The Journal of Child Psychology and Psychiatry

Journal of Child Psychology and Psychiatry **:* (2024), pp **–** doi:10.1111/jcpp.13990

Contamination bias in the estimation of child maltreatment causal effects on adolescent internalizing and externalizing behavior problems

John M. Felt,[1] Ulziimaa Chimed-Ochir,[2] Kenneth A. Shores,[3] Anneke E. Olson,[2] Yanling Li,[2] Zachary F. Fisher,[2] Nilam Ram,[4,5] and Chad E. Shenk[2,6]

[1]Center for Healthy Aging, The Pennsylvania State University, University Park, PA, USA; [2]Department of Human Development and Family Studies, The Pennsylvania State University, University Park, PA, USA; [3]School of Education, University of Delaware, Newark, DE, USA; [4]Department of Communications, Stanford University, Stanford, CA, USA; [5]Department of Psychology, Stanford University, Stanford, CA, USA; [6]Department of Pediatrics, The Pennsylvania State University College of Medicine, Hershey, PA, USA

Background: When unaddressed, contamination in child maltreatment research, in which some proportion of children recruited for a nonmaltreated comparison group are exposed to maltreatment, downwardly biases the significance and magnitude of effect size estimates. This study extends previous contamination research by investigating how a dual-measurement strategy of detecting and controlling contamination impacts causal effect size estimates of child behavior problems. **Methods:** This study included 634 children from the LONGSCAN study with 63 cases of confirmed child maltreatment after age 8 and 571 cases without confirmed child maltreatment. Confirmed child maltreatment and internalizing and externalizing behaviors were recorded every 2 years between ages 4 and 16. Contamination in the nonmaltreated comparison group was identified and controlled by either a prospective self-report assessment at ages 12, 14, and 16 or by a one-time retrospective self-report assessment at age 18. Synthetic control methods were used to establish causal effects and quantify the impact of contamination when it was not controlled, when it was controlled for by prospective self-reports, and when it was controlled for by retrospective self-reports. **Results:** Rates of contamination ranged from 62% to 67%. Without controlling for contamination, causal effect size estimates for internalizing behaviors were not statistically significant. Causal effects only became statistically significant after controlling contamination identified from either prospective or retrospective reports and effect sizes increased by between 17% and 54%. Controlling contamination had a smaller impact on effect size increases for externalizing behaviors but did produce a statistically significant overall effect, relative to the model ignoring contamination, when prospective methods were used. **Conclusions:** The presence of contamination in a nonmaltreated comparison group can underestimate the magnitude and statistical significance of causal effect size estimates, especially when investigating internalizing behavior problems. Addressing contamination can facilitate the replication of results across studies. **Keywords:** Child maltreatment; contamination; synthetic control method; causal estimation; internalizing behaviors; externalizing behaviors.

Introduction

Contamination occurs when members of an established control or comparison condition seek out, receive, or are exposed to the treatment under investigation (Cuzick, Edwards, & Segnan, 1997; Delgado-Rodríguez & Llorca, 2004). The presence of contamination artificially minimizes treatment effects between groups, resulting in a downward bias when estimating the significance and magnitude of treatment effects (Hirano, Imbens, Rubin, & Zhou, 2000; Jo, 2002; Magill, Knight, McCrone, Ismail, & Landau, 2019). Contamination has recently been examined in observational research on child maltreatment, where some proportion of individuals recruited to a nonchild maltreatment comparison group were exposed to maltreatment. Contamination downwardly biased the significance and magnitude of child maltreatment effects, spanning a range of physical and psychiatric health outcomes assessed in childhood, adolescence, and young adulthood (Scott, Smith, & Ellis, 2010; Shenk, Noll, Peugh, Griffin, & Bensman, 2016; Shenk, Rausch, Shores, Allen, & Olson, 2022). As in other substantive areas (Craven, Marsh, Debus, & Jayasinghe, 2001; Roobol et al., 2009; Simmons et al., 2015), failure to detect and control contamination in child maltreatment research can bias treatment effects in a manner that results in discovery or replication failures and underestimate the developmental consequences of maltreatment.

Existing child maltreatment research has dealt with contamination using a dual-measurement strategy where official case reports and self-report methods are used to detect contamination. This dual-measurement strategy established and maintained child maltreatment and control conditions using official case reports, which tend to be highly specific indicators of child maltreatment exposure. Then, self-report methods, which tend to be more sensitive relative to official case reports (U.S. Department of Health and Human Services, 2020;

Conflict of interest statement: No conflicts declared.

Finkelhor, Turner, Shattuck, & Hamby, 2015), were used to screen for maltreatment exposure in the established comparison group. Contamination in these studies was identified as any self-reported or subsequently occurring official case report of maltreatment in the established comparison group. Statistical models were then conducted with and without the contaminated cases to illustrate the degree of bias attributable to contamination when estimating child maltreatment effects.

The current study advances research on contamination bias in child maltreatment research in two ways. First, we used a dual-measurement strategy for detecting contaminated cases by establishing and maintaining child maltreatment and comparison groups using official case reports. We then screened the comparison group cases using two different self-report methods, prospective and retrospective, providing an opportunity to evaluate the presence and impact of contamination across commonly used methods for assessing maltreatment. Second, we examine changes in treatment effect estimates for child maltreatment before and after correcting for contamination using synthetic control methods (SCMs; Abadie, Diamond, & Hainmueller, 2010). SCMs are a causal modeling approach appropriate for observational research with longitudinal designs. SCMs can be used to equate child maltreatment and comparison groups on levels of the outcome prior to child maltreatment exposure, creating an effective counterfactual for measuring the effect of maltreatment on subsequent occasions of the outcome. This modeling approach ensures that there are no significant differences between groups on the outcome until the onset of child maltreatment if balancing is achieved, which then accommodates any observed or unobserved time-varying confounders (Abadie et al., 2010). However, causal estimates from the SCM modeling approach will continue to be biased in the presence of contamination. Therefore, contamination will need to be corrected in this, as well as any, modeling approach estimating maltreatment effects (see Shenk et al., 2023) for an in-depth introduction to contamination in child maltreatment research and ways to control it).

