nature mental health

Article

https://doi.org/10.1038/s44220-024-00345-8

DNA methylation mediates the link between adversity and depressive symptoms

Received: 14 September 2023

Accepted: 3 October 2024

Published online: 02 December 2024

Check for updates

Alexandre A. Lussier[1,2,3,17] ✉, Brooke J. Smith[1,17], Jonah Fisher[4], Mannan Luo[5,6,7], Janine Cerutti[1], Lisa Schneper[8], Trey Smith[4], Charlotte A. M. Cecil[5,7], Janine F. Felix[5,9], Colter Mitchell[4], Daniel A. Notterman[8], Kerry J. Ressler[2,10], Daniel J. Schaid[11], Andrew J. Simpkin[12], Matthew J. Suderman[13], Esther Walton[14], Andrew D. A. C. Smith[15] & Erin C. Dunn[1,2,3,16] ✉

Experiences of childhood adversity can double the risk for depression. Although the mechanisms underlying this relationship remain unclear, DNA methylation (DNAm) has emerged as a potential pathway to explain the link between adversity and depression. We thus investigated whether epigenome-wide DNAm statistically mediates the association between childhood adversity and adolescent depressive symptoms. Specifically, we performed epigenome-wide mediation analyses to investigate the role of blood-based DNAm (age 7 years) in linking seven types of adversity (ages 0–7 years) to depressive symptoms (age 10.6 years). Primary analyses were conducted in the Avon Longitudinal Study of Parents and Children and replicated in the Future of Families and Child Wellbeing Study and Generation R Study. We identified 70 cytosine–guanine dinucleotides (CpGs) that mediated 10–73% of the correlation between adversity and depressive symptoms, with DNAm differences at 39 of these CpGs showing protective effects. Our findings suggest DNAm reflects a biological pathway linking childhood adversity to depression and a potential mechanism towards resilience.

Depression is a serious global health problem that is estimated to affect over 264 million people worldwide[1]. Decades of research have shown that exposure to childhood adversity is one of the strongest and most consistent contributors to risk for depression and other psychiatric disorders[2]. In recent meta-analyses and systematic reviews, nearly 50% of people with depression report having experienced childhood maltreatment[3], with childhood adversity explaining up to 40% of premature morbidity and mortality[4]. However, the biological pathways through which adversity influences disease risk are poorly understood.

One possibility is that childhood adversity shapes mental health risk through epigenetic changes, such as DNA methylation (DNAm). DNAm is an epigenetic process wherein methyl groups are added to DNA, generally at cytosine–guanine dinucleotides (CpG), and can influence gene expression without changing DNA sequences[5]. Systematic reviews from over 100 human and animal studies show that various types of childhood adversity are linked to DNAm changes, contributing as much as a 5–17% difference in DNAm levels at a given CpG between those exposed and those unexposed to adversity[6,7]. Recent work also suggests that there may be sensitive periods between three and five years of age[8,9] when adversity has a stronger impact on DNAm.

DNAm patterns have also been linked, primarily in adults, to future risk for depressive disorders and symptoms[10–13]. Baseline DNAm profiles of individuals diagnosed with major depressive disorder (MDD) predicted their MDD status six years later, suggesting that DNAm may serve as a biomarker for treatment response or other symptom changes[14]. Even more intriguingly, recent experimental studies in humans have shown that psychiatric interventions might reverse DNAm signatures previously associated with internalizing psychopathology[15–18].

Studies using statistical mediation analyses are poised to assess whether DNAm is a biological bridge linking adversity exposure to

A full list of affiliations appears at the end of the paper. ✉e-mail: alussier@mgh.harvard.edu; erindunn@purdue.edu

Table 1 | Results of the SLCMA

Adversity	Sample size	Timepoint selected	Sensitive period	Total effect	*P* value	CI
Caregiver physical or emotional abuse	675	8 months	Very early childhood (0–3 years)	−0.81	0.70	(−1.97, 0.67)
Sexual or physical abuse (by anyone)	663	18 months	Early childhood (3–5 years)	1.81	0.06	(−0.03, 3.15)
Maternal psychopathology	653	33 months	Very early childhood (0–3 years)	1.12	0.02	(0.19, 1.85)
One-adult in the household	667	47 months	Early childhood (3–5 years)	0.91	0.35	(−0.43, 1.97)
Family instability	656	57 months	Early childhood (3–5 years)	1.21	0.37	(−0.47, 2.45)
Financial hardship	627	61 months	Early childhood (3–5 years)	1.24	0.09	(−0.10, 2.25)
Neighborhood disadvantage	660	21 months	Very early childhood (0–3 years)	1.04	0.11	(−0.13, 1.97)

Results from the SLCMA, which identified the theoretical life course model with the best explanatory power between each adversity and the depressive symptoms as measured by the Short Mood and Feelings Questionnaire (SMFQ) score at age 10.6 years. For all adversities, a sensitive period life course hypothesis was selected. The timepoint selected is shown here, along with its corresponding life course model, the effect estimate (total effect) for the sensitive period exposure on SMFQ score, *P* value and 95% confidence interval (CI). Post-selective inference was performed using a two-sided max-|*t*|-test.

depression. As reported in a 2021 scoping review[19] exploring DNAm as a mediator of the relationship between environmental exposures and chronic disease, only one study had investigated the relationship between negative childhood experiences and later depression[20]. That study focused on five candidate genes. To our knowledge, no studies have explored whether epigenome-wide DNAm patterns in peripheral tissues (such as blood) mediate the relationship between adversity and later depressive symptoms. We thus investigated the extent to which DNAm statistically mediates, or explains, the relationship between common types of childhood adversity and depressive symptoms in early adolescence.

Results

Analytic sample

We analyzed childhood adversity and depressive symptoms at age 10.6 years using data from the Accessible Resource for Integrated Epigenomic Studies (ARIES), a subset of the Avon Longitudinal Study of Parents and Children (ALSPAC), with epigenome-wide DNAm measured at age 7 years[21–23]. Specifically, we focused on seven types of childhood adversity, measured four to six times from birth to age 7 years, previously associated with epigenetic marks[6–9] and elevated risk for depression in youth[24,25]: (1) caregiver emotional or physical abuse; (2) sexual or physical abuse (by anyone); (3) maternal psychopathology; (4) one-adult household; (5) family instability; (6) financial hardship; (7) neighborhood disadvantage (Supplementary Table 1; $N = 627–675$).

Participants in the largest analytic sample (that is, those with complete data on caregiver physical/emotional abuse and DNAm) did not differ from the ARIES sample (Supplementary Table 2). The ARIES sample differed from the full ALSPAC cohort on all demographic variables except sex and depressive symptoms. Most differences were small, except for maternal smoking and maternal education, in which ARIES mothers smoked less and had higher education levels. Depressive symptom scores averaged 3.70 (median, 3; range, 0–21) out of a possible score of 26. The prevalence of children with possible depression (score ≥ 8)[26] within our analytic sample was 12%, comparable to the estimated prevalence among US adolescents[27]. In our analytic samples, exposure to sexual/physical abuse had the lowest prevalence at 3.3%, and maternal psychopathology had the highest at 13.5% (Extended Data Fig. 1).

Sensitive period exposures predict depressive symptoms

We first investigated whether the timing of adversity influenced the relationship between childhood adversity and depressive symptoms using the Structured Life Course Modeling Approach (SLCMA; Extended Data Fig. 2)[28–31]. For all adversities, exposure during very early childhood (ages 0–2 years) or early childhood (ages 3–5 years) explained the most variation in depressive symptoms at age 10.6 years (Table 1). These time periods were used as the exposure measures in subsequent mediation analyses.

Maternal psychopathology showed the strongest association with depressive symptoms ($P = 0.02$). Across all adversities except caregiver abuse, adversity exposure was linked to increased depressive symptoms, explaining up to 14.4% of the variation in depressive symptoms (37.9% correlation; Table 1). Counterintuitively, exposure to caregiver physical/emotional abuse was associated with lower depressive symptoms. However, this association was the weakest identified ($P = 0.70$) and had large confidence intervals overlapping with zero, probably reflecting a null relationship. Despite this finding, we examined all seven adversities, because (1) the link between childhood adversity and depressive symptoms is well established; (2) associations may be weaker compared to other epidemiological studies given our modest sample size; (3) the absence of a significant total effect does not reflect the absence of mediation, as mediated and non-mediated effects can have opposing signs that counteract one another, making it appear as though the total effect is null[32,33].

Mediation was detected at 70 CpGs across the epigenome

Next, we assessed mediation though DNAm by fitting multiple mediator models for each adversity, our analytic set of 278,586 CpGs, and depressive symptoms at age 10.6 years (Extended Data Fig. 2). Seventy CpG sites across all seven adversities showed evidence of being mediators, each explaining 10–73% of the observed correlation between exposure to adversity during the selected sensitive period and depressive symptoms. Of these 70 sites, three were selected in two different adversities (cg10953317, cg24059871 and cg22239534). Thirty-seven sites were significant at the Monte Carlo assigned $P < 0.05$ level. See Supplementary Table 3 for the details on all 70 sites, including their effect estimates, associations with genetic variation (methylation quantitative trait loci; 19 in total) and annotated gene.

