) between pregnancies. An example of this is maternal genetic influences.

Structural equation models: estimating magnitude of genetic and environmental confounding

We estimated how much of the variation in the outcomes that was due to each variance source in eight separate univariate models. To maximize power, in a separate model, we used the full 1983–2009 cohort to estimates fractions of explained variance for SDP. Table 5 presents the variance components from univariate models, and Appendix Table 1 in Supplementary Material presents the modelled values from which the variance components were derived. As can be seen additive genetic effects (i.e., the heritability) explained 69 % of the variance for SDP, and between 27 and 86 % of the variance in the outcomes.

We then used SEM to fit separate bivariate models for each outcome, parameter estimates and standard errors from the full bivariate models are presented in Appendix Table 2 in Supplementary Material. We then identified the best-fitting model for each bivariate relationship (the resulting best-fitting models are presented in Appendix Table 3 in Supplementary Material; decision steps are presented in Appendix Table 4a–h in Supplementary

Table 4 Effects of maternal smoking during pregnancy on outcome, both unrelated (cohort) analyses and within relative analyses; estimate (95 % confidence interval)

	Effect measure	Cohort		Within relatives			
		Crude	Adjusted	Half-cousins	Full-cousins	Half-siblings	Full-siblings
Birth weight (g)	Regression coefficients	−201 (−203 to −198)	−181 (−184 to −179)	−205 (−218 to −191)	−185 (−192 to −177)	−135 (−146 to −123)	−92 (−97 to −86)
Preterm birth	Odds ratio	1.34 (1.31–1.36)	1.29 (1.27–1.32)	1.23 (1.14–1.34)	1.19 (1.14–1.26)	1.18 (1.04–1.34)	1.21 (1.12–1.30)
Born small for gestational age	Odds ratio	2.52 (2.48–2.55)	2.45 (2.41–2.48)	2.41 (2.23–2.59)	2.29 (2.19–2.39)	1.83 (1.63–2.06)	1.64 (1.54–1.75)
Low academic achievement	Odds ratio	2.84 (2.80–2.89)	2.67 (2.63–2.71)	1.91 (1.74–2.09)	1.62 (1.54–1.70)	1.07 (0.92–1.25)	0.93 (0.87–1.01)
General cognitive ability[a] (9-point score)	Regression coefficients	−0.63 (−0.65 to −0.62)	−0.57 (−0.59 to −0.55)	−0.24 (−0.43 to−0.04)	−0.27 (−0.35 to −0.18)	−0.03 (−0.41 to 0.36)	−0.00 (−0.10–0.09)
Criminality	Odds ratio	1.98 (1.95–2.02)	1.90 (1.86–1.94)	1.62 (1.40–1.87)	1.32 (1.23–1.42)	0.92 (0.67–1.28)	0.95 (0.86–1.05)
Violent criminality	Odds ratio	3.02 (2.90–3.14)	2.77 (2.66–2.89)	2.07 (1.57–2.73)	1.59 (1.37–1.86)	0.82 (0.45–1.51)	0.94 (0.75–1.17)
Drug misuse	Odds ratio	2.13 (2.07–2.19)	2.00 (1.94–2.05)	1.55 (1.25–1.93)	1.47 (1.32–1.64)	0.83 (0.47–1.46)	0.92 (0.78–1.07)

Note All models except Crude model, adjusted for gender, birth year in categories (for birth weight in 3 year intervals; for preterm birth and born small for gestational age in 4 year intervals; for low academic achievement in 2 year intervals; for general cognitive ability, criminality, violent criminality and drug misuse in 1 year intervals; all starting in 1983), and maternal age at childbirth in categories (for birth weight <17, 17–19, then 2 year intervals, 42–44 and >44; for preterm birth <20, then 4 year intervals, >44; for born small for gestational age <20, then three year intervals, 44–45, >45; for low academic achievement <19, then 4 year intervals, >42; for general cognitive ability <20, then 2 year intervals, >39; for criminality, violent criminality and drug misuse <20, then 5 year intervals, >39)

[a] Sub-cohort includes only males

Table 5 Variance explained by each variance source (95 % confidence intervals)