We fit SCMs across three different scenarios that reflect realistic degrees of contamination encountered by applied researchers: (a) baseline or no correction of contamination, (b) contamination corrected via prospective self-report assessments, and (c) contamination corrected via retrospective self-report assessments. We fit these models to longitudinal data on children's internalizing and externalizing behavior problems during adolescence, given the strong consensus that child maltreatment affects these outcomes (Gilbert et al., 2009). We hypothesize that exposure to child maltreatment causes increases in adolescent behavior problems and that the significance and magnitude of these effects will be higher when contamination is corrected using the dual-measurement strategy.

Methods

Participants

Participants were from the Longitudinal Studies of Child Abuse and Neglect (LONGSCAN; N = 1,354), a multisite, multi-wave prospective cohort study of child maltreatment in the U.S. (Runyan et al., 1998). Data collection sites were in diverse regions of the U.S., where children were enrolled at or before the age of four and followed every 2 years through Age 18. Caregivers provided consent and parental permission for study participation. Children provided their own assent. Each LONGSCAN data collection site received approval from the respective Institutional Review Board, as well as the LONGSCAN Data Coordinating Center, prior to data collection.

Measures

Confirmed child maltreatment. The Modified Maltreatment Classification System (MMCS; English, Bangdiwala, & Runyan, 2005) was used to establish maltreatment and comparison groups. The MMCS is a research instrument where trained, independent raters review official case reports generated by a Child Protective Services (CPS) investigation of alleged child maltreatment. Raters code whether the information contained in the reports meets prespecified, standardized definitions of child maltreatment, resulting in improved detection of child maltreatment compared to CPS determinations (Runyan et al., 1998). MMCS raters reviewed case records for all LONGSCAN participants from birth to Age 18 and entered scores indicating confirmed child maltreatment (Yes = 1, No = 0) in 2-year intervals.

Contamination: prospective. The LONGSCAN Self-Reports of Physical, Sexual, and Psychological Abuse (SPSPA; Knight et al., 2000) were administered at the Age 12 and Age 16 LONGSCAN assessments. Item content for the SPSPA was generated based on established definitions in the original Maltreatment Classification System (Barnett, Manly, Cicchetti,, 1993) as well as the American Professional Society on the Abuse of Children (Hart, Brassard, & Karlson, 1996). The SPSPA has established psychometric properties that were evaluated using the LONGSCAN cohort (Everson et al., 2008). The Age 12 LONGSCAN assessment measured any prior exposure to child physical, sexual, and psychological abuse and the Age 16 LONGSCAN assessment measured any exposure to these same types of abuse since the Age 12 assessment. An indicator variable was created to indicate, for each child, the presence of any self-reported child abuse occurring prior to age 16 (Yes = 1, No = 0).

The About My Parents (AMP; LONGSCAN Investigators, 1998) scale is a 25-item, self-report measure of child neglect that was administered at the Age 12 LONGSCAN assessment to measure child neglect in elementary school and in the past year and at the Age 14 and Age 16 LONGSCAN assessments to measure child neglect in the past year. AMP items are phrased positively and rated on a 4-point Likert scale ranging from "0 = Never" to "3 = A lot." After removing four items measuring educational neglect and two items measuring supervisory neglect, the remaining 21 items demonstrated acceptable factor structure, measurement invariance, and internal consistencies (Dubowitz et al., 2011). The resulting AMP items were then reverse scored to indicate more severe exposure to child neglect. Endorsement of the most severe rating on any AMP item at any measurement occasion was used to indicate exposure to child neglect prior to age 16 (Yes = 1, No = 0).

Combining the responses obtained from the SPSPA and AMP administered at the Ages 12, 14, and 16 LONGSCAN assessments, a prospective contamination variable was created (Yes = 1, No = 0) to indicate any self-reported physical abuse, sexual abuse, psychological abuse, or neglect.

Contamination: retrospective. The SPSPA was again administered at the Age 18 LONGSCAN assessment to measure any lifetime instances of retrospectively self-reported exposure to physical, sexual, or psychological abuse. Using the same procedures described above, a retrospective contamination variable was created (Yes = 1, No = 0) to indicate any self-reported lifetime exposure to physical abuse, sexual abuse, or psychological abuse prior to age 18. AMP items were not administered at the Age 18 LONGSCAN assessment and therefore child neglect is not included in the retrospective contamination variable.

Behavior problems. The Child Behavior Checklist (CBCL; Achenbach, 1991) is a well-established, caregiver report instrument assessing the frequency of child internalizing behavior problems, such as worry and depressed mood, as well as externalizing behavior problems, such as noncompliance and delinquency, in the past 6 months. The CBCL was administered every 2 years from the Age 4 through Age 16 LONGSCAN assessments. Raw scores obtained from these seven measurement occasions were used in subsequent statistical modeling.

Analytic approach

The causal effects of child maltreatment on behavior problems were estimated through a series of partially pooled synthetic control methods (SCMs) for staggered adoption (Ben-Michael, Feller, & Rothstein, 2022). The SCM can be viewed as an extension of difference-in-difference estimators where, instead of using a simple average of the outcomes from nonmaltreated children to form a comparison group, a weighted average is instead used (Abadie, 2021; Abadie et al., 2010), whereby the weights are estimated so that the premaltreatment outcomes of the maltreated children and controls match as closely as possible (i.e., differences in premaltreatment outcomes between maltreated children and comparison children are zero or near zero). These weights typically range between 0 (not contributing to the comparison group) and 1.0 (comparison group is only that unit), though this staggered adoption approach uses an augmented SCM estimator that allows for negative weights to improve premaltreatment balancing. When balance is achieved, this weighted average is a synthetic nonmaltreated child that can approximate the potential outcomes of the maltreated child had they never been exposed to maltreatment.

The partially pooled SCM was developed to handle instances of staggered adoption of some intervention (Ben-Michael et al., 2022), which is appropriate for studying child maltreatment that first occurs at different ages across children. The partial pooling represents a middle ground between estimating separate SCMs for each maltreated child (i.e., estimate weights separately for each maltreated child and then average) and a fully pooled approach (i.e., average outcomes for all maltreated children before calculating weights). This balance is achieved through a hyperparameter, ν, where values closer to 0 weigh more heavily toward the separate SCM models, and ν parameters closer to 1 weigh more heavily toward the fully pooled model. The value of the hyperparameter is either chosen by the researcher prior to estimation or estimated as the value that minimizes the squared difference of the model fits of the no-pooled and fully pooled models. Given that no theory points to an ex ante correct value for ν, we choose to estimate ν.