Table 2 summarizes the number of mediation sites identified per adversity, as well as their estimated direct effect (that is, the effect of adversity on depression, independent of mediating effects), total indirect effect (that is, the portion of the total effect mediated by DNAm at each CpG) and total effect (sum of direct and total indirect effects). Because childhood adversity and depressive symptom data

https://doi.org/10.1038/s44220-024-00345-8

Table 2 | Mediators (CpGs) selected for each childhood adversity and their total standardized effect

Adversity	Number of mediators selected	Std TE[a]	Percent TE explained by mediation	Percent TE explained by direct effect[b]
Caregiver physical or emotional abuse	5	−0.049	−27	127
Sexual or physical abuse	6	0.100	−73	173
Maternal psychopathology	9	0.117	−22	122
One-adult in the household	10	0.062	−71	171
Family instability	8	0.068	10	90
Financial hardship	17	0.093	30	70
Neighborhood disadvantage	15	0.082	−13	113

[a]The total effect (TE) of childhood adversity on depressive symptoms, standardized so the estimate can be interpreted as the correlation between childhood adversity and depressive symptoms. [b]Direct effect is the effect of childhood adversity on depressive symptom score unrelated to mediation, listed here in terms of the percent of the total effect. The results of the seven mediation analyses use the Schaid–Sinnwell method. When the non-mediated (direct) effect has a percentage greater than 100%, it means that the mediated effect has the opposite sign and the total effect (sum of mediated and direct effects) will add to 100%.

were standardized before mediation analyses, these effect estimates are interpreted as partitioning the correlation between exposure and outcome, after adjusting for confounders. For example, 17 mediators contributed 30% of the observed total effect (or correlation) between financial hardship and depressive symptom score. The remaining 70% was due to the non-mediated (direct) effects.

For some adversities (sexual/physical abuse, maternal psychopathology, one-adult household and neighborhood disadvantage), the mediated effect was negative, and the direct effect was positive. In other words, DNAm reduced the total effect of childhood adversity on depressive symptoms, potentially indicating compensatory or protective mechanisms.

Mediating CpGs showed both risk and protective patterns

We identified four patterns among mediating CpGs, based on the signs of the coefficients relating adversity to DNAm (α) and DNAm to depressive symptoms (β) (Extended Data Fig. 3). The product of these coefficients ($\alpha\beta$) is the indirect effect of the mediating CpG, or the mediated effect. These patterns were: positive–positive (positive α and β), negative–negative (negative α and β), positive–negative (positive α and negative β) and negative–positive (negative α and positive β) (Fig. 1).

Overall, 44% of mediating CpGs had a net positive mediated effect (19 positive–positive; 12 negative–negative; Fig. 2a,b), where children exposed to adversity had changes in DNAm that correlated with increased depressive symptoms (that is, risk-increasing effects). By contrast, 56% of CpGs had a negative mediated effect (23 positive–negative; 16 negative–positive; Fig. 2c,d), whereby children exposed to adversity instead had changes in DNAm that reduced the effect of adversity on depressive symptoms (that is, protective effects).

Of note, CpGs associated with physical/sexual abuse were overrepresented for protective effects (all six CpGs; $P = 0.016$) and those linked to maternal psychopathology showed slight enrichment (seven of nine CpGs; $P = 0.09$). No differences in risk versus protective proportions were identified between adversities occurring during very early childhood or early childhood ($P = 0.81$; $\chi^2 = 0.058$; Fig. 1). These findings suggest some DNAm differences may suppress, rather than increase, future risk for depression.

Mediating CpGs were biologically relevant

We investigated the biological relevance of the 70 mediating CpGs through trait and gene ontology enrichment analyses in the EWAS Toolkit from the EWAS Atlas[34]. Mediating CpGs were enriched for sites previously associated with preterm birth (six CpGs, Bonferroni-adjusted $P < 6 \times 10^{-5}$), severe acute malnutrition in adults (one CpG, Bonferroni-adjusted $P < 0.006$) and postnatal maternal depression (one CpG, Bonferroni-adjusted $P < 0.003$) (Supplementary Fig. 1). Mediating CpGs were also enriched for 20 biological processes (false discovery rate (FDR) < 0.05), including several related to immune function and negative epigenetic regulation of gene expression (Supplementary Fig. 2). Mediating loci were not enriched in promoters, enhancers or CpG islands (Supplementary Fig. 3). However, the overall number of mediators identified was lower than random chance in the TSS1500 region, and positive–positive-patterned mediators were enriched in gene bodies ($P < 0.05$) (Supplementary Fig. 4). Brain–blood correlations for mediating loci were generally positive ($r_{avg} = 0.26$), with seven loci showing correlations >0.8 (Supplementary Fig. 5), suggesting concomitant DNAm differences could potentially co-occur in central tissues[35]. Mediating CpGs were no more likely to show changes in DNAm over time than by random chance, suggesting that the mediated effects were not due to age-dependent differences in DNAm[36] (Supplementary Information and Supplementary Fig. 6). Finally, no genes overlapped with those previously identified in genome-wide association studies (GWAS) of major depressive disorder, as indexed by the GWAS catalog[37], nor did they overlap with known GWAS signals in colocalization analyses (Supplementary Information).

As cell type is a major driver of DNAm, we investigated whether cell-type proportions mediated the relationship between adversity and depressive symptoms (Supplementary Information). We did not find any evidence of mediation through any of the six bioinformatically predicted cell types across all adversities (lowest $P = 0.44$; mean mediated effect = 0.73%), suggesting mediated effects were specific to DNAm itself (Supplementary Fig. 7).

Ever-exposed models of adversity showed distinct mediation patterns

Given the novelty of the SLCMA to test different life course hypotheses relating adversity to depression, we compared the respective R^2 values of our results to those obtained from a traditional approach comparing the presence versus absence of childhood adversity. In all instances except one, the sensitive period selected in the SLCMA explained more variability (that is, larger R^2 values) in depressive symptoms than an ever-exposed hypothesis.

Mediation analyses using the ever-exposed model revealed a similar number of mediators (74 versus 70 CpGs; Supplementary Table 4). However, only six of these 74 CpGs overlapped with our primary results. This finding suggests that different ways of categorizing exposure (sensitive period versus ever-exposed) are not biologically interchangeable. However, less emphasis should be placed on individual CpG overlap and more on the group of mediators, because our analytic approach selected mediators representing a set that may work together.

Protective effects of DNA methylation were replicated

We sought to replicate our primary findings in two independent cohorts: the Future of Families and Child Wellbeing Study (FFCWS; $N = 1{,}319$–$1{,}677$; saliva DNAm)[38] and the Generation R Study (GenR; $N = 312$–392; blood DNAm)[39]. These cohorts differed from ALSPAC in multiple respects, including a generally higher prevalence of adversity (Extended Data Fig. 1), differences in depressive symptom measures, sensitive periods identified in the SLCMA (Supplementary Table 5) and sociodemographic characteristics such as race/ethnicity (details are provided in the Supplementary Information).

Given these differences between cohorts, we focused on replicating the mediated effects of our primary set of 70 mediating CpGs, rather than recapitulating the primary analyses in additional datasets (Extended Data Fig. 4). Here, we present mediated effects corrected for the direction of the adversity–depression relationship, which allowed

https://doi.org/10.1038/s44220-024-00345-8

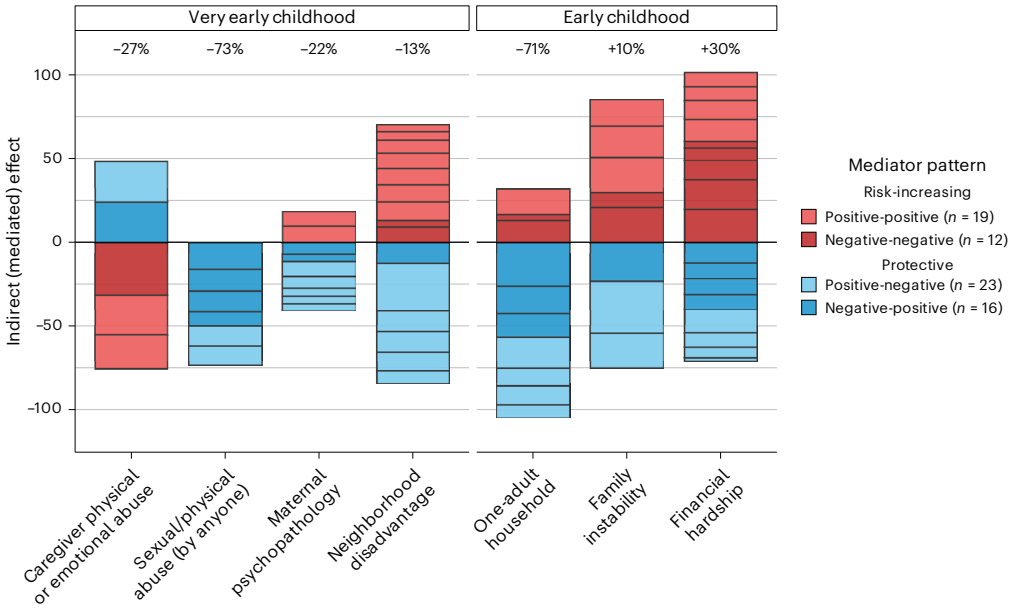


Fig. 1 | Summary of indirect (mediated) effects across adversities. Adversities are grouped based on the sensitive period when they had the greatest impact on depressive symptoms. Bars represent the CpGs selected as mediators for each adversity, with each box representing the indirect, or mediated, effect of that CpG. The total indirect effect for the collection of identified CpGs for each type of adversity is shown at the top of the graph, reflected in the difference between positive and negative mediated effects. Colors represent the net effect of the CpG on the adversity–depression relationship, where red reflects a risk-increasing effect and blue a protective effect on depressive symptoms. Note that the colors are flipped for caregiver abuse due to this adversity having a small negative relationship with depressive symptoms in our analytic sample. However, as our sample size was modest, this negative effect may instead reflect a null relationship between this adversity and depressive symptoms.

us to determine whether CpGs had risk-increasing or protective effects on depressive symptoms. Because *P* values may be unreliable within and across epigenome-wide studies[9,40,41], we used the magnitude and direction of the indirect (mediated) effects as metrics for replication, focusing primarily on the risk-increasing or protective effects of DNAm on depressive symptoms.