Maternal phenotype	Additive genetics	Environment unique to one mother	Environment shared between sisters who are mothers	Spouse effect	Environment unique to each pregnancy
Maternal smoking during pregnancy[a] (%)	69 (67–70)	3 (2–4)	4 (3–5)	16 (15–16)	8 (8–9)
Offspring phenotype	**Additive genetics**	**Environment shared between siblings**	**Environment shared between cousins**	**Paternal effect**	**Unique individual environment**
Birth weight (%)	50 (47–53)	16 (15–17)	7 (7–8)	0 (0–0)	27 (26–28)
Preterm birth (%)	27 (19–35)	24 (21–27)	8 (6–9)	0 (0–0)	41 (37–46)
Born small for gestational age (%)	43 (34–52)	16 (12–19)	7 (6–9)	0 (0 –0)	34 (29–38)
Low academic achievement (%)	86 (78–94)	1 (−1–3)	12 (11–14)	0 (−1–1)	1 (−4–5)
General cognitive ability (%)	33 (26–39)	19 (10–29)	0 (0–0)	1 (−8–10)	46 (43–50)
Criminal convictions (%)	39 (19–60)	2 (−5–9)	4 (1–7)	0 (0–0)	54 (43–66)
Violent criminal convictions (%)	63 (42–84)	2 (−5–8)	6 (2–11)	0 (0–0)	29 (16–42)
Drug/alcohol misuse (%)	35 (30–40)	6 (4–8)	3 (1–4)	1 (1–2)	56 (51–60)

Note Estimates are from univariate models. Wald-type confidence intervals, hence negative values may exist in intervals. Parameters were fitted using the non-squared, and, if applicable, non-standardized variance parameters, therefore standard errors are calculated using the delta method

[a] Calculated using the full 1983–2009 cohort

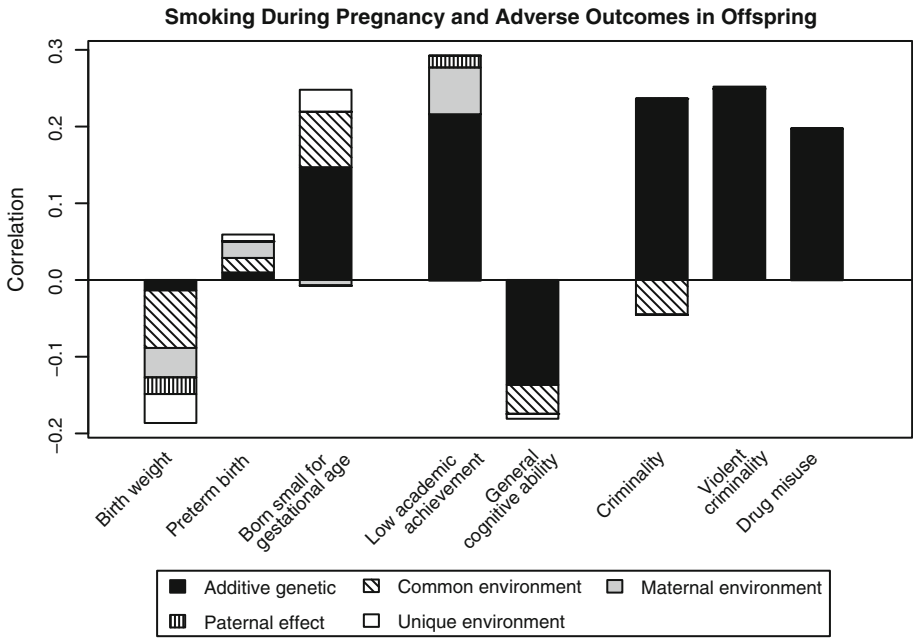

Fig. 2 Correlations between maternal smoking during pregnancy and outcomes in offspring and parts of the correlations explained by different sources of variance (*Note* Results from the best-fitting models, for full models see Appendix Figure 1 in Supplementary Material. Due to computational issues the model for the association between SDP and preterm birth, as well as with being born small for gestational age, was fitted in three steps; first SDP and outcomes were fitted separately, then the results from model fitting were used in the cross-phenotype analyses.)

