Final version published: Ekblad MO, Rolan E, Marceau K, Palmer R, Todorov A, Heath AC, Knopik VS. Disruptive Behavior in Siblings Discordant for Exposure to Maternal Smoking during Pregnancy: a Multi-rater approach. Nicotine Tob Res. 2019 Nov 18. pii: ntz214. doi: 10.1093/ntr/ntz214.

Disruptive Behavior in Siblings Discordant for Exposure to Maternal Smoking during Pregnancy: a Multi-rater approach

Mikael O. Ekblad PhD[1,2], Emily Rolan MS[1], Kristine Marceau PhD[1], Rohan Palmer PhD[3], Alexandre Todorov PhD[4], Andrew C. Heath D.Phil.[4], Valerie S. Knopik PhD[1]

[1]Department of Human Development and Family Studies, Purdue University, West Lafayette, Indiana, USA

[2]Department of General Practice, Institute of Clinical Medicine, Turku University and Turku University Hospital, Turku, Finland

[3]Department of Psychology, Emory University

[4]Department of Psychiatry, Midwest Alcoholism Research Center, Washington University School of Medicine, St Louis

Corresponding author: Mikael Ekblad, MD, PhD, Visiting Scholar

Address: Department of Human Development and Family Studies, Purdue University, Hanley Hall, Room 354, 1202 West State Street, West Lafayette, IN 47906

Key words: Disruptive behavior, Genetically-informed designs, Prenatal exposure, Health consequences, Sibling comparison

Word count: 5108

Abbreviations: CD, Conduct Disorder; DBD, Disruptive Behavior Disorder; CBCL, Child Behavior Checklist; ODD, Oppositional Defiant Disorder; SDP, Smoking during pregnancy; TRF, Teacher report form.

ABSTRACT

Introduction

Maternal smoking during pregnancy (SDP) is associated with disruptive behavior. However, there is debate whether the SDP-disruptive behavior association is a potentially causal pathway or rather a spurious effect confounded by shared genetic and environmental factors.

Methods

The Missouri Mothers and Their Children Study (MO-MATCH) is a sibling comparison study that includes families (n=173) selected for sibling pairs (aged 7 to 16 years) discordant for SDP. Critically, the sibling comparison design is used to disentangle the effects of SDP from familial confounds on disruptive behavior. An SDP severity score was created for each child using a combination of SDP indicators (timing, duration, and amount of SDP). Multiple informants (parents and teachers) reported on disruptive behavior (i.e., DSM-IV semi-structured interview, the Child Behavior Checklist, and Teacher Report Form).

Results

The variability in disruptive behavior was primarily a function of within-family differences (66-100%). Consistent with prior genetically-informed approaches, the SDP-disruptive behavior association was primarily explained by familial confounds (genetic and environmental). However, when using a multi-rater approach (parents and teachers), results suggest a potentially causal effect of SDP on disruptive behavior (b=0.09, SE=0.04, P=0.03). The potentially causal effect of SDP remained significant in sensitivity analyses.

Discussion

These findings suggest that familial confounding likely plays a complex role in the SDP-disruptive behavior association when examining both parent and teacher reports of behavior. Importantly, the current study highlights the importance of multiple raters, reflecting a more comprehensive measure of complex behaviors (e.g., disruptive behavior) to examine the teratogenic effects of SDP.

Implications

Our study provides additional evidence that controlling for genetic and family factors is essential when examining the effect of SDP on later behavioral problems, as it explains a portion of the association between SDP and later behavioral problems. However, we found a significant association between SDP and disruptive behavior when using a multi-rater approach that capitalizes on both parent and teacher report, suggesting that parent and teacher ratings capture a unique perspective that is important to consider when examining SDP-behavior associations.

INTRODUCTION

Maternal smoking during pregnancy (SDP) is one of the most important preventable prenatal factors associated with adverse effects on fetal development and youth postnatal health. However, the prevalence of SDP remains fairly common in many countries, in spite of the increasing knowledge of the potential harm of SDP exposure for the fetus.[1,2] SDP can significantly modulate genetically programmed brain development during fetal growth leading to alterations in brain structure and function.[3-5] These SDP-related variations in brain development could translate to later youth maladjustment, such as disruptive behavior.[6-8]

Oppositional Defiant Disorder (ODD) and Conduct Disorder (CD) are similar disorders, which differ primarily based on the severity of the symptoms. ODD and CD are often combined and studied as Disruptive Behavior Disorder (DBD), because they are similar externalizing disorders marked by high levels of aggression.[9] The combination of ODD and CD into DBD is further supported by evidence that ODD and CD share a great amount of genetic influence.[10] Although ADHD is sometimes included in DBD constructs, ODD and CD are differentiated from ADHD in the predominance of behavioral vs. attentional problems.[9] Compared to CD, ODD typically occurs earlier in the child's life and is defined by hostile, defiant, and disobedient behaviors directed at parents or teachers. ODD may lead to, or develop into, later CD.[11] CD is characterized by repetitive and persistent behaviors that involve violating the basic rights of other human beings or animals and breaking more severe rules (e.g. shoplifting or breaking into homes or cars in order to steal). Thus, in this paper, 'disruptive behavior' refers to ODD, CD, and a combination of ODD and CD symptoms. Given the prevalence of SDP and recent findings that call to question this association between SDP and later behavioral problems as a causal relationship,[12,13] the current study sought to utilize a sibling comparison design to disentangle the effects of SDP from familial confounds on disruptive behavior.