Overall, we largely replicated the role of DNAm in the relationship between adversity and depressive symptoms when combining the indirect effects of all CpGs for each adversity. There were different magnitudes of mediated effect across cohorts, sometimes with stronger associations identified in the replication samples (Fig. 3). Regarding the direction of total mediated effect per adversity, the risk-increasing or protective effects of DNAm were consistent across studies for caregiver abuse (risk-increasing in ALSPAC, FFCWS and GenR), maternal psychopathology (protective in ALSPAC and GenR), one-adult households (protective in ALSPAC and FFCWS) and financial hardship (risk-increasing in ALSPAC and GenR).

Similar to ALSPAC, we identified both risk-increasing and protective effects of DNAm across every adversity (caregiver abuse, maternal psychopathology, family instability and financial hardship) in both FFCWS and GenR. In the 35 CpGs measured across all three cohorts, 15 (43%) showed the same direction of indirect effects on depressive symptoms (risk-increasing or protective) in all studies, which was considerably higher than random chance ($P = 0.0022$; Fig. 3). When further parsing these loci into risk-increasing versus protective effects on depressive symptoms, protective CpGs were more strongly replicated (10 of 17 CpGs; $P = 0.00063$) than risk-increasing CpGs (6 of 18 CpGs; $P = 0.14$). See Supplementary Information for details of CpG-level replication in FFCWS (Supplementary Fig. 8 and Supplementary Tables 6 and 7) and GenR (Supplementary Fig. 9 and Supplementary Tables 8 and 9).

Discussion

The main finding of this prospective longitudinal study is that childhood adversity can partially influence DNAm signatures in ways that shape depressive symptoms in adolescence. Overall, we identified 70 CpG sites that mediated (or explained) between 10 and 73% of the correlation between adversity and depressive symptoms in ALSPAC. Interestingly, for most adversities—including sexual/physical abuse, maternal psychopathology, one-adult household and neighborhood disadvantage—mediating effects were primarily negative, suggesting that DNAm levels were protective against depressive symptoms. By contrast, for family instability and financial hardship, the mediating effect was positive, whereby DNAm levels at the identified CpG mediators worsened depressive symptoms.

There were diverse patterns of mediation among these 70 sites, suggesting that if mediation occurs through DNAm, it is not a directionally uniform process. These findings corroborate what is already known: epigenetic adaptation is complex, and childhood adversity may be a plasticity factor shaping epigenetic changes[8,42,43]. Several identified CpGs are in or near genes related to brain and cortical development and have been associated with developmental delays and disabilities, such as *RARB*[44], *SLC4A8*[45], *DIO2*[46], *DPYSL3*[47] and *HDAC4*[48]. Of note, two of these genes, *SLC4A8* and *DIO2*, were replicated in independent cohorts, both with protective effects of DNAm on depressive symptoms. Although these results suggest that DNAm may be on the causal pathway of adversity to depressive symptoms, mechanistic studies in model systems and central tissues are necessary to delineate the specific biological pathways underlying these relationships. In this context, DNAm changes in peripheral tissues (blood, saliva, etc.) may instead reflect a measurable signature of the underlying biological alterations driving disease risk. Potential mechanisms include changes in more central tissues, such as the brain, or upstream biological processes, such as chromatin modifications and structural alterations to DNA organization[49,50]. Thus, these findings should be supplemented with functional experiments in both animal-based and human-derived model systems to disentangle the specific role of DNAm in the progression and resolution of mental illness.

Fundamentally, our findings suggest that childhood adversity is linked to both deleterious and protective changes in DNAm. From a

https://doi.org/10.1038/s44220-024-00345-8

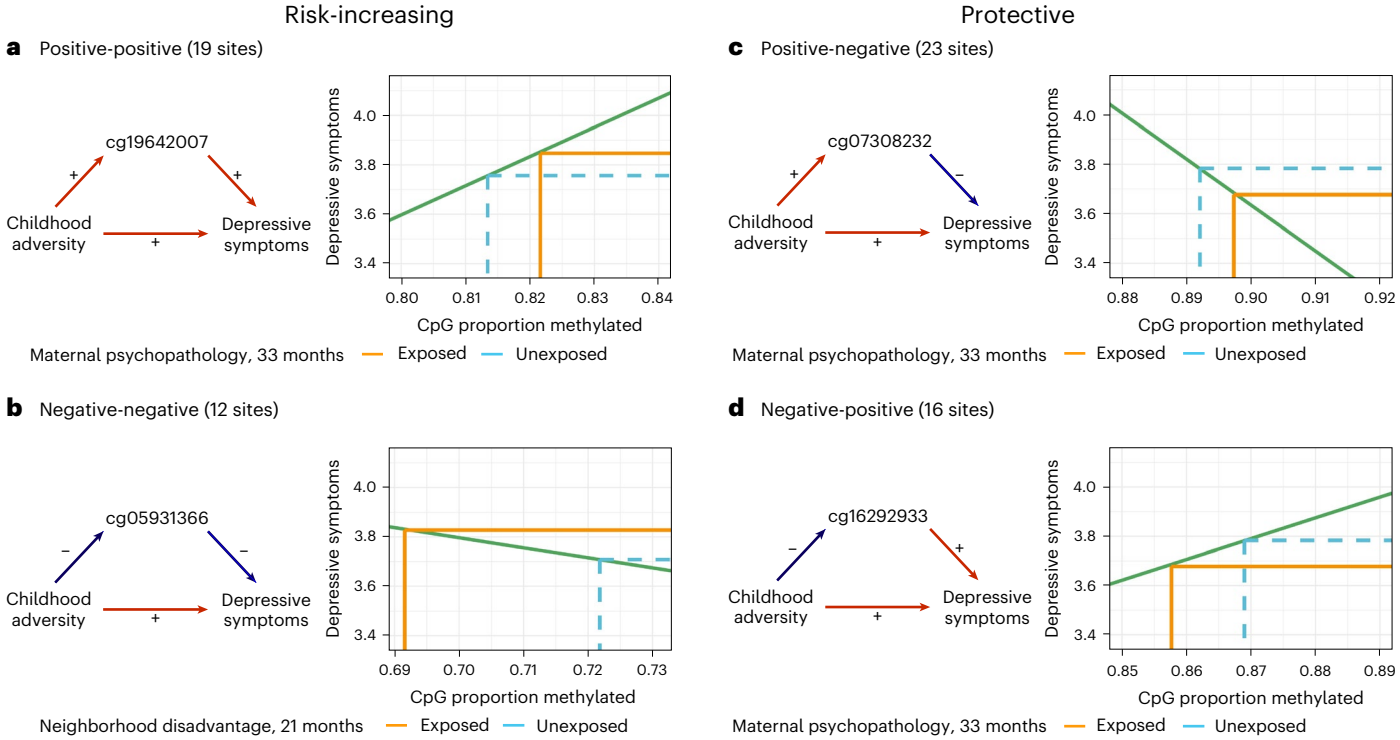


Fig. 2 | Direction of mediation patterns for individual CpG sites. a–d, The four mediating patterns: positive–positive (**a**), negative–negative (**b**), positive–negative (**c**) and negative–positive (**d**). The directed acyclic graph on the left illustrates the relationships among the childhood adversity, CpG mediator and depressive symptoms. Red lines indicate a positive effect and black lines a negative effect. Blue and orange lines represent mean values for unexposed versus exposed groups, respectively. In each zoomed-in plot, where the blue and orange lines meet the *x* axis reflects the mean DNAm values for each exposure group. The difference between the blue and orange lines on the *y* axis represents the indirect (mediated) effect of the CpG. The indirect effect is positive when the orange is above the blue, and negative when the blue is above the orange. Beta values are plotted here, for more intuitive biological interpretation (that is, the proportion of cells with DNAm). In **a**, average DNAm levels at the CpG were higher in the exposed group, which correlated with higher SMFQ scores. In **b**, unexposed individuals had higher DNAm levels on average at the CpG, and that higher DNAm correlated with lower SMFQ. In **c**, exposed children showed higher average DNAm levels at the CpG, whereas the SMFQ score decreased as DNAm levels increased. In **d**, exposed children had lower DNAm levels, but higher DNAm levels correlated with higher SMFQ scores.

clinical perspective, these DNAm patterns could potentially be used as biological indicators of risk or resilience to prevent new onsets of depression—or possibly even as metrics for treatments that reverse or minimize symptoms in people already suffering from depression. Recent intervention studies using psychotherapy suggest that DNAm may show concomitant responses to treatment, or act as markers of symptom resolution[16–18]. For example, an investigation into longitudinal DNAm profiles of soldiers with post-traumatic stress disorder (PTSD) found that successful treatment of the disorder with psychotherapy was linked to DNAm changes in 12 differentially methylated regions[15]. Notably, changes in DNAm at gene *ZFP57* were specifically linked to both the development and remission of PTSD[15]. Although DNAm may not be the direct causal mechanism in these relationships, our findings suggest that DNAm changes could potentially act as a peripheral biomarker for the efficacy of interventions at the molecular level. Should our findings be further replicated, these epigenetics signatures could also be leveraged as predictors of biological resilience and used in conjunction with other risk assessment tools to predict who may be more vulnerable to mental illness following childhood adversity. To this end, recent evidence suggests DNAm risk scores (MRS) of early-life exposures or disease processes could be used as predictors of risk[51–53]. Although our results are not yet clinically actionable, they provide important new insights into the potential applications of DNAm in the prevention and treatment of mental illness.