Material). The magnitude of the correlations, as well as the part of the correlations explained by each of the variance sources in the best-fitting model, are displayed in Fig. 2. For comparative reasons we produced a figure, similar to Fig. 2, which shows the correlation between SDP and the outcomes in the full model, before any model fitting was performed (Appendix Fig. 1 in Supplementary Material). The pattern of overlap differed between the pregnancy

outcomes and the cognitive and externalizing outcomes. In all the associations between SDP and pregnancy outcomes, non-shared environmental effects, which are consistent with a causal inference, were important, explaining 12–20 % of the total correlation. In contrast, non-shared environment factors only accounted for 0-4 % of the observed associations with the cognitive and externalizing outcomes. Genetic factors explained a majority of the associations between SDP and the long-term outcomes, accounting for at least 74 % of the correlations.

To exemplify the quantification of the correlations we here present the calculations for the SDP-low academic achievement association, following the approach outlined in Appendix A in Supplementary Material. The total covariance, which in this case is equivalent to correlation since both variables have a variance of 1, is (values from Appendix Table 3 in Supplementary Material)

$$\begin{aligned} Cov\left(SDP_{ijk}, OUT_{ijk}\right) &= \frac{1}{2}a_s a_o r_A + c_s c_o r_C + m_s m_o r_M \\ &\quad + p_s p_o r_P + e_s e_o r_E \\ &= \frac{1}{2}0.810 \cdot 0.893 \cdot 0.598 \\ &\quad + 0.344 \cdot 0.154 \cdot 0 \\ &\quad + 0.267 \cdot 0.366 \cdot 0.623 \\ &\quad + 0.296 \cdot 0.105 \cdot 0.494 \\ &\quad + 0.258 \cdot 0.185 \cdot (-0.018) \\ &= 0.216 + 0 + 0.061 + 0.015 \\ &\quad - 0.001 = 0.292. \end{aligned}$$

Thus the fractions explained by the different variance sources are

$$\begin{aligned} &\frac{(0.216 + 0 + 0.061 + 0.015 - 0.001)}{0.292} \\ &= 0.742 + 0 + 0.209 + 0.053 - 0.003, \end{aligned}$$

in the order *A*, *C*, *M*, *P* and *E*.

Sensitivity analyses

The within-sibling analyses utilize mothers who are discordant in SDP between pregnancies. If these mothers were very different from non SDP-discordant mothers, especially in the associations between SDP's the outcomes, our results may not generalize to other types of families. However, we found no support for SDP-discordant families being substantially different; we observed that the means/prevalences in outcome were roughly halfway between that of SDP-concordant non-smokers and smokers, indicating a liability in between the two concordant groups (Appendix Table 7 in Supplementary Material). Further, if SDP-discordant families are not generalizable to the general population we would not expect the agreement with within-cousin comparison seen in Table 4. If the SDP status in the first pregnancy affected the outcome in the second pregnancy, either directly through carry-over effects (e.g., smoking may induce a biological change in the mother, which carries over to following pregnancies) or through sibling contagion effect (e.g., the first offspring engages in criminal activities and influences the second offspring to do the same) the assumptions of sibling comparison would be violated and the within-sibling estimate would be biased. For the cognitive/behavioral outcomes we found no support for such effects being present when we conducted bi-directional analyses (Appendix Table 8 in Supplementary Material). As such, the results do not indicate that the assumptions in the sibling-comparison design, in as much as we could test them, account for the familial confounding of the associations between SDP and the long-term outcomes.

Discussion

Although SDP was associated with all outcomes in the domains we studied, in line with previous research (Cnattingius 2004; Huizink and Mulder 2006; Knopik 2009), we found support for different sources being responsible for the associations. Consistent with causal interpretations, the associations between SDP and pregnancy outcomes persisted when we compared siblings discordant for SDP (although the effect size of the association with preterm birth was relatively limited). For the long-term cognitive and externalizing outcomes, however, the analyses were not consistent with causal associations; when we compared siblings discordant for SDP, none of the long-term associations remained large, or statistically significant.