If the balance is achieved on the premaltreatment outcomes, the differences between the partially pooled synthetic control and the average of the maltreated children after maltreatment is the average treatment of the treated (ATT) causal estimate (Rubin, 1977). Balance on the premaltreatment outcomes can be evaluated using L2 norm statistics for the Global model and the Individual model (Ben-Michael et al., 2022). These L2 norm statistics are on the same scale as the outcome and are general measures of how well the weights used to balance achieve a mean difference of 0 between the synthetic controls and the maltreated children on the premaltreatment outcomes. Values close to 0 are evidence that the synthetic controls are reflecting what the maltreated children's outcomes would have been had they not been exposed to maltreatment. Statistical significance of the overall ATT and the ATT at each wave after the first maltreatment exposure is determined through 95% confidence intervals constructed via the wild bootstrap which accounts for serial correlation within units over time and is asymptotically correct if there are insufficient units for traditional clustered standard errors (see Ben-Michael et al., 2022; Mammen, 1993 for further details). Here, three partially pooled SCMs were estimated for each outcome (six total models) to highlight the impact of contamination on causal effect estimates. The impact of contamination on causal effect estimates was characterized by comparing the pattern of statistically significant results overall and by each wave after the first maltreatment exposure and the percent change in ATT estimates. Estimation was conducted using the AUGSYNTH package (Ben-Michael et al., 2022) in the R software language (R Core Team, 2018).

Multiple imputation of missing data. The overall percentage of missing data in the CBCL Internalizing and Externalizing behavior variables was 22.7%. The amount missing increased with the wave of data collection, ranging from 9.9% of children with missing data at the Age 4 assessment to 36% of children with missing data at the Age 16 assessment. In total, 63.4% (858/1354) of records were incomplete. The rate of missing cases was similar between those never maltreated and those ever maltreated (53% each). Multiple imputation of missing cases was implemented using the MICE package (van Buuren & Groothuis-Oudshoorn, 2011) in the R software language (R Core Team, 2018). Missing data on the CBCL Internalizing and Externalizing behaviors were imputed using the following indicators that were related to the probability of missingness and from theory (Ji, Chow, Schermerhorn, Jacobson, & Cummings, 2018): by the most recent previous assessment of the CBCL items, race and ethnicity, early life family income, age in years, gender, and the site where the study was implemented. Because data had a hierarchical structure (time nested within persons), the two-level predictive mean matching was used for multiple imputation of continuous variables. We generated five multiply imputed data sets for missing data on the CBCL Internalizing and Externalizing behavior subscales separately for the maltreated (always and eventually maltreated) and the nonmaltreated data (never and premaltreated) to prevent the exposure variable from artificially inflating associations with the CBCL items. Twenty-five data sets were then created by combining the 10 multiply imputed data sets (5 pre/never maltreated and 5 post/always maltreated) into all possible combinations of nonmaltreated and maltreated data. Rubin's rules were used to pool the parameter estimates obtained by fitting models to each of the 25 data sets independently. The models were estimated again on the complete case data (i.e., nonimputed data sets) for comparison. Results were comparable but the multiply imputed data yielded small confidence intervals, which was reasonable given the reduction in sample size in the complete case data.

Results

Confirmed maltreatment and prevalence of contamination

Each child's first exposure to confirmed child maltreatment occurred between the Age 10 and Age 16 waves of data collection with the age at the first exposure varying across children (see Figure 1 for a graphical depiction of the timing of first exposure to confirmed child maltreatment in the current study). With respect to the chronicity of confirmed maltreatment for those in the maltreatment group, $n = 51$ children were exposed to maltreatment at only one wave, $n = 11$ children were exposed to maltreatment at two waves, and $n = 1$ child was exposed to maltreatment at three waves of data collection.

Because SCM requires repeated premaltreatment measurement of outcomes, we restricted our sample to include individuals with at least three premaltreatment time points on child behavior problems; this data restriction yielded 571 children in the nonmaltreated control condition who never had a confirmed case of child maltreatment. Prospective self-reports of child maltreatment identified 392 (67%) children originally placed in the nonmaltreated comparison group who self-reported exposure to child maltreatment. When removing contaminated cases identified through prospective self-reports, the sample size in the comparison group is reduced to 179 children. Retrospective self-reports of child maltreatment identified 354 (62%) children in the nonmaltreated comparison group who self-reported exposure to child maltreatment. When removing contaminated cases identified through retrospective self-reports, the sample size in the comparison group is reduced to 217 children who have no confirmed cases of child maltreatment or retrospective self-reports of child maltreatment. Therefore, the prevalence rate of contamination was 67% from the prospective reports and 62% from the retrospective reports. Full sample characteristics are displayed in Table 1.

Internalizing behaviors

Full results for all three models on internalizing behaviors can be found in Table 2 and graphically displayed in Figure 2. Across all three models, L2 norm statistics indicated that balance on the pre-maltreatment outcome measures was achieved, with values ranging from 0 to 0.33 (i.e., near 0 difference between synthetic controls and maltreatment after balancing).

Model 1: Baseline. When contaminated cases are not removed and are left in the nonmaltreated comparison group, neither the overall causal effect ($ATT_{overall} = 1.86$, CI: [−0.84, 4.56], $v = 0.66$) nor the causal effects at any time point postmaltreatment were statistically significant (all CIs cover 0). This means that the causal effect of child maltreatment that occurred after age 8 on internalizing behaviors was not significantly different from 0 when ignoring contamination (i.e., those in the nonmaltreated comparison group who self-reported exposure to maltreatment).

Model 2: Contamination controlled via prospective self-reports. When contamination was identified and removed based on prospective self-reports, the overall causal effect of child maltreatment on internalizing behaviors was 39% higher than the estimate from the baseline model and became statistically significant ($ATT_{overall} = 2.58$, CI: [0.17 m 4.99], $v = 0.44$). This overall effect was driven by significant causal effects at 2 years ($ATT_{t+2} = 3.46$, CI: [0.56, 6.36], 20% higher ATT than baseline model) and at 4 years ($ATT_{t+4} = 3.49$, CI: [0.37, 6.61]. 37% higher ATT than the baseline model) following the first exposure to confirmed child maltreatment. This means that the causal effect of child maltreatment that occurred after age 8 on internalizing behaviors was detectable after controlling contamination by removing children from the nonmaltreated