We also partially replicated our results in two independent datasets, particularly around the protective role for DNAm in the relationship between childhood adversity and depressive symptoms. This finding contrasts the majority of epigenetic studies, which traditionally attribute a risk-increasing role to DNAm in human health and disease[49]. Instead, our results highlight DNAm as a potential resiliency factor or compensatory mechanism against the deleterious effects of adversity on psychopathology. Yet, these findings do not necessarily stand in opposition to current literature. Most epigenome-wide studies cannot disentangle risk-increasing and protective effects of DNAm on exposure–outcome relationships, as they lack the longitudinal study designs required to investigate mediation. Moreover, previous studies in human populations[54] and animal models[55] have identified a compensatory role for DNAm, though on a smaller scale than our study. For instance, increased DNAm in the glucocorticoid receptor gene (*NR3C1*) reduces the impact of maternal anxiety during pregnancy on child behavioral outcomes[54]. We also recently showed that adversity-associated DNAm differences can have both risk-increasing and protective effects on health outcomes[56], further supporting a potential role for DNAm in resilience. However, not all mediated effects were consistent across cohorts and adversities, suggesting there is additional granularity and complexity to uncover in these relationships. Additional epigenome-wide studies leveraging similarities and differences across contexts and populations are needed to triangulate the sociodemographic and biological factors that explain the negative and positive ways in which DNAm responds to life experience and influences future health outcomes[57].

Our study had several strengths, which allowed us to detect novel roles for DNAm in the relationship between adversity and depression. First, we used prospective longitudinal data from a population-based

https://doi.org/10.1038/s44220-024-00345-8

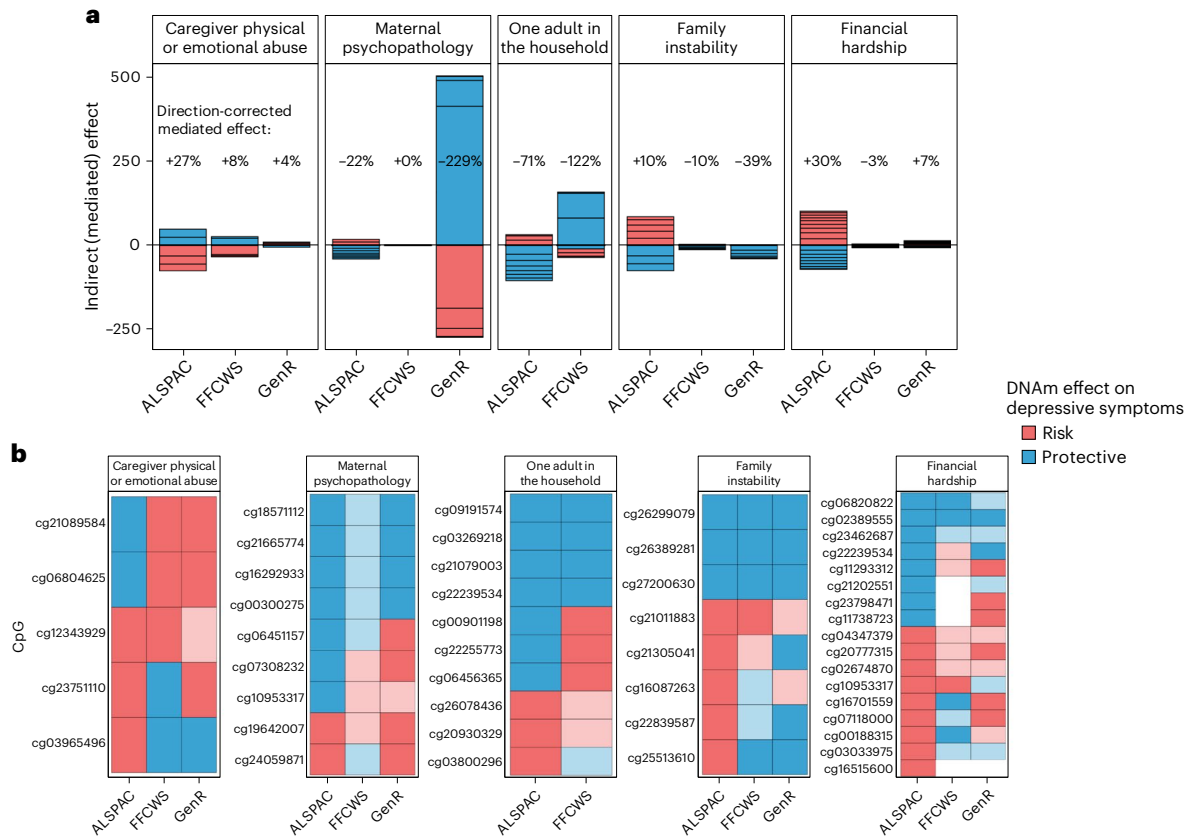


Fig. 3 | Summary of the replication results for mediated effects. Colors represent the effect of the CpG on the adversity–depression relationship, where red and blue reflect risk-increasing or protective effects on depressive symptoms, respectively. Only a subset of adversities were available in the FFCWS and GenR study. **a**, Summary of mediated effects across adversities and cohorts, with each box representing a specific CpG and its indirect (mediated) effect. The total indirect, or mediated, effect across all CpGs is corrected for the direction of the adversity–depression relationship to reflect either risk-increasing (positive) or protective (negative) effects of DNAm. In instances with >100% indirect effect, other mediating factors captured in the direct effect may act as strongly in the opposite direction to reduce the total effect (40). **b**, Direction of the mediated effect for each CpG measured in the replication cohorts. CpGs with an absolute indirect (mediated) effect of <1% are shown in lighter shades (that is, those with very small effects). White boxes reflect CpGs that were not measured in replication cohorts.

sample with repeated measures spanning over a decade. This data structure allowed us to explore relationships between the timing and frequency of childhood adversity and subsequent depressive symptoms. The prospective and longitudinal study design also minimized the likelihood of reverse causation and recall bias, two limitations often cited against social epigenetic studies and mediation analyses[58]. Furthermore, we used a novel mediation analysis technique that allowed all potential CpG mediators (n = 278,586) to have equal consideration for inclusion in the mediation model. As CpGs surviving the dimension-reduction process were included within the same model, our estimated mediated effects are more accurate than if the mediators had been considered separately[59]. Finally, we triangulated our findings in two separate studies[57], which measured slightly different dimensions of childhood adversity, DNAm and depressive symptoms. Triangulation allowed us to make stronger inferences about our main findings and identify cross-cutting features of the adversity–DNAm–depression relationship.

Our study also had limitations. As described in the Supplementary Information, our mediation model worked iteratively and used a penalty term to select the optimal number of model parameters. A large penalty term selects no mediators, whereas reducing the penalty term increases the number of mediators selected. Because of our sample size, we opted to balance false-positive and false-negative results by choosing a modest penalty term for all seven adversities. Ideally, the sample size would be much larger to accommodate a more stringent penalty term that simultaneously allows for greater mediation signal detection. Second, differences in the prevalence of adversity, measurements of adversity, DNAm and depressive symptoms, and the overall composition of our three cohorts may have influenced the sensitive periods identified in the SLCMA across cohorts and limited the reproducibility of our mediation results. Despite these differences, we reproduced cross-cutting features of DNAm as a mediator of the relationship between adversity and depressive symptoms, which may be more robust than those identified in identical cohorts. Third, we could not combine data across cohorts or meta-analyze epigenome-wide findings from all three cohorts due to data access and methodological considerations. Although future studies combining data across cohorts may identify CpGs with stronger and more consistent mediating effects, our discovery and replication approach yielded valuable insights into the role of DNAm in the relationship between adversity and depression. Fourth, our discovery data from ALSPAC came from a racially homogeneous group (97% European ancestry). Racially diverse groups are needed to increase the generalizability of future research. Finally, DNAm from peripheral tissues, like blood and saliva, is an imperfect proxy for DNAm levels in the brain, which limits the interpretability of our results in the context of brain-based disorders. Future studies should examine whether the DNAm changes we observed in this

study also occur in the central nervous system to help determine their role in brain-based disorders.

Conclusions

Together, our results suggest that DNAm levels may partially explain the relationship between exposure to childhood adversity and depressive symptoms in early adolescence. We also highlight a novel paradigm for epigenetic studies investigating the consequences of early-life stressors on human health and behavior, whereby DNAm may act both as a risk-increasing and protective mechanism in this relationship. Ultimately, these findings will extend our understanding of the complex biological mechanisms driving depression and aid in the identification of better treatments for mood disorders.

Data availability

All data are available by request from the ALSPAC Executive Committee for researchers who meet the criteria for access to confidential data (bristol.ac.uk/alspac/researchers/access). The FFCWS data analyzed in the current study are available with permission from the Future of Families and Childhood Wellbeing Study repository (ffcws.princeton.edu/documentation). Generation R Study data are not publicly available due to privacy or ethical restrictions. The study has an open policy in regard to collaboration with other research groups. Requests for collaboration should primarily be addressed to V. Jaddoe (ln.cmsumsare@eoddaj.v). These requests are discussed in the Generation R Study Management Team regarding their study aims, overlap with ongoing studies, logistic consequences and financial contributions.