Similar to our study, sibling-comparison studies of externalizing behavior (D'Onofrio et al. 2008), school performance (D'Onofrio et al. 2010b; Lambe et al. 2006), substance use (D'Onofrio et al. 2012), stress coping (Kuja-Halkola et al. 2010), criminality (D'Onofrio et al. 2010a), intellectual performance (Lundberg et al. 2010), and ADHD (Skoglund et al. 2013) have all suggested substantial familial confounding. Here, we took the analyses one step further and estimated the source of familial confounding. Consistent with the previously observed intra-generational correlation between SDP and maternal criminal behavior and other co-occurring risk factors (Ellingson et al. 2012), we found that genetic factors explained the main part of the associations with cognitive and externalizing outcomes in the offspring. These results are in line with the findings from family studies and in vitro fertilization studies of SDP and ADHD (Langley et al. 2012, Thapar et al. 2009). In a recent study on the association between SDP and conduct disorder using an adoption

design, the results suggested that the association was not due to familial confounding during the postnatal period (Gaysina et al. 2013). The finding, which is consistent with our results, suggests the familial confounding is due to factors present during the prenatal period (e.g., genes), rather than being exerted in the postnatal period.

One of the major strengths of the current study is the use of a populations based sample, where data has been collected prospectively. Furthermore we utilized the knowledge of familial relationships to estimate effects, which may be interpreted as being relatively free of confounding of factors shared within families. None of the sensitivity analyses indicated that the within-sibling results were due to the assumptions inherent in the sibling-comparison design that we were able to test. The extensive family information also allowed us to disentangle the relative contribution of genetic and environmental effects for the association between SDP and outcomes in offspring; sibling-comparison studies by themselves are unable to do so (D'Onofrio et al. 2013). Furthermore we predicted multiple outcomes in each of the domains to avoid misrepresenting inferences, and we found converging results.

The measure of SDP was a yes/no at approximately 15 weeks of gestation. Thus we did not investigate any dose–response relationship, neither in terms of how much the pregnant women smoked, nor for how long.

One of the reasons for writing this paper was to better understand the previously identified familial confounding for some of the SDP associations. To do this, we wanted to use the best available data. However, there were no methods developed for these types of family based analyses, and we therefore developed the method used in this paper. Admittedly, the parameterization used might not be the ultimately best one, especially not for the shared environmental parameters. We chose to separate the environment into variance parts *C*, *M*, *P*, and *E*. This is a somewhat arbitrary choice that might, or might not, be valid. In the best case scenario, we have captured the most important features of how environmental influences' on phenotypes are shared between relatives, and therefore estimates of cross-phenotypic additive genetic effects are unbiased. In the worst case scenario, our estimates of cross-phenotypic additive genetic effects are biased, but the within-family estimates testing causal inferences would remain unchanged. Further, we assume that maternal siblings have shared environmental effects (the *C* in offspring generation and *M* in parental generation), while paternal siblings have not. We do this since when parents divorce offspring tend to more often live with the mother (Statistics Sweden 1994). However, we did not explicitly validate this assumption in the present data. Because the main source of confounding was genetic, we did not try to further evaluate the different shared environmental parameters, because that would most likely not contribute to different interpretations of the data. If these methods are used for other research questions, further method developments might be needed.

Model fitting of the quantitative genetic *ACMPE* model presented some problems because the method is novel and the software has not been used for similar types of data sets. This lead to several drawbacks, first we were not able to fit the full bivariate models for the SDP-preterm birth and SDP-born small for gestational age associations since the models failed to converge. Instead we chose to use parameter estimates from separately fitted univariate models for SDP and outcomes in the bivariate analyses, where we estimated the cross-phenotype parameters. This may introduce bias, and spuriously increase precision in the parameters in the bivariate model. Second, we encountered problems in finding the global likelihood maximum. Focusing on finding the best fit for the models, we ran each model from a variety of starting values, and re-ran them from the previously fitted values, to ensure that a global likelihood maximum had been reached. The difficult optimization procedure, which solely focused on improvements in the likelihood, made the standard errors for some parameters not reliable (since the curvature of the likelihood around the fitted values, which is captured by the hessian matrix and used to calculate standard errors, obtained from a fitted model with starting values close to the final fitted values did not behave well for our models and data). To solve this problem we did not rely on standard errors in model fitting and inference from these models, rather, we used likelihood ratio tests (Appendix Tables 4a–h in Supplementary Material). We have neither considered dominant genetic effects nor assortative mating in our analyses. The measure for SDP is self-reported and may thus be subject to misclassification, which in turns leads to bias (toward null) of the associational estimates, which is particularly problematic for within-relative analyses (Frisell et al. 2012a; McGue et al. 2010). However, our exposure has been shown to be valid (Lindqvist et al. 2002), and we were able to estimate robust associations with pregnancy outcomes within families, suggesting measurement error alone cannot account for the findings.