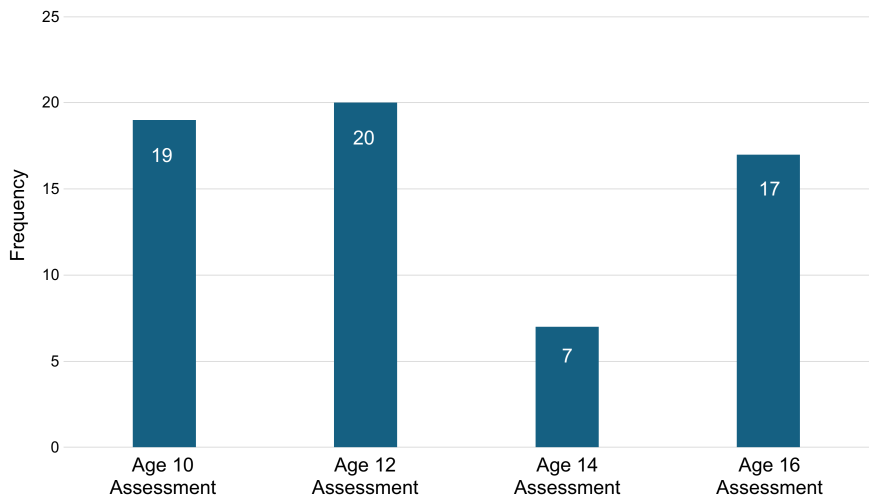

Figure 1 Developmental timing of first exposure to confirmed child maltreatment

Table 1 Demographics

	Maltreated group	Nonmaltreated Comparison Groups		
		No contamination removed	Contamination removed by prospective self-reports	Contamination removed by retrospective self-reports
	n = 63	*n* = 571	*n* = 179	*n* = 217
Sex (%)				
Male	51	51	45	49
Female	49	49	55	51
Race/Ethnicity (%)				
Black	60	62	61	71
Hispanic	3	6	6	5
White	25	22	22	19
Other	11	10	11	6
Data collection site (%)				
SW	25	31	34	42
NW	11	23	31	17
MW	17	10	8	10
SO	32	25	22	24
EA	16	11	6	7
Median family income	$10,000–$14,000	$10,000–$14,000	$10,000–$14,000	$5,000–$9,999

EA, East Site; MW, Midwest Site; NW, Northwest Site; SO, South Site; SW, Southeast Site.

Table 2 Internalizing behaviors—partially pooled synthetic control method by type of control of contamination

	Years since maltreatment	ATT	*SE*	Lower CI	Upper CI	
Model 1: Baseline model	Overall average	1.86	1.38	−0.84	4.56	–
	0	1.03	1.23	−1.38	3.44	–
	2	2.89	1.73	−0.50	6.28	–
	4	2.55	1.97	−1.31	6.41	–
	6	1.77	1.85	−1.86	5.40	–
	Years since maltreatment	ATT	*SE*	Lower CI	Upper CI	ATT percent increase
Model 2: Contamination controlled via prospective self-report control	Overall average	**2.58**	1.23	0.17	4.99	38.71
	0	1.56	1.31	−1.01	4.13	51.46
	2	**3.46**	1.48	0.56	6.36	19.72
	4	**3.49**	1.59	0.37	6.61	36.86
	6	2.11	1.70	−1.22	5.44	19.21
Model 3: Contamination controlled via retrospective self-report control	Overall average	**2.67**	1.14	0.44	4.90	43.55
	0	1.59	1.18	−0.72	3.90	54.37
	2	**4.40**	1.58	1.30	7.50	52.25
	4	**3.81**	1.64	0.60	7.02	49.41
	6	2.24	1.76	−1.21	5.69	26.55

Bold indicates statistically significant associations (i.e., 95% confidence intervals do not contain 0). Percent increase is the percentage increase in the causal estimand in comparison with the no control of contamination model results, and Nu is the optimal balance between the individual and the mean model, where scores closer to 0 favor the individual model and scores closer to 1 favor the mean model. ATT, Average Treatment of the Treated Causal Estimand; lower CI, lower 95% confidence interval bound; SE, standard error of the effect; upper CI, upper 95% confidence interval bound.

comparison group who prospectively self-reported exposure to maltreatment.

Model 3: Contamination controlled via retrospective self-reports. When contaminated cases were identified and removed based on retrospective reports and those cases were removed from the comparison group, the overall causal effect of child maltreatment on internalizing behaviors was 44% higher than the baseline model and became statistically significant ($ATT_{overall}$ = 2.67, CI: [0.44, 4.90], ν = 0.35). This overall effect was also driven by significant causal effects at 2 years (ATT_{t+2} = 4.40, CI: [1.30, 7.50], 52% higher than baseline model) and at 4 years (ATT_{t+4} = 3.81, CI: [0.37, 6.61], 49% higher than baseline model). This means that the causal effect of child maltreatment on internalizing behaviors was detectable after controlling contamination by removing children from the nonmaltreated comparison group who retrospectively reported exposure to maltreatment.

Externalizing behaviors

Full results for all three models on externalizing behaviors can be found in Table 3 and graphically