Code availability

The scripts used to complete the primary analyses in this manuscript are available on GitHub (github.com/thedunnlab/mediation).

Materials and methods

Sample and measures

The data were taken from ALSPAC, a 30-year ongoing study that recruited pregnant women in Avon, England with expected delivery dates between April 1991 and December 1992[21,22]. We analyzed data from ARIES, a subset of 1,018 mother–child pairs in ALSPAC who had complete phenotype data from at least five waves of data collection[23].

Depressive symptoms. Depressive symptoms were assessed at a mean age of 10.6 years using total scores from the child-completed SMFQ, a well-established measure of depressive symptoms in youth[60,61]. We used total SMFQ scores as our measure of depressive symptoms—rather than a binary indicator for probable depression—because continuous scores enable detection of more subtle variability, leading to greater precision and improved statistical power. We selected the age 10.6 years timepoint because it was the first self-reported measure of depressive symptoms collected after the DNAm assessment. See Supplementary Information for details and the full rationale for this measure.

DNA methylation. As described elsewhere[23], genome-wide DNAm was measured from blood at 485,577 CpGs using the Infinium Human Methylation 450K BeadChip microarray (Illumina) when the children were 7 years old. To reduce the dimensionality of our analyses, we performed pruning and sure independence screening, a dimension reduction method that selects CpGs with the highest marginal correlations with both the exposure (adversity) and outcome (depressive symptoms), resulting in an analytic set of 278,586 CpGs (Supplementary Information).

Covariates. We controlled for the following sociodemographic characteristics previously linked to adversity, depression and differences in DNAm[62–64]: child sex, birthweight, race and/or ethnicity, maternal age at birth, maternal education at birth, number of previous pregnancies and maternal smoking during pregnancy (Supplementary Information and Supplementary Table 2).

Data analysis

We analyzed each type of adversity separately in the three phases as described below and outlined in Extended Data Fig. 2. Our analytic samples included singleton-birth children with complete data on adversity, depressive symptoms, DNAm and all covariates (N = 627–675; Table 1).

Identifying the best explanatory life course exposure. As previous studies reported developmental timing differences in the association of adversity with DNAm[7–9], we used the SLCMA[28–31] to identify the life course hypothesis[65,66] for each adversity that explained the most variation in depressive symptoms at age 10.6 years. We considered five life course hypotheses: (1) a very early childhood sensitive period, in which adversity had a greater impact on adolescent depressive symptoms during the period 0–2 years of age; (2) an early childhood sensitive period, 3–5 years of age; (3) a middle childhood sensitive period, 6–7 years of age; (4) accumulation of risk, in which the impact of adversity increased with repeated exposures, regardless of timing; and (5) recency, in which adversity had a greater impact for more proximally occurring (rather than distally occurring) exposures. The single best-fitting life course model identified by the SLCMA for each adversity was used in subsequent mediation analyses. Details are provided in the Supplementary Information.

Assessing mediation through epigenome-wide DNAm. We fit multiple mediator models using the sparse group lasso penalized model, a structural equation model (SEM) approach developed by Schaid and Sinnwell (dubbed the 'Schaid–Sinnwell model')[67]. SEMs are optimal for situations with multiple mediators, because they simultaneously model the relationship among variables, account for correlation between mediators, and estimate the direct effect and indirect effect[59,68,69].

The Schaid–Sinnwell model is ideally suited to epigenome-wide mediation analyses as it groups the effect estimates for each mediator (that is, both exposure–mediator and mediator–outcome associations), rather than considering them separately. It also encourages sparseness (parsimony) of the parameters by using a penalty (shrinkage) parameter[67]. Also advantageously, it selects mediators based on the Bayesian information criterion (BIC), an approach that avoids model over-fitting[70]. Although the BIC is a model selection criterion not reliant on P values, we also report P values calculated using Monte Carlo approximations for comparability. Details on the implementation of the Schaid–Sinnwell model and parameter estimation are provided in the Supplementary Information.

Replication in independent cohorts. We sought to replicate our findings in two independent birth cohorts: (1) the FFCWS (n = 1,319–1,677)[38] and (2) GenR (n = 312–392)[39]. We performed CpG-specific mediation analyses using the mediation package in R for each adversity, matched to the timing and type of adversity measured in ALSPAC. We used this approach for two reasons: (1) we previously showed that the timing of exposures can influence different sets of CpGs[64], and (2) we attempted to limit the influence of cohort-levels differences to improve our ability to replicate our primary results. In FFCWS, we analyzed a subset of CpGs associated with five types of childhood adversity (caregiver abuse, maternal psychopathology, one-adult households, family instability and financial hardship), DNAm from saliva at age 9 years, and depressive symptoms at age 15 years. In GenR, we analyzed CpGs associated with four types of childhood adversity (caregiver abuse, maternal psychopathology, family instability and financial hardship), DNAm from blood at age 9 years and depressive symptoms at age 13–14 years. Details are provided in the Supplementary Information.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this Article.

References

1. WHO. *World Health Statistics 2020: Monitoring Health for the SDGs, Sustainable Development Goals*. viii, 77 (World Health Organization, 2020).
2. Li, M., d'Arcy, C. & Meng, X. Maltreatment in childhood substantially increases the risk of adult depression and anxiety in prospective cohort studies: systematic review, meta-analysis and proportional attributable fractions. *Psychol. Med.* **46**, 717–730 (2015).
3. Lippard, E. T. C. & Nemeroff, C. B. The devastating clinical consequences of child abuse and neglect: increased disease vulnerability and poor treatment response in mood disorders. *Am. J. Psychiatry* **177**, 20–36 (2020).
4. Grummitt, L. R. et al. Association of childhood adversity with morbidity and mortality in US adults: a systematic review. *JAMA Pediatr.* **175**, 1269–1278 (2021).
5. Jones, P. A. & Takai, D. The role of DNA methylation in mammalian epigenetics. *Science* **293**, 1068–1070 (2001).
6. Cecil, C. A. M., Zhang, Y. & Nolte, T. Childhood maltreatment and DNA methylation: a systematic review. *Neurosci. Biobehav. Rev.* **112**, 392–409 (2020).
7. Parade, S. H. et al. A systematic review of childhood maltreatment and DNA methylation: candidate gene and epigenome-wide approaches. *Transl. Psychiatry* **11**, 134 (2021).
8. Dunn, E. C. et al. Sensitive periods for the effect of childhood adversity on DNA methylation: results from a prospective, longitudinal study. *Biol. Psychiatry* **85**, 838–849 (2019).
9. Lussier, A. A. et al. Updates to data versions and analytic methods influence the reproducibility of results from epigenome-wide association studies. *Epigenetics* **17**, 1373–1388 (2022).
10. Nöthling, J., Malan-Müller, S., Abrahams, N., Hemmings, S. M. J. & Seedat, S. Epigenetic alterations associated with childhood trauma and adult mental health outcomes: a systematic review. *World J. Biol. Psychiatry* **21**, 493–512 (2020).
11. Li, M. et al. What do DNA methylation studies tell us about depression? A systematic review. *Transl. Psychiatry* **9**, 68 (2019).
12. Park, C. et al. Stress, epigenetics and depression: a systematic review. *Neurosci. Biobehav. Rev.* **102**, 139–152 (2019).
13. Jovanova, O. S. et al. DNA methylation signatures of depressive symptoms in middle-aged and elderly persons: meta-analysis of multiethnic epigenome-wide studies. *JAMA Psychiatry* **75**, 949–959 (2018).
14. Clark, S. L. et al. A methylation study of long-term depression risk. *Mol. Psychiatry* **25**, 1334–1343 (2020).
15. Vinkers, C. H. et al. Successful treatment of post-traumatic stress disorder reverses DNA methylation marks. *Mol. Psychiatry* **26**, 1264–1271 (2021).
16. Ziegler, C. et al. MAOA gene hypomethylation in panic disorder-reversibility of an epigenetic risk pattern by psychotherapy. *Transl. Psychiatry* **6**, e773 (2016).
17. Thomas, M. et al. Increased BDNF methylation in saliva, but not blood, of patients with borderline personality disorder. *Clin. Epigenetics* **10**, 109 (2018).
18. Hoye, J. R. et al. Preliminary indications that the attachment and biobehavioral catch-up intervention alters DNA methylation in maltreated children. *Dev. Psychopathol.* **32**, 1486–1494 (2020).
19. Fujii, R., Sato, S., Tsuboi, Y., Cardenas, A. & Suzuki, K. DNA methylation as a mediator of associations between the environment and chronic diseases: a scoping review on application of mediation analysis. *Epigenetics* **17**, 759–785 (2021).
20. Peng, H. et al. Childhood trauma, DNA methylation of stress-related genes, and depression: findings from two monozygotic twin studies. *Psychosom. Med.* **80**, 599–608 (2018).
21. Boyd, A. et al. Cohort profile: the 'Children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).
22. Fraser, A. et al. Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
23. Relton, C. L. et al. Data resource profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). *Int. J. Epidemiol.* **44**, 1181–1190 (2015).
24. Dunn, E. C. et al. What life course theoretical models best explain the relationship between exposure to childhood adversity and psychopathology symptoms: recency, accumulation or sensitive periods? *Psychol. Med.* **48**, 2562–2572 (2018).
25. McLaughlin, K. A. Future directions in childhood adversity and youth psychopathology. *J. Clin. Child Adolesc. Psychol.* **45**, 361–382 (2016).
26. Thapar, A. & McGuffin, P. Validity of the shortened Mood and Feelings Questionnaire in a community sample of children and adolescents: a preliminary research note. *Psychiatry Res.* **81**, 259–268 (1998).
27. Merikangas, K. R. et al. Lifetime prevalence of mental disorders in US adolescents: results from the National Comorbidity Survey Replication-Adolescent Supplement (NCS-A). *J. Am. Acad. Child Adolesc. Psychiatry* **49**, 980–989 (2011).
28. Smith, B. J., Smith, A. D. A. C. & Dunn, E. C. in *Sensitive Periods of Brain Development and Preventive Interventions Current Topics in Behavioral Neurosciences* (ed. Andersen, S. L.) Ch. 9, 215–234 (Springer, 2021).
29. Smith, A. D. et al. A structured approach to hypotheses involving continuous exposures over the life course. *Int. J. Epidemiol.* **45**, 1271–1279 (2016).
30. Smith, A. D. et al. Model selection of the effect of binary exposures over the life course. *Epidemiology* **26**, 719–726 (2015).
31. Zhu, Y. et al. A structured approach to evaluating life course hypotheses: moving beyond analyses of exposed versus unexposed in the omics context. *Am. J. Epidemiol.* **190**, 1101–1112 (2021).
32. Shrout, P. E. & Bolger, N. Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychol. Methods* **7**, 422–445 (2002).
33. Loeys, T., Moerkerke, B. & Vansteelandt, S. A cautionary note on the power of the test for the indirect effect in mediation analysis. *Front. Psychol.* **5**, 1549 (2014).
34. Li, M. et al. EWAS Atlas: a curated knowledgebase of epigenome-wide association studies. *Nucleic Acids Res.* **47**, D983–D988 (2019).
35. Braun, P. R. et al. Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals. *Transl. Psychiatry* **9**, 47 (2019).
36. Mulder, R. H. et al. Epigenome-wide change and variation in DNA methylation in childhood: trajectories from birth to late adolescence. *Hum. Mol. Genet.* **30**, 119–134 (2021).
37. Sollis, E. et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res.* **51**, D977–D985 (2023).
38. Reichman, N. E., Teitler, J. O., Garfinkel, I. & McLanahan, S. S. Fragile families: sample and design. *Child. Youth Serv. Rev.* **23**, 303–326 (2001).
39. Jaddoe, V. W. et al. The Generation R Study: design and cohort profile. *Eur. J. Epidemiol.* **21**, 475–484 (2006).
40. Amrhein, V. & Greenland, S. Remove, rather than redefine, statistical significance. *Nat. Hum. Behav.* **2**, 4 (2018).