Prevention of SDP remains important; in our analyses we add further support of SDP being a causal risk factor for birth/pregnancy-related complications. However, we find no such support for adolescent/adult outcomes in the cognitive and behavioral problems. Nevertheless, the observed associations are real; mothers who smoke while pregnant have offspring with greater risk of many adverse outcomes throughout life. Our results suggest that the sources of the long-term associations originate in families (primarily due to shared genetic variation), however. Although this should not be interpreted as a deterministic feature, which is immune against interventional efforts, there are

nevertheless important consequences. Understanding the underlying associations between SDP and offspring outcomes is necessary for appropriate prevention, intervention and future research efforts. For example, an imaging study was recently conducted where measurable differences between offspring of exposed and not exposed to SDP during reward anticipation were observed (Muller et al. 2013). The potential for genetic variants passed down from mother was noted as a limitation but not examined further, in contrast to imaging work on schizophrenia and working memory (Karlsgodt et al. 2007). Thus, Muller et al. cannot be certain that it is smoking that caused the observed differences or if the differences were caused by genetic variants passed down from the smoking mother. Thus, to avoid wasted resources, the information that genetic effects are of substantial importance for the association between SDP and long-term outcomes should be considered in intervention and prevention, as well as in basic (e.g., clinical neuroscience) research.

Acknowledgments Preliminary partial results were presented at the 42nd Annual Meeting of the Behavior Genetics Association; June 24, 2012; Edinburgh, Scotland. The study was supported by Grants from the National Institute of Child Health and Human Development (B.M.D., Grant Number HD061817); by the Swedish Research Council through the Swedish Initiative for Research on Microdata in the Social And Medical Sciences (SIMSAM) framework grant no 340-2013-5867, by the and Swedish Research Council for Health, Working Life and Welfare, and by the Swedish Prison and Probation Services.

Conflict of Interest Author Kuja-Halkola R, Author D'Onofrio BM, Author Larsson H and Author Lichtenstein P declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000. Data was merged and anonymized by Statistics Sweden, an independent governmental agency. The key linking the personal number to the data was destroyed immediately after merging; therefore no informed consent was required.

References

Agrawal A, Knopik VS, Pergadia ML, Waldron M, Bucholz KK, Martin NG, Heath AC, Madden PAF (2008) Correlates of cigarette smoking during pregnancy and its genetic and environmental overlap with nicotine dependence. Nicotine Tob Res 10:567–578

Bates D, Maechler M, Bolker B (2012) lme4: Linear mixed-effects models using S4 classes [computer program]

Boker S, Neale M, Maes H, Wilde M, Spiegel M, Brick T, Spies J, Estabrook R, Kenny S, Bates T, Mehta P, Fox J (2011) OpenMx: an open source extended structural equation modeling framework. Psychometrika 76:306–317

Boker SM, Neale MC, Maes HH, Wilde MJ, Spiegel M, Brick TR, Estabrook R, Bates TC, Mehta P, von Oertzen T, Gore RJ, Hunter MD, Hackett DC, Karch J, Brandmaier AM (2012) OpenMx User Guide, Release 1.2.0-1919

Centre for Epidemiology (2012) The Swedish Medical Birth Register—A Summary of Content and Quality Socialstyrelsen

Cnattingius S (2004) The epidemiology of smoking during pregnancy: smoking prevalence, maternal characteristics, and pregnancy outcomes. Nicotine Tob Res 6(Suppl 2):S125–S140

D'Onofrio BM, Turkheimer E, Eaves LJ, Corey LA, Berg K, Solaas MH, Emery RE (2003) The role of the Children of Twins design in elucidating causal relations between parent characteristics and child outcomes. J Child Psychol Psychiatry 44:1130–1144

D'Onofrio BM, Van Hulle CA, Waldman ID, Rodgers JL, Harden KP, Rathouz PJ, Lahey BB (2008) Smoking during pregnancy and offspring externalizing problems: an exploration of genetic and environmental confounds. Dev Psychopathol 20:139–164

D'Onofrio BM, Singh AL, Iliadou A, Lambe M, Hultman CM, Grann M, Neiderhiser JM, Langstrom N, Lichtenstein P (2010a) Familial confounding of the association between maternal smoking during pregnancy and offspring criminality: a population-based study in Sweden. Arch Gen Psychiatry 67:529–538

D'Onofrio BM, Singh AL, Iliadou A, Lambe M, Hultman CM, Neiderhiser JM, Langstrom N, Lichtenstein P (2010b) A quasi-experimental study of maternal smoking during pregnancy and offspring academic achievement. Child Dev 81:80–100

D'Onofrio BM, Rickert ME, Langstrom N, Donahue KL, Coyne CA, Larsson H, Ellingson JM, Van Hulle CA, Iliadou AN, Rathouz PJ, Lahey BB, Lichtenstein P (2012) Familial confounding of the association between maternal smoking during pregnancy and offspring substance use and problems. Arch Gen Psychiatry 69:1140–1150

D'Onofrio BM, Lahey BB, Turkheimer E, Lichtenstein P (2013) Critical need for family-based, quasi-experimental designs in integrating genetic and social science research. Am J Public Health 103(Suppl 1):S46–S55

Donovan SJ, Susser E (2011) Commentary: advent of sibling designs. Int J Epidemiol 40:345–349

Ellingson JM, Rickert ME, Lichtenstein P, Langstrom N, D'Onofrio BM (2012) Disentangling the relationships between maternal smoking during pregnancy and co-occurring risk factors. Psychol Med 42:1547–1557

Frisell T, Lichtenstein P, Langstrom N (2011) Violent crime runs in families: a total population study of 12.5 million individuals. Psychol Med 41:97–105

Frisell T, Oberg S, Kuja-Halkola R, Sjolander A (2012a) Sibling comparison designs: bias from non-shared confounders and measurement error. Epidemiology 23:713–720

Frisell T, Pawitan Y, Langstrom N (2012b) Is the association between general cognitive ability and violent crime caused by family-level confounders? PLoS ONE 7:e41783

Gaysina D, Fergusson DM, Leve LD, Horwood J, Reiss D, Shaw DS, Elam KK, Natsuaki MN, Neiderhiser JM, Harold GT (2013) Maternal smoking during pregnancy and offspring conduct problems: evidence from 3 independent genetically sensitive research designs. JAMA Psychiatry 70:956–963

Gilman SE, Gardener H, Buka SL (2008) Maternal smoking during pregnancy and children's cognitive and physical development: a causal risk factor? Am J Epidemiol 168:522–531

Glantz MD, Chambers JC (2006) Prenatal drug exposure effects on subsequent vulnerability to drug abuse. Dev Psychopathol 18:893–922

Heath AC, Kendler KS, Eaves LJ, Markell D (1985) The resolution of cultural and biological inheritance: informativeness of different relationships. Behav Genet 15:439–465

Huizink AC, Mulder EJ (2006) Maternal smoking, drinking or cannabis use during pregnancy and neurobehavioral and cognitive functioning in human offspring. Neurosci Biobehav Rev 30:24–41

Karlsgodt KH, Glahn DC, van Erp TG, Therman S, Huttunen M, Manninen M, Kaprio J, Cohen MS, Lonnqvist J, Cannon TD (2007) The relationship between performance and fMRI signal during working memory in patients with schizophrenia, unaffected co-twins, and control subjects. Schizophrenia Res 89:191–197

Kendler KS, Neale MC, MacLean CJ, Heath AC, Eaves LJ, Kessler RC (1993) Smoking and major depression. A causal analysis. Arch Gen Psychiatry 50:36–43

Kendler KS, Sundquist K, Ohlsson H, Palmer K, Maes H, Winkleby MA, Sundquist J (2012) Genetic and familial environmental influences on the risk for drug abuse: a national Swedish adoption study. Arch Gen Psychiatry 69:690–697

Knopik VS (2009) Maternal smoking during pregnancy and child outcomes: real or spurious effect? Dev Neuropsychol 34:1–36

Knopik VS, Maccani MA, Francazio S, McGeary JE (2012) The epigenetics of maternal cigarette smoking during pregnancy and effects on child development. Dev Psychopathol 24:1377–1390

Kuja-Halkola R, D'Onofrio BM, Iliadou AN, Langstrom N, Lichtenstein P (2010) Prenatal smoking exposure and offspring stress coping in late adolescence: no causal link. Int J Epidemiol 39:1531–1540

Lahey BB, D'Onofrio BM (2010) All in the family: comparing siblings to test causal hypotheses regarding environmental influences on behavior. Curr Dir Psychol Sci 19:319–323

Lambe M, Hultman C, Torrang A, Maccabe J, Cnattingius S (2006) Maternal smoking during pregnancy and school performance at age 15. Epidemiology 17:524–530

Langley K, Heron J, Smith GD, Thapar A (2012) Maternal and paternal smoking during pregnancy and risk of ADHD symptoms in offspring: testing for intrauterine effects. Am J Epidemiol 176:261–268

Lindqvist R, Lendahls L, Tollbom O, Aberg H, Hakansson A (2002) Smoking during pregnancy: comparison of self-reports and cotinine levels in 496 women. Acta Obstet Gynecol Scand 81:240–244

Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekbom A (2009) The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. Eur J Epidemiol 24:659–667

Lundberg F, Cnattingius S, D'Onofrio B, Altman D, Lambe M, Hultman C, Iliadou A (2010) Maternal smoking during pregnancy and intellectual performance in young adult Swedish male offspring. Paediatr Perinat Epidemiol 24:79–87

McGue M, Osler M, Christensen K (2010) Causal inference and observational research: the utility of twins. Persp Psychol Sci 5:546–556

Muller KU, Mennigen E, Ripke S, Banaschewski T, Barker GJ, Buchel C, Conrod P, Fauth-Buhler M, Flor H, Garavan H, Heinz A, Lawrence C, Loth E, Mann K, Martinot JL, Pausova Z, Rietschel M, Strohle A, Struve M, Walaszek B, Schumann G, Paus T, Smolka MN (2013) Altered reward processing in adolescents with prenatal exposure to maternal cigarette smoking. JAMA Psychiatry 70:847–856

Neale MC, Cardon LR (1992) Methodology for genetic studies of twins and families. Kluwer Academic Publishers, Dordrecht

Neuhaus JM, Kalbfleisch JD (1998) Between- and within-cluster covariate effects in the analysis of clustered data. Biometrics 54:638–645

Neuhaus JM, McCulloch CE (2006) Separating between- and within-cluster covariate effects by using conditional and partitioning methods. J R Stat Soc Ser B 68:859–872

Paradis AD, Fitzmaurice GM, Koenen KC, Buka SL (2011) Maternal smoking during pregnancy and criminal offending among adult offspring. J Epidemiol Community Health 65:1145–1150

R Development Core Team (2012) R: A language and environment for statistical computing [computer program]

Silberg J, Rutter M, D'Onofrio B, Eaves L (2003) Genetic and environmental risk factors in adolescent substance use. J Child Psychol Psychiatry 44:664–676

Skoglund C, Chen Q, D'Onofrio BM, Lichtenstein P, Larsson H (2013) Familial confounding of the association between maternal smoking during pregnancy and ADHD in offspring. J Child Psychol Psychiatry 55:61–68

Slotkin TA (2013) Maternal smoking and conduct disorder in the offspring. JAMA Psychiatry 70:901–902

Statistics Sweden (1994) Fakta om den svenska familjen. Demografiska rapporter (1994:2). Stockholm

Thapar A, Rice F, Hay D, Boivin J, Langley K, van den Bree M, Rutter M, Harold G (2009) Prenatal smoking might not cause attention-deficit/hyperactivity disorder: evidence from a novel design. Biol Psychiatry 66:722–727

Therneau T (2012) A Package for Survival Analysis in S. R package version 2.36-14. [computer program]

Turkheimer E, Harden KP (2014) Behavior genetic research methods: testing quasi-causal hypotheses using multivariate twin data. In: Reis HT, Judd CM (eds) Handbook of research methods in personality and social psychology, 2nd edn. Cambridge University Press, Cambridge

Wakschlag LS, Pickett KE, Cook E Jr, Benowitz NL, Leventhal BL (2002) Maternal smoking during pregnancy and severe antisocial behavior in offspring: a review. Am J Public Health 92:966–974