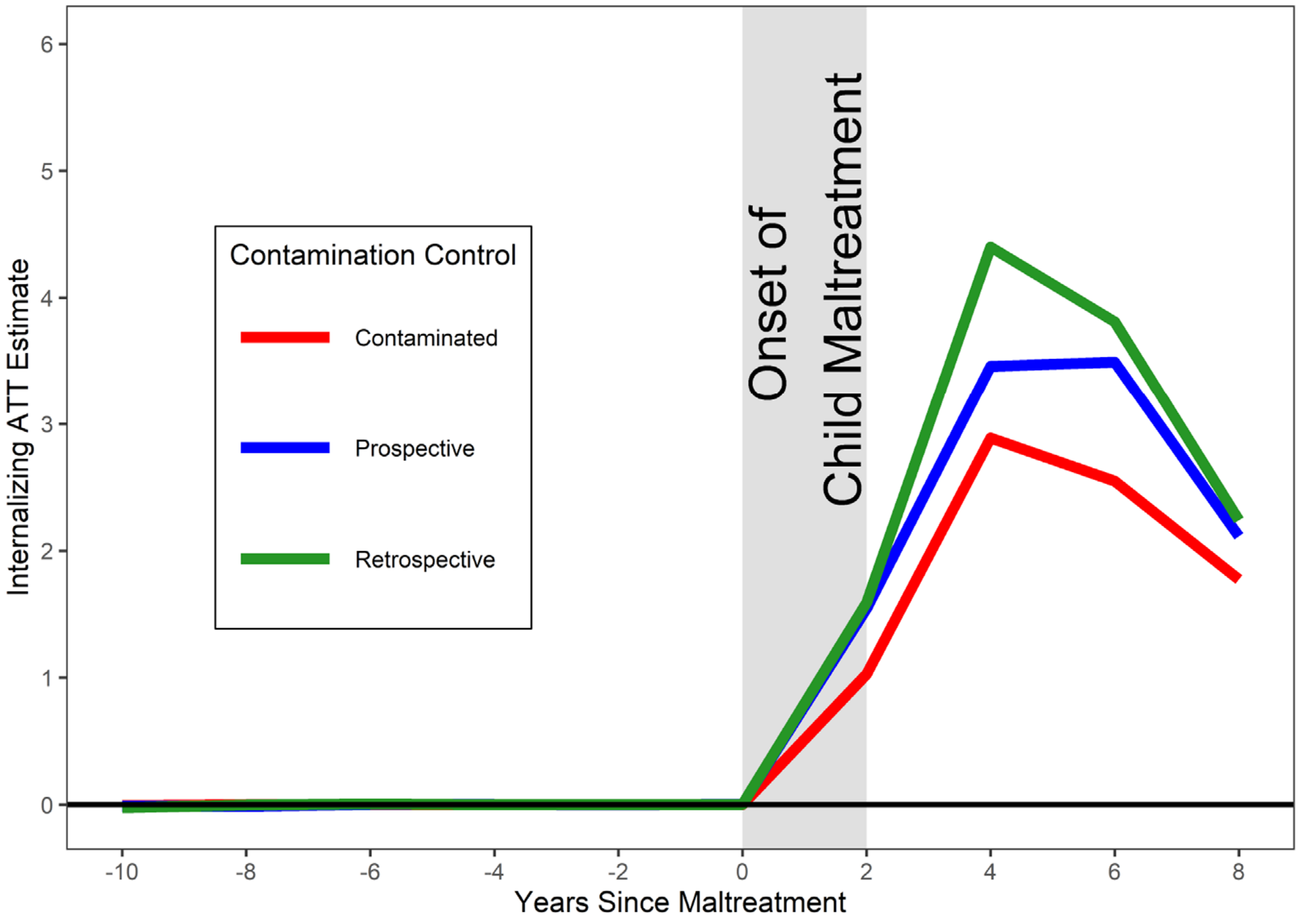

Figure 2 Internalizing behaviors average treatment of the treated causal estimand estimates by the method of contamination control

Table 3 Externalizing behaviors—partially pooled synthetic control method by type of control of contamination

	Years since maltreatment	ATT	*SE*	Lower CI	Upper CI	
Model 4: Baseline model	Overall average	3.14	1.86	−0.51	6.79	–
	0	1.69	1.66	−1.56	4.94	–
	2	**4.86**	2.37	0.21	9.51	–
	4	2.97	2.15	−1.24	7.18	–
	6	3.44	3.24	−2.91	9.79	–
	Years since maltreatment	ATT	SE	Lower CI	Upper CI	ATT percent increase
Model 5: Contamination controlled via prospective self-report	Overall average	**3.39**	1.71	0.05	6.74	8.06
	0	1.85	1.64	−1.36	5.06	9.47
	2	**4.85**	1.85	1.22	8.48	−0.21
	4	3.34	2.05	−0.68	7.36	12.46
	6	3.63	3.17	−2.58	9.84	5.52
Model 6: Contamination controlled via retrospective self-report	Overall average	3.35	1.78	−0.14	6.84	6.69
	0	2.00	1.71	−1.35	5.35	18.34
	2	**5.11**	1.97	1.25	8.97	5.14
	4	3.40	2.21	−0.94	7.73	14.31
	6	4.05	3.16	−2.14	10.24	17.73

Bold indicates statistically significant associations (i.e., 95% confidence intervals do not contain 0). Percent increase is the percentage increase in the causal estimand in comparison to the no control of contamination model results, and Nu is the optimal balance between the individual and the mean model, where scores closer to 0 favor the individual model and scores closer to 1 favor the mean model. ATT, Average Treatment of the Treated Causal Estimand; lower CI, lower 95% confidence interval bound; *SE*, standard error of the effect; upper CI, upper 95% confidence interval bound.

displayed in Figure 3. Across all three models, L2 norm statistics indicated that balance on the pre-maltreatment outcome measures was achieved, with values ranging from 0 to 0.10 (i.e., near 0 difference between synthetic controls and maltreatment after balancing).