41. McShane, B. B., Gal, D., Gelman, A., Robert, C. & Tackett, J. L. Abandon statistical significance. *Am. Stat.* **73**, 235–245 (2019).
42. Teicher, M. H. & Samson, J. A. Childhood maltreatment and psychopathology: a case for ecophenotypic variants as clinically and neurobiologically distinct subtypes. *Am. J. Psychiatry* **170**, 1114–1133 (2014).
43. Dunn, E. C. Childhood adversity as a plasticity factor that modifies the effect of subsequent life experience on psychopathology. *JAMA Netw. Open* **1**, e185358 (2018).
44. Srour, M. et al. Gain-of-function mutations in RARB cause intellectual disability with progressive motor impairment. *Hum. Mutat.* **37**, 786–793 (2016).
45. Chen, L. M., Haddad, G. G. & Boron, W. F. Effects of chronic continuous hypoxia on the expression of SLC4A8 (NDCBE) in neonatal versus adult mouse brain. *Brain Res.* **1238**, 85–92 (2008).
46. Bárez-López, S., Obregon, M. J., Bernal, J. & Guadaño-Ferraz, A. Thyroid hormone economy in the perinatal mouse brain: implications for cerebral cortex development. *Cerebral Cortex* **28**, 1783–1793 (2018).
47. Tsutiya, A. et al. Human CRMP4 mutation and disrupted Crmp4 expression in mice are associated with ASD characteristics and sexual dimorphism. *Sci. Rep.* **7**, 16812 (2017).
48. Pinto, D. et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am. J. Hum. Genet.* **94**, 677–694 (2014).
49. Cecil, C. A. M., Neumann, A. & Walton, E. Epigenetics applied to child and adolescent mental health: progress, challenges and opportunities. *JCPP Adv.* **3**, e12133 (2023).
50. O'Donnell, K. J. & Meaney, M. J. Epigenetics, development and psychopathology. *Annu. Rev. Clin. Psychol.* **16**, 327–350 (2020).
51. van den Oord, C. L. J. D. et al. DNA methylation signatures of childhood trauma predict psychiatric disorders and other adverse outcomes 17 years after exposure. *Mol. Psychiatry* **27**, 3367–3373 (2022).
52. Kilanowski, A. et al. Methylation risk scores for childhood aeroallergen sensitization: results from the LISA birth cohort. *Allergy* **77**, 2803–2817 (2022).
53. Barbu, M. C. et al. Epigenetic prediction of major depressive disorder. *Mol. Psychiatry* **26**, 5112–5123 (2021).
54. Cao-Lei, L. et al. Epigenetic modifications associated with maternal anxiety during pregnancy and children's behavioral measures. *Cells* **10**, 2421 (2021).
55. Guo, C. et al. DNA methylation protects against cisplatin-induced kidney injury by regulating specific genes, including interferon regulatory factor 8. *Kidney Int.* **92**, 1194–1205 (2017).
56. Schuurmans, I. K., Dunn, E. C. & Lussier, A. A. DNA methylation as a possible causal mechanism linking childhood adversity and health: results from a two-sample Mendelian randomization study. *Am. J. Epidemiol.* 10.1093/aje/kwae072 (2024).
57. Lawlor, D. A., Tilling, K. & Davey Smith, G. Triangulation in aetiological epidemiology. *Int. J. Epidemiol.* **45**, 1866–1886 (2016).
58. Barker, E. D., Walton, E. & Cecil, C. A. M. Annual research review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology. *J. Child Psychol. Psychiatry* **59**, 303–322 (2018).
59. VanderWeele, T. J. & Vansteelandt, S. Mediation analysis with multiple mediators. *Epidemiol. Methods* **2**, 95–115 (2014).
60. Lundervold, A. J., Hinshaw, S. P., Sorensen, L. & Posserud, M. B. Co-occurring symptoms of attention deficit hyperactivity disorder (ADHD) in a population-based sample of adolescents screened for depression. *BMC Psychiatry* **16**, 46 (2016).
61. Patton, G. C. et al. Predicting female depression across puberty: a two-nation longitudinal study. *J. Am. Acad. Child Adolesc. Psychiatry* **47**, 1424–1432 (2008).
62. Liu, J. et al. Socioeconomic changes predict genome-wide DNA methylation in childhood. *Hum. Mol. Genet* **32**, 709–719 (2023).
63. Richmond, R. C. et al. Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Hum. Mol. Genet.* **24**, 2201–2217 (2015).
64. Lussier, A. A. et al. Association between the timing of childhood adversity and epigenetic patterns across childhood and adolescence: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC) prospective cohort. *Lancet Child Adolesc. Health* **7**, 532–543 (2023).
65. Kuh, D., Ben Shlomo, Y. & Ezra, S. (eds). *A Life Course Approach to Chronic Disease Epidemiology*, 2nd edn (Oxford, 2004).
66. Ben-Shlomo, Y. & Kuh, D. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *Int. J. Epidemiol.* **31**, 285–293 (2002).
67. Schaid, D. J. & Sinnwell, J. P. Penalized models for analysis of multiple mediators. *Genet. Epidemiol.* **44**, 408–424 (2020).
68. Hoyle, R. *Handbook of Structural Equation Modeling* 1st edn (Guilford Press, 2012).
69. Preacher, K. J. & Hayes, A. F. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* **40**, 879–891 (2008).
70. Vrieze, S. I. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychol. Methods* **17**, 228–243 (2012).

Acknowledgements

This work was supported by the National Institute of Mental Health of the National Institutes of Health (grant no. R01MH113930 awarded to E.C.D.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. E.C.D. received funding support from One Mind. A.A.L. is supported by a postdoctoral fellowship from the Canadian Institutes of Health Research. E.W. is funded by CLOSER, whose mission is to maximize the use, value and impact of longitudinal studies (www.closer.ac.uk). CLOSER was funded by the Economic and Social Research Council (ESRC) and the Medical Research Council (MRC) between 2012 and 2017. Its initial five-year grant was extended to March 2021 by the ESRC (grant no. ES/K000357/1). E.W. and C.A.M.C. are also supported by the European Union's Horizon 2020 research and innovation program (grant no. 848158). D.J.S. is supported by the US Public Health Service and National Institutes of Health (grant no. R35GM140487). The funders took no role in the design, execution, analysis or interpretation of the data or in the writing up of the findings. We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and Wellcome (grant no. 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf); this research was specifically funded by grants from the BBSRC (BBI025751/1; BB/I025263/1), IEU (MC_UU_12013/1; MC_UU_12013/2; MC_UU_12013/8), the National Institute of Child and Human Development (R01HD068437), NIH (5RO1AI121226-02) and CONTAMED EU (212502). This publication is the work of the authors, who will serve as guarantors for the contents of this paper. The general design of the Generation R Study was made possible by financial support from Erasmus MC, Erasmus University Rotterdam, the Netherlands Organization for Health Research and

Development and the Ministry of Health, Welfare and Sport. EWAS data were funded by a grant from the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA; project no. 050-060-810), by funds from the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, and by a grant from the National Institute of Child and Human Development (R01HD068437). This project received funding from the European Union's Horizon 2020 research and innovation program (848158, EarlyCause). Funding for the Future of Families and Child Wellbeing Study (formerly the Fragile Families and Child Wellbeing Study) was provided by the Eunice Kennedy Shriver National Institute of Child Health and Human Development through grants nos. R01-HD36916, R01-HD39135 and R01-HD40421 and by a consortium of private foundations. Epigenetic data generation and processing were funded by grants by the National Institute on Minority Health and Health Disparities (R01 MD011716), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R01 HD076592) and the National Institute of Mental Health (R01 MH103761).