In diverse samples and developmental periods, SDP has been found as a robust predictor of ODD,[14] CD[6,15-19] and DBD.[20-22] A recent meta-analysis by Ruisch et al.,[21] including 25 non-genetically-informed studies, showed an association between SDP and CD. Importantly, Gaysina et al.[17] found an association between SDP and CD for children of both genetically-related and genetically-unrelated mothers in a study evaluating data from three independent samples. They concluded it is unlikely that genetic and postnatal environmental factors (e.g., inconsistent parenting skills) fully explain the association between SDP and CD.[17] Robinson et al.[23] showed that quitting smoking in early pregnancy prevented behavior problems in offspring when compared to offspring exposed to continued SDP. However, the difference in behavior between groups may be due to maternal characteristics that hindered the ability to stop smoking,[24] which could be a precursor for poorer behavior outcomes as opposed to maternal duration of SDP.[25]

There has been increasing debate surrounding the association between SDP and later behavioral problems and whether the association is a potentially causal effect or a spurious effect confounded by shared genetic and environmental factors.[7,26] A growing body of research has shown that the association between SDP and CD is entirely explained by previously unmeasured environmental, familial or genetic factors.[12,13] D'Onofrio et al.[12] investigated the effect of SDP on four to ten-year-old children's externalizing problems, including conduct, oppositional and attention problems, using maternal reports on the Behavioral Problem Index. They used hierarchical linear models to compare siblings discordant for SDP and found that the siblings did not differ with respect to CD and ODD symptoms.[12] Their

following study, with the same data and design that included youth ranging in age from four to 13-year-old children, showed that familial confounds attenuated the association between SDP and most of the developmental outcomes including ODD and CD symptoms.[13] Maughan et al.[24] found a robust association between SDP and multi-rater (parent and teacher) report of CD symptoms, but after adjusting for parental antisocial behavior and genetic factors, only heavy SDP was associated with CD. Thus, previous, seemingly robust, associations between SDP and later behavioral problems may be a function of familial confounding as opposed to the teratogenic effect of SDP.

An important factor, and potential gap in the current literature, is reducing measurement error in complex behaviors when examining the role of SDP in the development of youth disruptive behavior. ODD and CD share many similar symptoms as they both rest on a continuum of youth disruptive behavior.[27] Further, disruptive behavior in childhood is normative to an extent before it increases to clinically significant DBD.[11,28] Given the complexity of disruptive behavior, proper assessment of symptoms by multiple raters is essential, (e.g., to capture disruptive behavior symptoms from multiple contexts by different assessors).[11] The various assessment methods of disruptive behavior include DSM-based interviewing methods,[11,29] parent, teacher, or youth rating scales,[11,30,31] and observational assessment location (e.g., in school, at home, or during testing).[11,32] Parent and teacher report of disruptive behavior might not necessarily measure the same construct and in fact, in the field of ADHD, this has been suggested to be the case. Parent and teacher reports may reflect valid and important differences in behavior and functional impairment across home and school settings (e.g., Narad et al.[33], Piancentini et al.[34], Sutin et al.[35]). Further, Hart and colleagues[36] suggest that there is evidence that parent and teacher ratings account for unique variance in functional impairment and capture a unique perspective that is important to consider. Thus, the use of multiple raters allows researchers to capture a broader measure of disruptive behavior symptoms in non-clinical samples. This is particularly important when looking at the effects of SDP on child behavioral outcomes as parents have personal knowledge of their own smoking behavior during the pregnancy with each child, which has the potential to bias their ratings.[37,38]

To accommodate both parent and teacher ratings, the DSM-IV field trials and prior work in the externalizing behavior domain have used an algorithm in which each symptom reported by either the parent or teacher during a structured interview was counted as a positive symptom.[38-41] Thus, an approach that tests associations with a composite “or” rule measure would most accurately reflect clinical symptom levels across home and school settings from a DSM-IV perspective.[34] Surprisingly, it has been rarely used within genetically-informed studies of the association between SDP and child behavioral outcomes.[38] Most studies testing the SDP-DBD association to date have used parent reports on rating scales.[12,13,20] A multi-rater approach to the SDP-DBD association allows a more comprehensive investigation of the potential causal pathways behind disruptive behavior, and may offer an avenue to identify youth that are at risk for later maladjustment. Because such problems are likely to escalate over time if left untreated, and therefore, can predispose the child to later psychiatric problems,[42] it is important to examine potential critical developmental windows for prevention and intervention. In order to address this gap in the literature, the current study utilizes multiple raters (parent and teacher) to determine optimal symptom counts of both ODD and CD separately, and together as DBD, that more comprehensively captures youth behavior problems from multiple contexts and situations.

The sibling comparison design is one of several quasi-experimental designs that can be used to strengthen causal inferences regarding environmental risks by ruling out specific forms of confounding,

including confounding by gene-environment correlation.[7,26,43] Specifically, the present study sought to examine the potentially causal role of SDP on disruptive behaviors using a sibling comparison design comprised of sibling pairs (aged 7 to 16 years) discordant for SDP and a multi-rater report of youth disruptive behavior. Based on our work with other externalizing disorders in this sample (e.g., Knopik et al.[38], Marceau et al.[44]), we hypothesized that SDP would predict disruptive behavior, above and beyond familial confounding, supporting a potentially causal role of SDP for later youth behavioral problems.

METHODS

Participants and Procedure

Data for the current study were drawn from the Missouri Mothers and Their Children study (MO-MATCH).[45] The MO-MATCH is a family study, specifically designed to leverage the sibling comparison approach to examine the effects of prenatal smoking exposure on child behavioral problems and associated learning and cognitive deficits. Briefly the recruitment methods; families of the MO-MATCH study were identified using birth registers obtained from the Missouri Department of Health and Senior Services Bureau of Health Informatics. Birth register data in Missouri was utilized to recruit mothers who changed smoking behavior between two pregnancies for child's birth years 1998–2005 (N > 4000 identified). A total of 1,520 mothers were contacted for a screening interview to verify SDP information from the birth record. Of those 1,520 screened, 413 of these mothers' interview information (i) confirmed a history of regular smoking (i.e., smoking more than 100 cigarettes lifetime)[46] and (ii) aligned with the birth record. These 413 were invited to participate in the study. The final study population consisted of 173 families including 344 pregnancies. A total of 54% (N = 94) of fathers participated to the study. The recruitment methods, exclusion criteria, and differences between (i) the demographics of the MO-MATCH sample and the demographic of the state of Missouri and between (ii) families where fathers did versus did not participate are described in detail elsewhere (Supplementary material and Knopik et al.[45]).