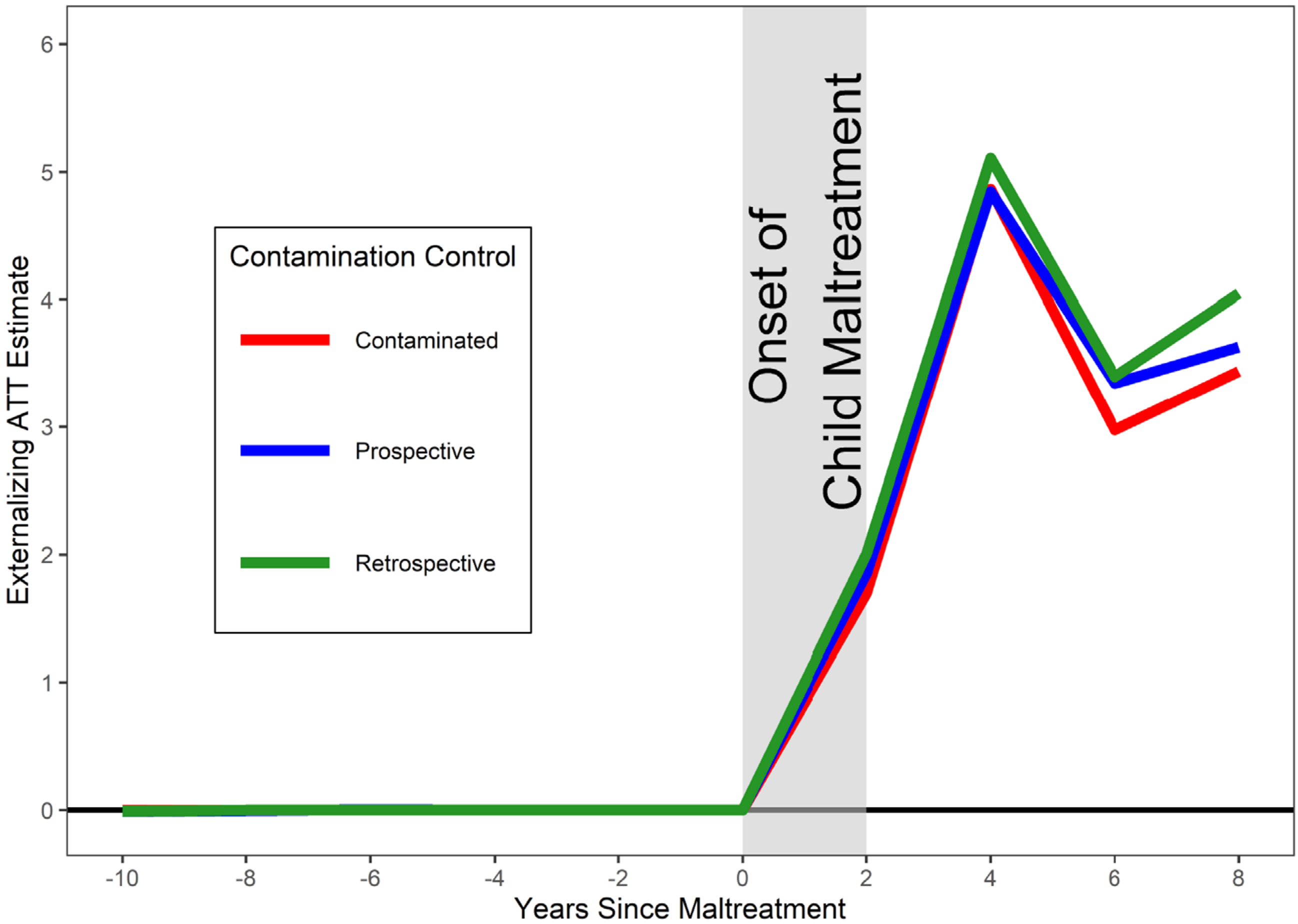

Figure 3 Externalizing behaviors ATT estimates by the method of contamination control

Model 4: Baseline. When contaminated cases were not removed and left in the nonmaltreated comparison group, the overall causal effect of child maltreatment on externalizing behaviors was not statistically significant ($ATT_{overall}$ = 3.14, CI: [−0.51, 6.79], v = 0.59). However, there was a significant causal effect of child maltreatment on externalizing behaviors at 2 years after the confirmed maltreatment was first recorded (ATT_{t+2} = 4.86, CI: [0.21, 9.51]). This means that the causal effect of child maltreatment that first occurred after age 8 on externalizing behaviors was only statistically significant 2 years after confirmed child maltreatment when ignoring contamination.

Model 5: Contamination controlled via prospective self-reports. When contamination was identified and removed based on ages 12 and 16 prospective self-reports, the overall causal effect of child maltreatment on externalizing behaviors was 8.1% higher than the baseline model and became statistically significant ($ATT_{overall}$ = 3.39. CI: [0.05, 6.74], v = 0.550). This means that the overall causal effect of child maltreatment occurring after age 8 on externalizing behaviors became significant only after controlling contamination by removing children from the nonmaltreated comparison group who prospectively reported exposure to maltreatment. Like in the baseline model, there was a significant causal effect on externalizing behaviors 2 years after the first exposure to confirmed maltreatment (ATT_{t+2} = 4.85, CI: [1.22, 8.48]), although the increase in effect magnitude was negligible. The percent increase in ATTs across all waves relative to the baseline model ranged from 0% to 12.5%.

Model 6: Contamination controlled via retrospective self-reports. When contamination was identified and removed based on Age 18 retrospective self-reports, the overall causal effect of child maltreatment on externalizing behaviors was 6.7% higher than baseline model but was still not statistically significant ($ATT_{overall}$ = 3.35, CI: [−0.14, 6.84], v = 0.42). However, as in the baseline model, the causal effect of child maltreatment on externalizing behaviors at 2 years after the confirmed maltreatment was first recorded was statistically significant (ATT_{t+2} = 5.11, CI: [1.25, 8.97], 9% higher than baseline model). This means that the causal effect of child maltreatment that first occurred after age 8 on externalizing behaviors was only statistically significant 2 years after confirmed child maltreatment when contamination was controlled by removing children from the nonmaltreated comparison group who retrospectively reported exposure to maltreatment.

Discussion

The current study demonstrated the prevalence and impact of unaddressed contamination on the causal effects of child maltreatment on behavior problems. We leveraged the longitudinal design and multiple measurement strategies for child maltreatment status in the LONGSCAN cohort to quantify differences in the magnitude and statistical significance of causal effects across different data scenarios. The repeated assessments of behavior problems and child maltreatment status of the LONGSCAN cohort allowed us to use SCMs for causal estimates, an innovative approach to generate higher quality counterfactual controls that are balanced on premaltreatment outcomes (Abadie et al., 2010; Ben-Michael et al., 2022). Given the strengths of the study design and the use of an innovative causal inference method, results from this study have important implications for future research and the need to address contamination to generate accurate causal effect estimates.

The prevalence of contamination in the LONGSCAN cohort was high and impacted causal effect estimates for child behavior problems. In line with previous estimates (Shenk et al., 2022), contamination rates from this study ranged between 62% and 67% depending on whether prospective or retrospective self-reports were used to identify maltreatment. The impact contamination had on causal effect estimates was most pronounced for internalizing behavior problems. In our study, the overall causal effect of child maltreatment on internalizing behavior problems was not statistically significant until contamination was addressed, with effect size magnitudes being as much as 54% higher after removing contaminated cases. Researchers who fail to quantify and address contamination might conclude that there were no causal effects of child maltreatment on internalizing behaviors or underestimate the magnitude of effect sizes. That is, if unaddressed, contamination can undermine the replication of robust associations in the extant literature (Gilbert et al., 2009). These findings underscore the importance of using a longitudinal dual-measurement strategy to monitor child maltreatment status and identify contamination throughout the duration of the study.

Contamination had a smaller impact on the significance and magnitude of causal effect estimates of child maltreatment on externalizing behaviors. Correcting contamination via prospective reports did yield a significant overall effect of child maltreatment on externalizing behavior problems that were not observed in the model ignoring contamination, thereby promoting replication. However, correcting contamination produced effect size estimates that were only 0% to 18% higher, a stark contrast from the magnitude of difference in effect size with internalizing behaviors. Thus, correcting contamination in SCMs of child maltreatment effects on externalizing behaviors may improve the significance and accuracy of causal estimates but with smaller effects relative to internalizing behaviors.

The results from this study may be in part due to subsetting our data to those who had at least three premaltreatment assessments and where maltreatment and control conditions were matched on the levels and trends of the outcomes during these three assessments to facilitate the use of the SCM. As such, the earliest first exposure to child maltreatment in our analytic sample occurred at approximately age 10. The current results are consistent with a growing body of research suggesting that first exposure to child maltreatment in adolescence remains statistically significant with internalizing behaviors but where the effect may be weaker for externalizing behaviors (Bauer, Hammerton, Fraser, Fairchild, & Halligan, 2021; Dunn, Nishimi, Gomez, Powers, & Bradley, 2018; Keiley, Howe, Dodge, Bates, & Pettit, 2001; Russotti et al., 2021). It is also important to note that results from SCMs are best characterized as “time since” first exposure to child maltreatment. Our results therefore reflect the time course of child maltreatment effects 6 years following exposure to maltreatment, a pattern that may yield different results than “age-based” models that do not account for the timing of maltreatment exposure or equate maltreatment and comparison groups on the levels and trends of the outcome before exposure.

Findings from this study need to be interpreted with certain limitations. First, our analytic sample was restricted to those who first experienced maltreatment no earlier than the Age 10 LONGSCAN assessment to ensure at least three premaltreatment time points for modeling purposes. As such, our findings are specifically about child maltreatment effects for those who first experienced maltreatment at entry to adolescence. Future work using SCMs to establish causal associations on children who first experience maltreatment at a younger age will require researchers to collect data at earlier ages to provide premaltreatment assessment points. This is admittedly a difficult feat as children are most vulnerable to child maltreatment in early childhood (U.S. Department of Health and Human Services, 2020). Thus, a key limitation of panel data approaches such as SCM is that these models cannot be applied to settings when maltreatment occurs before it is possible to obtain premaltreatment outcome data. Notably, contamination rates derived from the retrospective self-report of child maltreatment status may also underestimate the amount of contamination present in the control condition because this measure excluded neglect. Future work should include retrospective reports of neglect to better identify possible contamination. Additionally, the SCM modeling approach is not yet capable of modeling multiple exposures to the same treatment. As such, we are unable to account for any influence that chronicity of confirmed maltreatment, which occurred in 12 of the 63 members of the maltreatment group, may have had on the outcomes. Finally, we operationalized contamination as the presence of self-reported child maltreatment in a control group of children without a confirmed case of maltreatment. Some researchers may choose to use self-report methods to create child maltreatment and control groups. However, contamination can still occur in this type of classification if children do not self-report maltreatment but have an alleged, confirmed, or substantiated history of maltreatment. Despite the limitations, our findings, in line with previous work on contamination in child maltreatment research, provide strong evidence that contamination is a feature of child maltreatment research and needs to be measured and addressed to facilitate replication of findings.

Conclusions

Contamination in child maltreatment research biases the significance and magnitude of statistical effect size estimates (Scott et al., 2010; Shenk et al., 2016, 2022). This study used an innovative causal inference approach to show that contamination can hide significant causal effects even in well-established consequences of child maltreatment. Whether generating statistical or causal associations, child maltreatment researchers should consider whether and how contamination may be present in their comparison conditions. Our findings reveal that a dual-measurement strategy using both confirmed cases of child maltreatment and self-report methods is sufficient to detect large amounts of contamination and that ignoring contamination can downwardly bias the significance and magnitude of causal effect estimates. Removing children from the nonmaltreated control groups who self-report child maltreatment, or modeling them as their own group, should be considered when correcting contamination and estimating maltreatment effects.

Acknowledgments

The authors would like to acknowledge the funding agencies that supported work on this manuscript: National Science Foundation, Division of Behavioral and Cognitive Science (BCS-2041333; Shenk, PI), and the National Institute of Child Health and Human Development (R03HD104739; Shenk, PI). The authors have declared that they no competing or potential conflicts of interest.

Data availability

Data used in this analysis come from the publicly available LONGSCAN data set (https://www.ndacan.acf.hhs.gov/datasets/dataset-details.cfm?ID=170).

Correspondence

John M. Felt, Center for Healthy Aging at Penn State, 431 Biobehavioral Health Building, University Park, PA 16802, USA; Email: jzf434@psu.edu

Key points

- Contamination, which happens when children who have been maltreated are erroneously assigned to or retained in a nonmaltreated comparison condition, can downwardly bias the significance, magnitude, or direction of associations for child maltreatment outcomes.
- Using a novel causal inference method, the synthetic control method for staggered adoption, we estimated causal effects between child maltreatment and behavior problems before and after controlling contamination.
- Failing to identify and address contamination can make it difficult to identify significant causal effects, even for well-established outcomes.
- Removing contamination from nonmaltreated comparison conditions increased the significance and magnitude of causal effect estimates.
- Child maltreatment researchers should use a dual-measurement strategy to identify contamination and obtain more accurate statistical and causal effect estimates.

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, *59*, 391–425.

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, *105*, 493–505.

Achenbach, T. (1991). *Integrative guide for the 1991 CBCL/4-18, YSR, and TRF profiles*. Burlington, VT: University of Vermont.

Barnett, D., Manly, J., & Cicchetti, D. (1993). Defining child maltreatment: The interface between policy and research. *Child Abuse, Child Development, and Social Policy*, *8*, 7–73.

Bauer, A., Hammerton, G., Fraser, A., Fairchild, G., & Halligan, S.L. (2021). Associations between developmental

timing of child abuse and conduct problem trajectories in a UK birth cohort. *BMC Psychiatry*, *21*, 1–13.
Ben-Michael, E., Feller, A., & Rothstein, J. (2022). Synthetic controls with staggered adoption. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *84*, 351–381.
Craven, R.G., Marsh, H.W., Debus, R.L., & Jayasinghe, U. (2001). Diffusion effects: Control group contamination threats to the validity of teacher-administered interventions. *Journal of Educational Psychology*, *93*, 639–645.
Cuzick, J., Edwards, R., & Segnan, N. (1997). Adjusting for non-compliance and contamination in randomized clinical trials. *Statistics in Medicine*, *16*, 1017–1029.
Delgado-Rodríguez, M., & Llorca, J. (2004). Bias. *Journal of Epidemiology & Community Health*, *58*, 635–641.
Dubowitz, H., Villodas, M.T., Litrownik, A.J., Pitts, S.C., Hussey, J.M., Thompson, R., ... Runyan, D. (2011). Psychometric properties of a youth self-report measure of neglectful behavior by parents. *Child Abuse & Neglect*, *35*, 414–424.
Dunn, E.C., Nishimi, K., Gomez, S.H., Powers, A., & Bradley, B. (2018). Developmental timing of trauma exposure and emotion dysregulation in adulthood: Are there sensitive periods when trauma is most harmful? *Journal of Affective Disorders*, *227*, 869–877.
English, D.J., Bangdiwala, S.I., & Runyan, D.K. (2005). The dimensions of maltreatment: Introduction. *Child Abuse & Neglect*, *29*, 441–460.
Everson, M.D., Smith, J.B., Hussey, J.M., English, D., Litrownik, A.J., Dubowitz, H., ... Runyan, D.K. (2008). Concordance between adolescent reports of childhood abuse and child protective service determinations in an at-risk sample of young adolescents. *Child Maltreatment*, *13*, 14–26.
Finkelhor, D., Turner, H.A., Shattuck, A., & Hamby, S.L. (2015). Prevalence of childhood exposure to violence, crime, and abuse: Results from the national survey of children's exposure to violence. *JAMA Pediatrics*, *169*, 746–754.
Gilbert, R., Widom, C.S., Browne, K., Fergusson, D., Webb, E., & Janson, S. (2009). Burden and consequences of child maltreatment in high-income countries. *The Lancet*, *373*, 68–81.
Hart, S.N., Brassard, M.R., & Karlson, H.C. (1996). Psychological maltreatment. In J. Brier, L. Berliner, J.A. Bulkley, C. Jenny, & T. Reid (Eds.), *The APSAC handbook on child maltreatment* (pp. 72–89). Thousand Oaks: Sage Publications.
Hirano, K., Imbens, G.W., Rubin, D.B., & Zhou, X.-H. (2000). Assessing the effect of an influenza vaccine in an encouragement design. *Biostatistics*, *1*, 69–88.
Ji, L., Chow, S.M., Schermerhorn, A.C., Jacobson, N.C., & Cummings, E.M. (2018). Handling missing data in the modeling of intensive longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*, 715–736.
Jo, B. (2002). Statistical power in randomized intervention studies with noncompliance. *Psychological Methods*, *7*, 178–193.
Keiley, M.K., Howe, T.R., Dodge, K.A., Bates, J.E., & Pettit, G.S. (2001). The timing of child physical maltreatment: A cross-domain growth analysis of impact on adolescent externalizing and internalizing problems. *Development and Psychopathology*, *13*, 891–912.
Knight, E.D., Runyan, D.K., Dubowitz, H., Brandford, C., Kotch, J., Litrownik, A., & Hunter, W. (2000). Methodological and ethical challenges associated with child self-report of maltreatment: Solutions implemented by the LONGSCAN consortium. *Journal of Interpersonal Violence*, *15*, 760–775.
LONGSCAN Investigators. (1998). *About my parents*. University of North Carolina. Available from: http://www.unc.edu/depts/sph/longscan/
Magill, N., Knight, R., McCrone, P., Ismail, K., & Landau, S. (2019). A scoping review of the problems and solutions associated with contamination in trials of complex interventions in mental health. *BMC Medical Research Methodology*, *19*, 4.
Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *The Annals of Statistics*, *21*, 255–285.
R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statsitical Computing. Available from: https://www.r-project.org/
Roobol, M.J., Kerkhof, M., Schröder, F.H., Cuzick, J., Sasieni, P., Hakama, M., ... & Kwiatkowski, M. (2009). Prostate cancer mortality reduction by prostate-specific antigen–based screening adjusted for nonattendance and contamination in the European Randomised Study of Screening for Prostate Cancer (ERSPC). *European Urology*, *56*, 584–591.
Rubin, D.B. (1977). Assignment to treatment group on the basis of a covariate. *Journal of Educational Statistics*, *2*, 1–26.
Runyan, D.K., Curtis, P.A., Hunter, W.M., Black, M.M., Kotch, J.B., Bangdiwala, S., ... & Landsverk, J. (1998). Longscan: A consortium for longitudinal studies of maltreatment and the life course of children. *Aggression and Violent Behavior*, *3*, 275–285.
Russotti, J., Warmingham, J.M., Duprey, E.B., Handley, E.D., Manly, J.T., Rogosch, F.A., & Cicchetti, D. (2021). Child maltreatment and the development of psychopathology: The role of developmental timing and chronicity. *Child Abuse & Neglect*, *120*, 105215.
Scott, K.M., Smith, D.R., & Ellis, P.M. (2010). Prospectively ascertained child maltreatment and its association with DSM-IV mental disorders in young adults. *Archives of General Psychiatry*, *67*, 712–719.
Shenk, C.E., Noll, J.G., Peugh, J.L., Griffin, A.M., & Bensman, H.E. (2016). Contamination in the prospective study of child maltreatment and female adolescent health. *Journal of Pediatric Psychology*, *41*, 37–45.
Shenk, C.E., Rausch, J.R., Shores, K.A., Allen, E.K., & Olson, A.E. (2022). Controlling contamination in child maltreatment research: Impact on effect size estimates for child behavior problems measured throughout childhood and adolescence. *Development and Psychopathology*, *1–13*, 1287–1299.
Shenk, C.E., Shores, K.A., Ram, N., Felt, J.M., Chimed-Ochir, U., Olson, A.E., & Fisher, Z.F. (2023). Contamination in observational research on child maltreatment: A conceptual and empirical review with implications for future research. *Child Maltreatment*. https://doi.org/10.1177/10775595231224472
Simmons, N., Donnell, D., Ou, S., Celentano, D.D., Aramrattana, A., Davis-Vogel, A., ... & Latkin, C. (2015). Assessment of contamination and misclassification biases in a randomized controlled trial of a social network peer education intervention to reduce HIV risk behaviors among drug users and risk partners in Philadelphia, PA and Chiang Mai, Thailand. *AIDS and Behavior*, *19*, 1818–1827.
U.S. Department of Health and Human Services, Administration for Children and Families, & Administration on Children, Youth, and Families, Children's Bureau. (2020). *Child maltreatment*. Available from: https://www.acf.hhs.gov/cb/report/child-maltreatment-2020
van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*, 1–67.

Accepted for publication: 26 February 2024