Author contributions

A.A.L. and B.J.S. conceptualized the study, ran the analyses, interpreted the results, and wrote the manuscript. J.F. and T.S. ran the analyses in FFCWS. M.L. ran the analyses in GenR. J.C. assisted with the ALSPAC analyses and interpretation of primary results. L.S. ran the DNAm microarrays for FFCWS and provided specific input on the FFCWS results. C.A.M.C. and J.F. provided access to GenR data and specific input on the GenR results. C.M. and D.A.N. provided access to FFCWS data and specific input on the FFCWS results. D.J.S. and A.D.A.C.S. provided further input and review of the statistical analyses. K.J.R., D.J.S., A.J.S., M.J.S., E.W. and A.D.A.C.S. helped with the interpretation of the results. E.C.D. conceptualized the study, provided funding and helped write the manuscript. All authors provided critical input on the findings, their interpretation and the manuscript.

Competing interests

The authors declare no competing interests.

Inclusion And Ethics

Ethical approval for the study was obtained from the ALSPAC Law and Ethics Committee and the Local Research Ethics Committees. Consent for biological samples was collected in accordance with the Human Tissue Act of 2004. Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee. Secondary analyses of these data were approved with oversight by the Mass General Brigham Institutional Review Boards (IRB) (Protocol 2017P001110). The Medical Ethical Committee of the Erasmus Medical Centre approved the Generation R Study protocol; data collection and ethical issues were described in detail elsewhere. Informed consent was obtained from all individual participants included in the Generation R study. The FFCWS study protocol was approved by the Institutional Review Board at Princeton University. Informed consent was obtained from all participating families, with parents or legal guardians consenting on behalf of minors, who also provided their assent. All data collection, storage, and analysis procedures were designed to protect participants' anonymity, and families were compensated for their participation.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s44220-024-00345-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44220-024-00345-8.

Correspondence and requests for materials should be addressed to Alexandre A. Lussier or Erin C. Dunn.

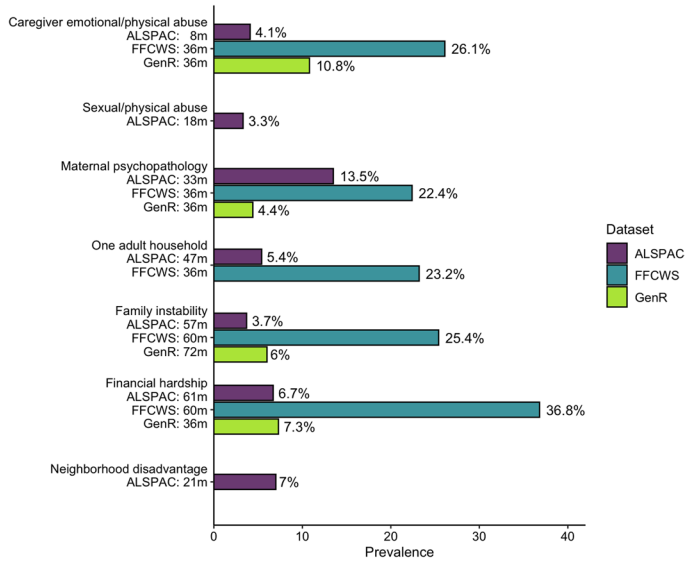
Peer review information *Nature Mental Health* thanks Carlos Alfonso Tovillo-Zarate and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

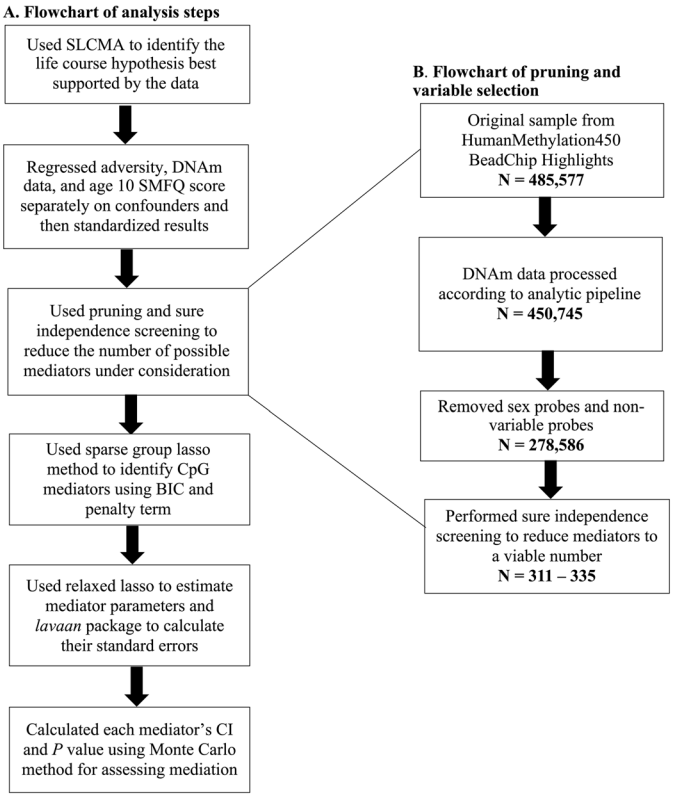
[1]Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA. [2]Department of Psychiatry, Harvard Medical School, Boston, MA, USA. [3]Stanley Center for Psychiatric Research, The Broad Institute of Harvard and MIT, Cambridge, MA, USA. [4]Institute for Social Research, University of Michigan, Ann Abor, MI, USA. [5]Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands. [6]Department of Psychology, Education and Child Studies, Erasmus MC, Erasmus University Rotterdam, Rotterdam, the Netherlands. [7]Department of Child and Adolescent Psychiatry/Psychology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands. [8]Department of Molecular Biology, Princeton University, Princeton, NJ, USA. [9]Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands. [10]McLean Hospital, Belmont, MA, USA. [11]Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA. [12]School of Mathematics, Statistics and Applied Mathematics, University of Galway, Galway, Ireland. [13]MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. [14]Department of Psychology, University of Bath, Bath, UK. [15]Mathematics and Statistics Research Group, University of the West of England, Bristol, UK. [16]Department of Sociology, College of Liberal Arts, Purdue University, West Lafayette, IN, USA. [17]These authors contributed equally: Alexandre A. Lussier, Brooke J. Smith. ✉e-mail: alussier@mgh.harvard.edu; erindunn@purdue.edu

https://doi.org/10.1038/s44220-024-00345-8



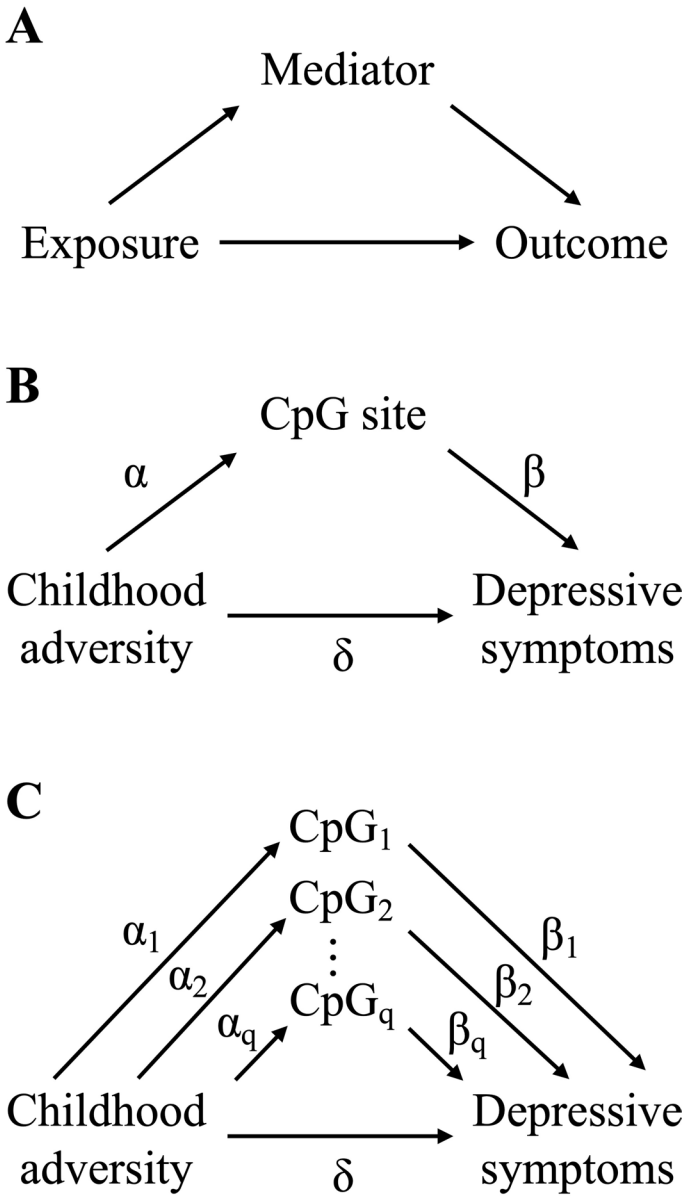
Extended Data Fig. 1 | Prevalence of exposures within each childhood adversity's analytic sample and across cohorts. Prevalence of adversity exposure within each analytic sample for the ALSPAC cohort ranged from 3.3% in sexual/physical abuse (by anyone) to 13.5% in maternal psychopathology. These numbers varied in the FFCWS and GenR cohorts, with higher prevalence observed in FFCWS across all adversities and slightly higher prevalence in GenR. The timepoints shown are those selected from the SLCMA in ALSPAC and the best matched timepoint in the replication cohorts.