The MO-MATCH study sample and assessment methods have also been described in detail previously,[45] and so only an overview of methods used in this particular study is presented here. The MO-MATCH assessments included maternal diagnostic semi-structured interviews about their pregnancies and child mental health and behavioral history. For interview, an age-appropriate version of the Missouri Assessment of Genetics Interview for Children (MAGIC) was used.[29] Mother- and teacher-reported questionnaires on disruptive behavior outcomes were also utilized. Families provided the names of up to two teachers per child to contact for teacher-report questionnaires. Assessments occurred when the children were 7–16 years old and the mean age difference was 2.8 years (SD = 1.54) between the siblings (Child 1 [older sibling] average age = 12.99, standard deviation [SD] = 1.95, 54% male; Child 2 [younger sibling] average age = 10.19, SD = 1.80, 51% male). Parents were primarily of Caucasian ancestry (96%, n = 250). Please see Table 1 for additional details on the sample.

The study was approved by the Institutional Review Boards of Rhode Island Hospital, Washington University, the State of Missouri Department of Health and Senior Services, and Purdue University.

Measures

SDP. Maternal report of SDP was obtained using a modified version of the MAGIC—Parent on Child.[29] A previous investigation of the MO-MATCH data[47] compared the predictive ability of maternal report of

SDP relative to both birth record report and paternal report of maternal SDP and determined that of these three variables, retrospective self-report of maternal SDP was the best assessment of SDP. Thus, we focus here on maternal retrospective reports of SDP. We created an overall SDP severity score for each child by using the following indicators: (i) *Any SDP* (0 = No, 1 = Yes) across each pregnancy as a whole and specific to each trimester, and (ii) overall quantity smoked during pregnancy via mothers' estimate of the number of cigarettes smoked in each trimester. We incorporated timing of exposure in the SDP severity score as continued SDP has been found to be potentially more harmful.[48,49] Consistent with prior reports,[38,44,47,50,51] a single SDP severity score was created for each child (Supplementary material):

1: did not smoke during pregnancy;

2: smoked during first trimester only, 1–10 cigarettes per day;

3: smoked during first trimester only, 11–19 cigarettes per day;

4: smoked during first trimester only, 20+ cigarettes per day;

5: smoked beyond first trimester, 1–10 cigarettes per day (max of all trimesters);

6: smoked beyond first trimester, 11–19 cigarettes per day (max of all trimesters); and

7: smoked beyond first trimester, 20+ cigarettes per day (max of all trimesters).

Disruptive behavior. Disruptive behavior was assessed separately as symptom counts of (i) Oppositional Defiant Disorder (ODD), (ii) Conduct Disorder (CD), and (iii) a combination of ODD and CD (referred as DBD) scores by using multiple raters (i.e., parents and teachers) and assessment tools. Disruptive behavior symptoms were assessed via mother report on the MAGIC–Parent on Child DSM-IV semi-structured interview[29] and the child behavior checklist (CBCL),[30] as well as teacher report form (TRF).[31] The TRF was available for 220 of 344 children. TRF data was missing at random in relation to birth order (Chi-Square, p=0.71) and SDP severity (Chi-Square, p=0.23). Further, ninety-five children had TRF from two different. Therefore, we used the "or" rule for each item in these cases to obtain a single, maximum teacher-rated symptom score per child (for more detail see Knopik et al.[45]).[34,52] Two TRFs resulted in higher symptoms scores compared to only one TRF. Having two vs. only one TRF was random in relation to birth order (Chi-Square, p=0.27) and SDP severity (Chi-Square, p=0.62).

The CBCL and TRF assess behavioral symptoms over the past 6 months on a scale of 0 (not true) to 2 (very true/often true). Symptom scores according to the CBCL, measuring ODD, included five items (α = 0.77) and, measuring CD, included 17 items (α = 0.84). Symptom scores according to the TRF, measuring ODD included five items (α = 0.79) and, measuring CD, included 13 items (α = 0.83). Thus, the total combined DBD symptom scores (a sum of ODD and CD scores) according to the CBCL (α = 0.88) included 22 items and according to the TRF (α = 0.89) included 18 items. A multi-rater composite score of DBD (between parent and teacher reports) was created in the same way as our previous study.[38] Specifically, this multi-rater was created by applying the "or" rule for each item assessed both in CBCL and TRF and for whom the data was available.[34,52] Some items were asked only on the CBCL or on the TRF; and these were included in the maximized multi-rater composite. Consistent with prior work,[34,47] in the case of missing TRF, the parent report (CBCL) alone was used to create the multi-rater score. The CBCL/TRF composite score for ODD consisted of 6 items (α = 0.78), CD of 18 items (α = 0.85), and DBD of 24 items

(α = 0.89). The outcome scores for the CBCL, TRF, and multi-rater composites were square root transformed due to skew in the residuals of the hypothesis-testing models. Table 1 depicts the descriptive statistics for the sample.

The MAGIC–Parent on Child[29] is a DSM-IV semi-structured interview where mothers answered yes/no whether the child exhibited each of the eight DSM-IV ODD and 15 CD symptoms within the past year. We created sum scores indicating the number of symptoms endorsed within the ODD (α = 0.76), the CD (α = 0.63), and total DBD symptoms (α = 0.78). The MAGIC symptom scores were also square root transformed due to skew in the residuals of the hypothesis-testing models. The correlations of the outcome variables are shown in Supplementary Table 1.

Child and Familial covariates. Covariates were chosen to be consistent with other genetically-informed studies of SDP-behavioral problem associations.[12,38,44,53] The covariates included 1) maternal characteristics: marital status, age, and education at birth of each child, 2) child characteristics: birth order, sex, 3) second-hand smoke exposure during pregnancy by the father, and 4) families' qualification for food stamps at the time of delivery. The information on qualification for food stamps was collected from the birth record and was used to control for other maternal and family characteristics that potentially confound the association of SDP and disruptive behavior symptoms. Birth order was significantly (and negatively) correlated with age (r= -0.87), which leads to a multicollinearity problem when modeling these data. In order to be consistent with previous studies,[38,44] we included birth order as a covariate because the majority of the mothers smoked in the first pregnancy (64%) but not the second. Supplementary analyses with age as a covariate instead of birth order resulted in no change in our main findings (see supplementary materials and Supplementary Table 2).

Statistical Analysis

Given the use of a sibling comparison approach and the need to account for non-independence of data, a series of hierarchical linear models (HLM) were executed using SAS PROC MIXED (SAS 9.4 System. Cary, NC, SAS Institute Inc). HLM enables researchers to assess the within- and between-family associations of SDP with disruptive behavior. The analytic approach (Figure 1) used here is identical to the approach detailed in previous studies on the association between SDP and behavioral outcomes such as ADHD and inhibitory control. [38,44,51]

Unconditional Model.

First, an unconditional 'intercept-only' model was fitted to the data to decompose the variance in the symptom measures into within-family (e.g. individual child-level) and between-family (e.g. family level) variation via intra-class correlations (ICC, Supplementary material).[54]

Standard Model.

The standard model (Model 1) can be thought of as an approximation of the standard between-family comparison seen in the literature. This model does not capitalize on the family structure (or sibling comparison aspect) of the data; although, it does adjust for the nesting structure of children within families. The covariates included child sex, child birth order, mother education, maternal age, marital status, food stamp usage at birth, and prenatal second-hand smoke exposure by fathers.

Sibling Comparison Model.

The sibling comparison model (Model 2) allows for the evaluation of siblings within families that are discordant for exposure to SDP (i.e., mother smoked [or smoked more] while pregnant with one sibling and did not smoke [or smoked less] while pregnant with the other sibling). By nature of the design, siblings are matched for confounding familial factors. In this model, we examined whether SDP is operating at the within-family level (e.g., contributing to differences in disruptive behavior symptoms in one sibling vs. another, within families) or between-family level (e.g., contributing to differences in overall, average levels of siblings' disruptive behavior symptoms across families).

In Model 2, to specifically assess both between- and within-family associations of SDP and disruptive behavior, a family SDP average score and a child-specific SDP severity scores were used. "Family average SDP severity" for each family was the average score for SDP severity across both siblings, which was used to assess the between-family effect of SDP on disruptive behavior. "Child-specific SDP severity" for each child was the resulting value when the family average SDP was subtracted from each child-specific SDP severity score used in Model 1 (i.e., within-family centering). Thus, if a mother smoked the same amount for both pregnancies, regardless of total severity, then both siblings would have a "child-specific SDP severity" score of zero. Correspondingly, the sibling for whom mothers smoked (or smoked more) would have a positive score of "child-specific SDP severity", whereas the sibling for whom mothers did not smoke (or smoked less) would have a negative score of the same magnitude.[47]

In this model, the "child-specific SDP severity" score was entered as a Level 1 predictor and assessed the within-family effect of SDP on disruptive behavior (i.e. comparing across siblings within a family). The "family-average SDP" score (assessing the between-family effect) was entered as a Level 2 predictor. The analysis for Model 2 allows the examination of any unique effect of SDP on child specific outcomes over and above familial and genetic factors that siblings share, and sibling-level (Level 1) and family-level covariates (a Level 2 variable; Supplementary material). The detailed HLM equation used for Model 2 can be found in Knopik et al.[38]

In sum (Figure 1), for each outcome variable, we fit one unconditional model and two conditional models (Models 1 and 2). In order to quantify how much of the within-family variance is explained by each conditional model, we computed the percentage of the (within-family) variance explained: ([unconditional individual child-level variance—conditional, e.g., Model 1, individual child-level variance]/unconditional individual child-level variance).[55]

RESULTS

Table 2 depicts a summary of the main study findings. More specific information on the models, including between-family effects, covariances, variances, model fit statistics, and percentage of the within-family variance explained by each conditional model are presented in Supplementary Tables **3-5**. In addition, consistent with prior work[38,44,51] Supplementary Table **6** provides a sensitivity analysis incorporating multiple alternative operationalizations of SDP and their associations with all disruptive behavior measures. Throughout most models and consistent with documented sex differences in prevalence rates of DBD (e.g., McCabe et al.[56], Demmer et al.[57]), child sex was generally a significant predictor of disruptive behavior symptoms, such that boys exhibited higher symptoms scores than girls.

Parent Report - Questionnaire

Decompositions of the variance in the unconditional model showed that the majority of the variance in parent-reported disruptive behavior was attributable to within-family differences (DBD: 74%; ODD: 74%; CD: 79%), and the remaining variance (21-26%) was attributable to between-family differences. This suggests that the variance of disruptive behavior by parent-report was primarily a function of differences between siblings rather than differences across families. In the standard model (model 1), SDP severity was not significantly associated with parent-report of disruptive behavior (DBD: b = 0.05, p = 0.19; ODD: b =0.02, p = 0.59; CD: b = 0.06, p = 0.07). Next, we used a sibling comparison model (model 2) to examine the within-family effect (child-specific SDP severity relative to family average) of SDP on parent-report of disruptive behavior. No significant within-family association between SDP and parent-report of disruptive behavior was found (DBD: b = 0.01, p = 0.78; ODD: b = -0.02, p = 0.56; CD: b = 0.04, p = 0.27).

Teacher Report - Questionnaire

The variance in teacher-report of disruptive behavior was primarily a function of differences between siblings (within-family variability; DBD: 95%; ODD: 100%; CD: 82%) rather than differences across families (between-family differences; 0-18%). In the standard model (model 1), SDP severity predicted more teacher-reported DBD symptoms (b = 0.12, p = 0.02) and CD symptoms (b = 0.10, p = 0.01). There was also a trend-level association between SDP severity and ODD symptoms (b = 0.07, p = 0.08). However, the sibling comparison model (model 2) assessing within-family effect of SDP for teacher-report of disruptive behavior showed only a trend-level association on DBD (b = 0.11, p = 0.07) and CD symptoms (b = 08, p = 0.07), and no significant association on ODD symptoms (b = 0.06, p = 0.19).

Multi-rater Composites - Questionnaire

The variance in multi-rater composite score of disruptive behavior was primarily a function of differences between siblings (within-family variability; DBD: 75%; ODD: 72%; CD: 82%) rather than differences across families (between-family differences; 18-25%). In the standard model (model 1), SDP severity predicted increased symptoms of disruptive behavior (DBD: b = 0.11, p = 0.003; ODD: b = 0.06, p = 0.03; CD: b = 0.10, p = 0.002). In the sibling comparison model (model 2), there was a significant within-family effect of SDP on the multi-rater composite score of DBD (b = 0.09, p = 0.03) and CD (b = 0.09, p = 0.01), but not on ODD (b = 0.04, p = 0.19). In general, the pattern of results remained in sensitivity analyses that considered different variations of SDP (Supplementary Table **6**).

Parent Report - Interview-Based

In interview-based maternal reports of disruptive behavior, a smaller proportion of the variance was attributable to within-family differences (DBD: 66%; ODD: 75%; CD: 76%) compared to the other measurements of disruptive behavior. The between-family differences in variance ranged from 24-34%. In the standard model (model 1), SDP severity was related to all interview-based disruptive behavior symptoms across measures (DBD: b = 0.06, p = 0.005; ODD: b = 0.06, p = 0.01; CD b = 0.04, p = 0.01). In the sibling comparison model (model 2), the SDP severity score showed only a trend-level association on CD symptoms (b = 0.03, p = 0.07) and was unrelated to the other interview-based disruptive behavior symptoms scores (DBD: b = 0.04, p = 0.10; ODD: b = 0.03, p = 0.26).

DISCUSSION

This is the first genetically-informed report of the association between SDP and disruptive behavior that considers multiple assessment methods (parent questionnaire, teacher questionnaire, and maternal interview report), as well as symptom dimensions (ODD, CD, and DBD). Consistent with prior genetically-informed approaches,[12,13] the associations between SDP and disruptive behavior, as measured by DSM-IV-based interview as well as symptom-based CBCL and TRF, were primarily explained by familial confounds rather than a potentially causal teratogenic effect. However, we found a significant association between SDP and both DBD and CD symptoms using a multi-rater approach that capitalizes on both parent and teacher report (CBCL and TRF). This pattern of results remained in sensitivity analysis. With regard to the multi-rater findings, a similar positive association between SDP and CD was reported by Gaysina et al.[17] in their study using a complementary genetically sensitive research design from three different studies including two with parent report and one with multi-rater measurements. Overall, the current study provides additional evidence that controlling for genetic and family factors is essential when examining SDP as an important risk factor for later behavioral outcomes and further points to the importance of considering child behavior in quasi-experimental designs of SDP.

A unique aspect of the current study was the use of multiple raters (i.e. parent and teacher reports) and varying measures of disruptive behavior (i.e., ODD, CD, and a composite DBD score).). In particular, the multi-rater approach encompassed greater symptoms across raters and environments, and possibly allows us to better capture youths' behavior. Importantly, in prior studies comparing parent and teacher reports of externalizing behavior problems, teachers' reports were more indicative of referrals to mental health services[58] and more predictive of future behavioral disorders.[35] However, both parents and teachers were uniquely important to consider together, as one cannot act as a proxy for the other because they offer measures of children in distinct contexts and situations.[58] In our analyses, neither teacher nor parent report alone reached significance. This could be related to decreased power for teacher reports, as there was data on teacher report for only 220 of 344 children leading to fewer complete sibling pairs in the analyses that used teacher report. However, teacher reports seem to drive our significant findings in the multi-rater composite scores as we found trend-level associations between SDP severity and teacher-report of DBD and CD and not with parent reports. Critically, teachers' reports are important, as they are able to report on child behavior relative to a broad reference group and in an environment where youth have greater opportunity to engage in problem behavior.[37] In addition, teachers may be privy to children's behavior in a critical context outside the home where increased, clinical level, symptoms occur. In sum, our findings suggest that when examining complex behaviors, including a comprehensive measure of disruptive behavior may be important when examining the role of SDP.

The strength of this study is the use of data from a specifically designed family study, which leverages the sibling comparison approach (with siblings discordant for SDP) to examine the effects of SDP on child behavioral outcomes. The use of more comprehensive study designs is fundamental when disentangling the possible effects of SDP from the genetic and environmental factors on child behavioral outcomes. However, findings should be interpreted in light of the following limitations. First, retrospective report of SDP is used and biochemical confirmation of SDP (e.g., maternal saliva or serum cotinine measurements) is not available. Biochemical validation can decrease misclassification of SDP.[59] Further, there is also a risk for recall bias when using retrospective information of SDP, especially for assessing specific details about the quantity of SDP.[60] However, based on our recruitment methods, retrospective report of SDP has been shown to be reliable and accurate in this sample.[47] To address possible

limitations of SDP measurement, we performed sensitivity analyses that consider timing, duration, and amount of SDP and found our findings to be consistent across SDP measures. Second, while we have not examined reasons why these mothers changed their smoking behaviors from one pregnancy to another, we do have some data to address this and will examine these questions in depth in future extensions of this paper.

Third, our study sample is limited by sample size; in particular, we have fewer teacher reports than parent report, resulting in a smaller sample size for the separate teacher analyses (i.e. in which a trend for statistical significance was found). However, including teacher report in this study offers a more objective examination of SDP-disruptive behavior associations as the teachers are unaware of SDP exposure and therefore lack bias that parents may experience when reporting on both SDP and their children's behavior. In 95 cases, multiple teachers reported on a single participant. This could bias our findings, such that those youth would have higher reports of disruptive behavior compared to those who did not have a teacher report or who had a single teacher report on behavior. A fourth potential limitation is the age of the children (varying from 7 to 16 years) because the children may not have passed through the period of risk for exhibiting disruptive behavior symptoms, which might contribute to low prevalence of symptoms and power. Some of these limitations may explain the null findings in the standard model in parent-reported symptom dimensions, which contradicted findings from previous studies using samples of similar ages to MO-MATCH.[13,20]

Fifth, we are also aware of the number of tests conducted. While the pattern of results remain consistent in sensitivity analyses using various definitions of SDP, significant multi-rater results do not survive a conservative Bonferroni correction. Thus, further work with larger samples needs to be done to determine how best to interpret the information provided by different reporters. Finally, despite our comprehensive approach utilizing the sibling comparison design, there are undoubtedly unmeasured variables that differ between siblings that are not included in these analyses.[26]

In conclusion, significant associations between SDP and disruptive behavior found in the standard model were explained by familial and genetic confounds in the sibling comparison model, with the exception of the multiple rater model. A potentially causal effect of SDP on disruptive behavior was seen when using a multi-rater approach capitalizing on both parent and teacher report of disruptive behavior in the sibling comparison model; however, results need to be interpreted with caution given limitations noted above. Critically, a novel aspect of our study highlights the importance of incorporating multiple raters to capture complex behavior in youth. These findings suggest that familial confounds as well as both parent and teacher reports are important considerations in the potential association between SDP and disruptive behavior in non-clinical samples. It is clear that genetic and familial factors explain at least some proportion of the previously found robust association between SDP and later behavioral problems. However, smoking pregnant women should be encouraged and supported to partake in smoking cessation because the adverse effects of SDP on fetal health and birth outcomes are undisputable.

Figure 1. The Analysis Plan of the Study

FUNDING

This work was supported by grants to Dr. Ekblad: the Foundation for Pediatric Research (Finland), the Orion Research Foundation sr, the Emil Aaltonen's Foundation, the Paulo Foundation, the Maud Kuistila Memorial Foundation, and the Turku University Hospital Research Foundation. The data collection was

funded by the National Institute of Health; Grant numbers: DA023134 (Knopik), DA17671 (Knopik), AA07728 (Heath), AA09022 (Heath), AA11998 (Heath), HD049024 (Heath), AA017688 (Heath), AA021492 (Heath), R01DA04272 (Palmer), K01DA039288 (Marceau).

DECLARATION OF INTERESTS

None declared.

REFERENCES

1. Lange S, Probst C, Rehm J, Popova S. National, regional, and global prevalence of smoking during pregnancy in the general population: a systematic review and meta-analysis. *Lancet Glob Health.* 2018;6(7):e769-e76.

2. Ekblad M, Gissler M, Korkeila J, Lehtonen L. Trends and risk groups for smoking during pregnancy in Finland and other Nordic countries. *Eur J Public Health.* 2014;24(4):544-51.

3. Bublitz MH, Stroud LR. Maternal smoking during pregnancy and offspring brain structure and function: review and agenda for future research. *Nicotine Tob Res.* 2012;14(4):388–97.

4. Chatterton Z, Hartley BJ, Seok MH, et al. In utero exposure to maternal smoking is associated with DNA methylation alterations and reduced neuronal content in the developing fetal brain. *Epigenetics Chromatin.* 2017;10:4.

5. Ekblad M, Korkeila J, Lehtonen L. Smoking during pregnancy affects foetal brain development. *Acta Paediatr.* 2015;104(1):12-8.

6. Ekblad M, Gissler M, Lehtonen L, Korkeila J. Prenatal smoking exposure and the risk of psychiatric morbidity into young adulthood. *Arch Gen Psychiatry.* 2010;67(8):841-9.

7. Knopik VS. Maternal smoking during pregnancy and child outcomes: Real or spurious effect? *Dev Neuropsychol.* 2009;34(1):1–36.

8. Latimer K, Wilson P, Kemp J, et al. Disruptive behaviour disorders: a systematic review of environmental antenatal and early years risk factors. *Child Care Health Dev.* 2012;38(5):611-28.

9. Tolan PH, Leventhal BL. Introduction: Connecting brain development, disruptive behavior, and children. In Disruptive Behavior Disorders. Springer, New York, NY; 2013:1-11.

10. Knopik VS, Bidwell LC, Flessner C, et al. DSM-IV defined conduct disorder and oppositional defiant disorder: an investigation of shared liability in female twins. *Psychol Med.* 2014;44(5):1053-64.

11. McKinney C, Morse M. Assessment of Disruptive Behavior Disorders: Tools and Recommendations. *Professional Psychology: Research and Practice.* Advance online publication. 2012. doi: 10.1037/a0027324.

12. D'Onofrio BM, Van Hulle CA, Waldman ID, et al. Smoking during pregnancy and offspring externalizing problems: an exploration of genetic and environmental confounds. *Dev Psychopathol.* 2008;20(1):139-64.

13. Ellingson JM, Goodnight JA, Van Hulle CA, Waldman ID, D'Onofrio BM. A sibling-comparison study of smoking during pregnancy and childhood psychological traits. *Behav Genet.* 2014;44(1):25-35.

14. Ellis LC, Berg-Nielsen TS, Lydersen S, Wichstrøm L. Smoking during pregnancy and psychiatric disorders in preschoolers. *Eur Child Adolesc Psychiatry.* 2012;21(11):635-44.

15. Ekblad M, Lehtonen L, Korkeila J, Gissler M. Maternal Smoking during Pregnancy and the Risk of Psychiatric Morbidity in Singleton Sibling Pairs. *Nicotine Tob Res.* 2017;19(5):597-604.

16. Estabrook R, Massey SH, Clark CA, et al. Separating Family-Level and Direct Exposure Effects of Smoking During Pregnancy on Offspring Externalizing Symptoms: Bridging the Behavior Genetic and Behavior Teratologic Divide. *Behav Genet.* 2016;46(3):389-402.

17. Gaysina D, Fergusson DM, Leve LD, et al. Maternal smoking during pregnancy and offspring conduct problems: evidence from 3 independent genetically sensitive research designs. *JAMA Psychiatry.* 2013;70(9):956-63.

18. Marceau K, Rolan E, Leve LD, et al. Parenting and prenatal risk as moderators of genetic influences on conduct problems during middle childhood. *Dev Psychol.* 2019. doi: 10.1037/dev0000701. [Epub ahead of print]

19. Wakschlag LS, Lahey BB, Loeber R, Green SM, Gordon RA, Leventhal BL. Maternal smoking during pregnancy and the risk of conduct disorder in boys. *Arch Gen Psychiatry.* 1997;54(7):670-6.

20. Dolan CV, Geels L, Vink JM, et al. Testing Causal Effects of Maternal Smoking During Pregnancy on Offspring's Externalizing and Internalizing Behavior. *Behav Genet.* 2016;46(3):378-88.

21. Ruisch IH, Dietrich A, Glennon JC, Buitelaar JK, Hoekstra PJ. Maternal substance use during pregnancy and offspring conduct problems: A meta-analysis. *Neurosci Biobehav Rev.* 2018;84:325-36.

22. Wakschlag L, Pickett K, Cook E, Benowitz N, Leventhal B. Maternal smoking during pregnancy and severe antisocial behavior in offspring: A review. *Am J Public Health.* 2002;92:966–74.

23. Robinson M, McLean NJ, Oddy WH, et al. Smoking cessation in pregnancy and the risk of child behavioural problems: a longitudinal prospective cohort study. *J Epidemiol Community Health.* 2010;64(7):622-9.

24. Maughan B, Taylor C, Taylor A, Butler N, Bynner J. Pregnancy smoking and childhood conduct problems: a causal association? *J Child Psychol Psychiatry*. 2001;42(8):1021-8.

25. Agrawal A, Knopik VS, Pergadia ML, et al. Correlates of cigarette smoking during pregnancy and its genetic and environmental overlap with nicotine dependence. *Nicotine Tob Res.* 2008;10(4):567-78.

26. D'Onofrio BM, Lahey BB, Turkheimer E, Lichtenstein P. Critical need for family-based, quasi-experimental designs in integrating genetic and social science research. *Am J Public Health.* 2013;103 Suppl 1:S46-55.

27. de la Osa N, Granero R, Trepat E, Domenech JM, Ezpeleta L. The discriminative capacity of CBCL/1½-5-DSM5 scales to identify disruptive and internalizing disorders in preschool children. *Eur Child Adolesc Psychiatry*. 2016;25(1):17-23.

28. McKinney C, Renk K. Emerging research and theory in the etiology of oppositional defiant disorder: Current concerns and future directions. *IJBCT.* 2007;3:349–71.

29. Todd RD, Joyner CA, Heath AC, Neuman RJ, Reich W. Reliability and stability of a semistructured DSM–IV interview designed for family studies. *J Am Acad Child Adolesc Psychiatry.* 2003;42(12):1460-8.

30. Achenbach TM. Manual for the Child Behavior Checklist/4–18 and 1991 profile. Burlington, VT: University of Vermont, Department of Psychiatry. 1991.

31. Achenbach TM. Manual of the Teacher's Report Form and 1991 profile. Burlington, VT: University of Vermont, Department of Psychiatry. 1991.

32. Wakschlag LS, Briggs-Gowan MJ, Carter AS, et al. A developmental framework for distinguishing disruptive behavior from normative misbehavior in preschool children. *J Child Psychol Psychiatry.* 2007;48(10):976-87.

33. Narad ME, Garner AA, Peugh JL, et al. Parent-teacher agreement on ADHD symptoms across development. *Psychol Assess.* 2015;27(1):239-48.

34. Piacentini JC, Cohen P, Cohen J. Combining discrepant diagnostic information from multiple sources: are complex algorithms better than simple ones? *J Abnorm Child Psychol.* 1992;20(1):51-63.

35. Sutin AR, Flynn HA, Terracciano A. Maternal cigarette smoking during pregnancy and the trajectory of externalizing and internalizing symptoms across childhood: Similarities and differences across parent, teacher, and self reports. *J Psychiatr Res.* 2017;91:145-8.

36. Hart EL, Lahey BB, Loeber R, Hanson KS. Criterion validity of informants in the diagnosis of disruptive behavior disorders in children: a preliminary study. *J Consult Clin Psychol.* 1994;62(2):410-4.

37. Deater-Deckard K, Plomin R. An adoption study of etiology of teacher and parent reports of externalizing behavior problems in middle childhood. *Child Development.* 1999;70(1):144-54.

38. Knopik VS, Marceau K, Bidwell LC, et al. Smoking during pregnancy and ADHD risk: A genetically informed, multiple-rater approach. *Am J Med Genet B Neuropsychiatr Genet.* 2016;171(7):971–81.

39. Bidwell LC, Willcutt EG, McQueen MB, et al. A family based association study of DRD4, DAT1, and 5HTT and continuous traits of attention-deficit hyperactivity disorder. *Behav Genet.* 2011;41(1):165-74.

40. Lahey BB, Applegate B, McBurnett K, et al. DSM-IV field trials for attention deficit hyperactivity disorder in children and adolescents. *Am J Psychiatry.* 1994;151(11):1673-85.

41. Willcutt EG, Chhabildas N, Kinnear M, et al. The internal and external validity of sluggish cognitive tempo and its relation with DSM-IV ADHD. *J Abnorm Child Psychol.* 2014;42(1):21-35.

42. Pelletier J, Collett B, Gimpel G, Crowley S. Assessment of disruptive behaviors in preschoolers: Psychometric properties of the disruptive behavior disorders rating scale and school situations questionnaire. *Journal of Psychoeducational Assessment.* 2006;24(1):3–18.

43. Kendler KS. Causal Inference in Psychiatric Epidemiology. *JAMA Psychiatry.* 2017;74(6):561-2.

44. Marceau K, Cinnamon Bidwell L, Karoly HC, et al. Within-Family Effects of Smoking during Pregnancy on ADHD: the Importance of Phenotype. *J Abnorm Child Psychol*. 2018;46(4):685-99.

45. Knopik VS, Heath AC, Marceau K, et al. Missouri mothers and their children: A family study of the effects of genetics and the prenatal environment. *Twin Res Hum Genet.* 2015;18(5):485–96.

46. Ryan H, Trosclair A, Gfroerer J. Adult current smoking: differences in definitions and prevalence estimates--NHIS and NSDUH, 2008. *J Environ Public Health.* 2012;2012:918368.

47. Knopik VS, Marceau K, Palmer RH, Smith TF, Heath AC. Maternal smoking during pregnancy and offspring birth weight: A genetically-informed approach comparing multiple raters. *Behav Genet*. 2016;46(3):353–64.

48. Dwyer JB, McQuown SC, Leslie FM. The dynamic effects of nicotine on the developing brain. *Pharmacol Ther.* 2009;122(2):125–39.

49. Hebel JR, Fox NL, Sexton M. Dose-response of birth weight to various measures of maternal smoking during pregnancy. *J Clin Epidemiol.* 1988;41:483–9.

50. Bidwell LC, Marceau K, Brick LA, et al. Prenatal Exposure Effects on Early Adolescent Substance Use: Preliminary Evidence From a Genetically Informed Bayesian Approach. *J Stud Alcohol Drugs.* 2017;78(5):789-94.

Note: References 51 onwards are listed in supplementary material.

Table 1. Sample characteristics

Study variables	Child 1 (older sibling)				Child 2 (younger sibling)			
	N	Mean	(SD)	Variance	N	Mean	(SD)	Variance
SDP Severity	173	3.95	2.05		171	2.04	1.77	
Disruptive behavior disorder								
CBCL	162	1.76	1.28	1.65	160	1.74	1.26	1.58
TRF	109	1.04	1.27	1.61	111	1.38	1.29	1.68
CBCL/TRF multi-rater total	165	2.13	1.25	1.56	164	2.13	1.31	1.70
MAGIC	173	1.09	0.84	0.71	171	1.03	0.82	0.67
Oppositional defiant disorder								
CBCL	162	1.21	0.92	0.85	160	1.19	0.90	0.81
TRF	109	0.65	0.85	0.73	111	0.86	0.94	0.87
CBCL/TRF multi-rater total	165	1.48	0.83	0.69	164	1.48	0.89	0.79
MAGIC	173	0.97	0.76	0.58	171	0.94	0.73	0.27
Conduct disorder								
CBCL	162	1.09	1.12	1.25	160	1.10	1.09	1.18
TRF	109	0.73	1.00	1.00	111	0.93	1.06	1.12
CBCL/TRF multi-rater total	165	1.37	1.16	1.34	164	1.36	1.19	1.40
MAGIC	173	0.27	0.56	0.58	171	0.23	0.52	0.53
Covariates								
Maternal age at birth	162	26.48	5.55		163	29.22	5.75	
Maternal education (in years) at birth	162	13.28	2.12		163	13.5	1.94	
Second-hand smoke exposure by fathers	171	1.84	1.44		161	1.15	1.43	
	N	%			N	%		
Sex (male)	91	53%			87	51%		
Marital status (percent married) at birth	155	85%			159	83%		
Food stamp usage at birth	149	9%			150	13%		
Family demographics (at assessment)	N	Mean	(SD)					
Maternal age	162	39.83	5.62					
Paternal age	80	44.04	6.34					
Child 1 age	173	12.99	1.95					
Child 2 age	170	10.19	1.80					
Child age difference	170	2.79	1.54					
	Maternal				Paternal			
	N	%			N	%		
Education								
Less than HS	7	4%			9	10%		
HS	30	18%			19	20%		
1–2 years college	50	30%			14	15%		
3–4 years college	46	27%			17	18%		
More than college	29	17%			21	22%		
Not reported	7	4%			14	15%		
Mother's marital status								
Never married	6	4%						
Married	130	77%						
Separated	5	3%						
Divorced	26	15%						
Widowed	2	1%						

Table 1. Sample characteristics cont.

	Maternal		Paternal	
	N	%	N	%
Parent psychopathology diagnosis (MAGIC Adult on Self interview)				
Conduct disorder	1	<1%	1	1%
Present oppositional defiant disorder	2	1%	0	
Lifetime ADHD (any type)	9	5%	7	7%
Present depressive episode	8	5%	0	
Generalized anxiety	12	7%	0	

Smoking during pregnancy (SDP), Child Behavior Checklist (CBCL), Teacher Report Form (TRF), Missouri Assessment of Genetics Interview for Children-Parent Interview (MAGIC), High School (HS). N's reflect the sample size with valid data and vary due to random missing data in the birth record. The outcome composite scores were square root transformed due to skew in the residuals.

Table 2. Summary of SDP-Disruptive Behavior Associations

	Model 1 - Standard	Model 2 - Sib-comparison
Outcome variable	Child-specific SDP severity	Child-specific SDP severity relative to family average (within-family effect)
Parent report (CBCL)	*b* (SE)	*b* (SE)
DBD	0.05 (0.04)	0.01 (0.04)
ODD	0.02 (0.03)	-0.02 (0.03)
CD	0.06† (0.03)	0.04 (0.04)
Teacher report (TRF)		
DBD	0.12* (0.05)	0.11† (0.06)
ODD	0.07† (0.04)	0.06 (0.05)
CD	0.10* (0.04)	0.08† (0.05)
Multi-rater (CBCL/TRF)		
DBD	0.11* (0.04)	0.09* (0.04)
ODD	0.06* (0.03)	0.04 (0.03)
CD	0.10* (0.03)	0.09* (0.04)
Interview-based (MAGIC)		
DBD	0.06* (0.02)	0.04 (0.03)
ODD	0.06* (0.02)	0.03 (0.02)
CD	0.04* (0.02)	0.03† (0.02)

Unstandardized beta-weights (b), Standard Error (SE), *P<0.05, †P<0.10. Each row represents a different outcome variable and summarizes only the SDP severity findings from the larger models noted along the top. All parameter and variance estimates and model fit information from each of the models are provided in supplemental materials.