https://doi.org/10.1038/s44220-024-00345-8



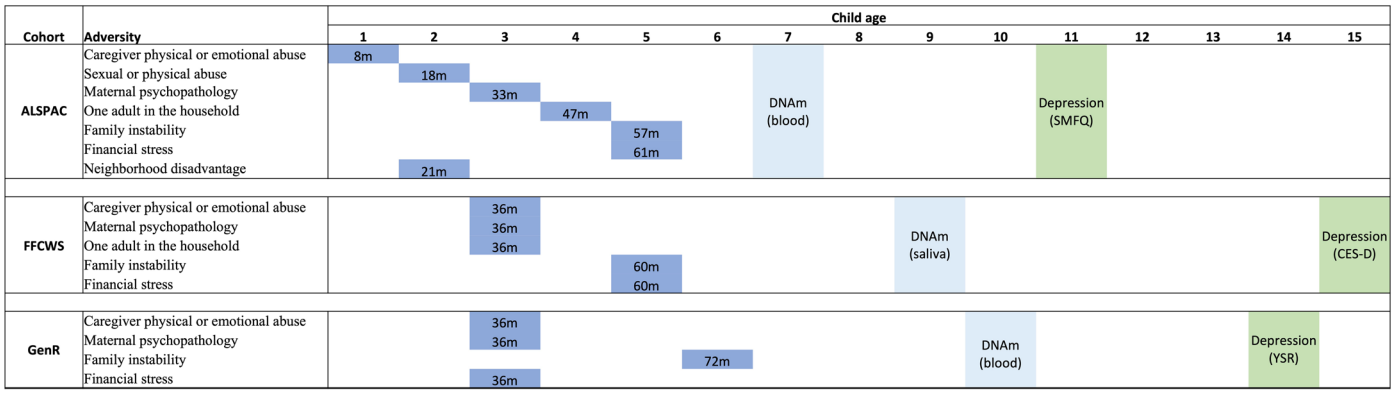
Extended Data Fig. 2 | Flowchart of analysis steps for each adversity-outcome relationship. SLCMA, structured life course modeling approach; DNAm, DNA methylation; SMFQ, Short Mood and Feelings Questionnaire; CpG, cytosines preceding a guanine nucleotide; BIC, Bayesian information criterion; CI, confidence interval. **A.** depicts the steps of the mediation analysis performed 7 times, once for each of the 7 adversities we studied (N = 627-675). **B.** highlights the steps taken to reduce the mediators under consideration to a viable number for each of the 7 adversities.

https://doi.org/10.1038/s44220-024-00345-8



Extended Data Fig. 3 | Single and multiple mediator structures. CpG = DNA region where a cytosine nucleotide is followed by a guanine nucleotide; α_i = effect estimate of childhood adversity on CpG_i; β_i = effect estimate of CpG_i DNA methylation on depressive symptoms; Directed acyclic graph (DAG) depicting the relationships between the study variables. **A.** Shows a general exposure-mediator-outcome relationship with a single mediator. **B.** Shows an exposure-mediator-outcome relationship within the context of our study. **C.** Shows a simplified version of our multiple mediator analysis where q represents the total number of mediators considered in the analysis after sure independence screening.

https://doi.org/10.1038/s44220-024-00345-8



Extended Data Fig. 4 | Summary of variables across ALSPAC, FFCWS, and GenR cohorts. Primary analyses in ALSPAC focused on exposures to adversity during sensitive periods (dark blue), DNA methylation (DNAm) at age 7, and depressive symptoms at age 10.6. Replication analyses in the FFCWS cohort focused on exposures to five adversities directly comparable to those in ALSPAC; comparability was based on both the type of adversity measure examined and sensitive period explored. DNAm was measured from saliva at age 9 and depressive symptoms were indexed at age 15. Replication analyses in the GenR cohort focused on exposures to four adversities directly comparable to those in ALSPAC, both in terms of adversity measurement and sensitive period explored. DNAm was measured in blood at age 10 and depression was assessed at age 13-14.

nature portfolio

Corresponding author(s): Alexandre Lussier

Last updated by author(s): Sep 3, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	The scripts used to complete the primary analyses in this manuscript are available on our github (https://github.com/thedunnlab/mediation). R packages used included: mediation, meffil, lavaan, coloc, lars, regmed.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data are available by request from the ALSPAC Executive Committee for researchers who meet the criteria for access to confidential data (bristol.ac.uk/alspac/researchers/access). The FFCWS data analyzed in the current study are available with permission from the Future of Families and Childhood Wellbeing Study repository (fragilefamilies.princeton.edu/documentation). Generation R Study data are not publicly available due to privacy or ethical restrictions. The study has an

open policy in regard to collaboration with other research groups. Requests for collaboration should primarily be addressed to Prof. Dr. Vincent Jaddoe (ln.cmsumsare@eoddaj.v). These requests are discussed in the Generation R Study Management Team regarding their study aims, overlap with ongoing studies, logistic consequences and financial contributions.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	We used sex assigned at birth (0=male, 1=female) as a covariate in our primary analyses. Gender information was not available for our analyses. We did not stratify by sex due to power considerations, which would have limited our ability to detect statistical mediation in our primary analyses.
Reporting on race, ethnicity, or other socially relevant groupings	Due to the low levels of diversity in the ALSPAC, we used a covariate to reflect white versus non-white participants. This variable was based on parent self-reported race/ethnic group at birth (0=non-White, 1=White).
Population characteristics	Childhood adversities were collected between ages 0 and 7 years. DNA methylation (DNAm) was measured at age 7 years. Depressive symptoms were measured at age 10.8 years. The study population was primarily white and of generally higher socio-economic position, estimated using maternal education (see Table S2 for details on the population characteristics).
Recruitment	Data came from the Avon Longitudinal Study of Parents and Children (ALSPAC), a 30-year ongoing study that recruited pregnant women in Avon, England with expected delivery dates between April 1991-December 1992. We analyzed data from the Accessible Resource for Integrated Epigenomic Studies (ARIES), a subset of 1,018 mother-child pairs in ALSPAC who had complete phenotype data from at least 5 waves of data collection.
Ethics oversight	Ethical approval for the study was obtained from the ALSPAC Law and Ethics Committee and the Local Research Ethics Committees. Consent for biological samples was collected in accordance with the Human Tissue Act of 2004. Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee. Secondary analyses of these data were approved with oversight by the Mass General Brigham Institutional Review Boards (IRB) (Protocol 2017P001110).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Our analytic samples for primary included singleton-birth children with complete data on adversity, depressive symptoms, DNA methylation, and all covariates (N=627-675). No sample size calculations were performed as we used all data available within the ALSPAC dataset. Replication analyses were performed in two additional datasets, where complete cases were used once again.
Data exclusions	No data exclusions were present beyond using complete case analysis.
Replication	We sought to replicate our findings in two independent birth cohorts: 1) the Future of Families and Child Wellbeing Study (FFCWS; n=1,319-1,677); and 2) Generation R Study (GenR; n=312-392). We performed CpG-specific mediation analyses using the mediation package in R for each adversity, matched to the timing and type of adversity measured in ALSPAC. These cohorts differed from ALSPAC in multiple respects, including generally higher prevalence of adversity (Figure S1), differences in depressive symptom measures, sensitive periods identified in the SLCMA (Table S5), and sociodemographic characteristics such as race/ethnicity (details in Supplementary Materials). Because p-values may be unreliable within and across epigenome-wide studies, we used the magnitude and direction of the indirect (mediated) effects as metrics for replication, focusing primarily on the risk-increasing or protective effects of DNAm on depressive symptoms. Overall, we largely replicated the role of DNAm in the relationship between adversity and depressive symptoms when combining the indirect effects of all CpGs for each adversity.Similar to ALSPAC, we identified both risk-increasing and protective effects of DNAm across every adversity (caregiver abuse, maternal psychopathology, family instability, and financial hardship) in both FFCWS and GenR.
Randomization	Participants were not randomly allocated into experimental groups, as all data were obtained from population-wide studies. As allocation was not random, we controlled for the following sociodemographic characteristics previously linked to adversity, depression, and differences in DNAm: child sex, birthweight, race and/or ethnicity, maternal age at birth, maternal education at birth, number of previous pregnancies, and maternal smoking during pregnancy. To control for confounding and to help sustain our statistical power, we separately regressed the life course hypothesis selected for each adversity, DNAm data, and age 10.6 SMFQ score on the previously mentioned covariates. We took the residuals from these regressions and standardized them before using them in our mediation analysis. This approach ensured our results could be interpreted in terms of correlation between childhood adversity and depressive symptoms after adjusting for confounders.

Blinding	Blinding was not relevant to our study as we did not complete the data collection and analyses were performed on a population-level with directly examining the main predictors (exposed/unexposed to childhood adversity).

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA