



## Research paper

# Developing an individualized risk calculator for psychopathology among young people victimized during childhood: A population-representative cohort study



Alan J. Meehan<sup>a</sup>, Rachel M. Latham<sup>a</sup>, Louise Arseneault<sup>a</sup>, Daniel Stahl<sup>b</sup>, Helen L. Fisher<sup>a,1</sup>, Andrea Danese<sup>a,c,d,1,\*</sup>

<sup>a</sup> Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

<sup>b</sup> Department of Biostatistics & Health Informatics, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

<sup>c</sup> Department of Child & Adolescent Psychiatry, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

<sup>d</sup> National and Specialist CAMHS Trauma, Anxiety, and Depression Clinic, South London and Maudsley NHS Foundation Trust, London, UK

## ARTICLE INFO

## Keywords:

Victimization  
Psychopathology  
Risk prediction  
Risk calculator  
Resilience

## ABSTRACT

**Background:** Victimized children are at greater risk for psychopathology than non-victimized peers. However, not all victimized children develop psychiatric disorders, and accurately identifying which victimized children are at greatest risk for psychopathology is important to provide targeted interventions. This study sought to develop and internally validate individualized risk prediction models for psychopathology among victimized children.

**Methods:** Participants were members of the Environmental Risk (E-Risk) Longitudinal Twin Study, a nationally-representative British birth cohort of 2,232 twins born in 1994–1995. Victimization exposure was measured prospectively between ages 5 and 12 years, alongside a comprehensive range of individual-, family-, and community-level predictors of psychopathology. Structured psychiatric interviews took place at age-18 assessment. Logistic regression models were estimated with Least Absolute Shrinkage and Selection Operator (LASSO) regularization to avoid over-fitting to the current sample, and internally validated using 10-fold nested cross-validation.

**Results:** 26.5% ( $n = 591$ ) of E-Risk participants had been exposed to at least one form of severe childhood victimization, and 60.4% ( $n = 334$ ) of victimized children met diagnostic criteria for any psychiatric disorder at age 18. Separate prediction models for any psychiatric disorder, internalizing disorders, and externalizing disorders selected parsimonious subsets of predictors. The three internally validated models showed adequate discrimination, based on area-under-the-curve estimates (range = 0.66–0.73), and good calibration.

**Limitations:** External validation in wholly-independent data is needed before clinical implementation.

**Conclusions:** Findings offer proof-of-principle evidence that prediction modeling can be useful in supporting identification of victimized children at greatest risk for psychopathology. This has the potential to inform targeted interventions and rational resource allocation.

## 1. Introduction

Childhood victimization is associated with a range of internalizing (Nanni et al., 2012; Takizawa et al., 2014), externalizing (Braga et al., 2017; Capusan et al., 2016), and psychotic disorders (Fisher et al., 2013), likely reflecting a general vulnerability for psychopathology in victimized children (Schaefer et al., 2018). Yet, not all victimized

children develop psychopathology (Rutter, 2013). Accurately identifying which victimized children are at greatest risk for psychopathology is therefore important, in order to provide targeted support and to inform rational allocation of resources. Previous studies have described individual-, family-, and community-level factors that increase or decrease risk of psychopathology among victimized children (i.e., vulnerability or resilience factors, respectively), and thus, may improve

\* Corresponding author at: Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, 16 De Crespigny Park, London SE5 8AF, UK.

E-mail address: [andrea.danese@kcl.ac.uk](mailto:andrea.danese@kcl.ac.uk) (A. Danese).

<sup>1</sup> These authors contributed equally to this work.

<https://doi.org/10.1016/j.jad.2019.10.034>

Received 28 May 2019; Received in revised form 23 September 2019; Accepted 25 October 2019

Available online 05 November 2019

0165-0327/ © 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

risk detection (Fritz et al., 2018; Meng et al., 2018). However, these studies also typically had important limitations: they were based on cross-sectional or short-term longitudinal designs, making it difficult to ascertain whether victimization and potential resilience/vulnerability factors preceded the onset of psychopathology (Cicchetti, 2013; Meng et al., 2018); they described resilience/vulnerability factors in isolation, likely over-estimating their individual contribution compared to more realistic multivariate models (Fritz et al., 2018); and they identified factors that modify the average risk for psychopathology for subgroups of victimized children with shared features, but did not test if such factors can accurately predict which victimized children do or do not develop psychopathology (Yarkoni and Westfall, 2017). As such, the extent to which established resilience/vulnerability factors can inform individualized risk prediction is unclear.

To address these gaps, we investigated psychiatric risk prediction among members of a nationally-representative prospective British cohort. Building on prediction modeling methods developed for medical conditions (Damen et al., 2016; Hippisley-Cox and Coupland, 2015) and adult psychiatric disorders (Cannon et al., 2016; Fusar-Poli et al., 2017; Hafeman et al., 2017), we developed multivariate individualized risk prediction models (risk calculators) for psychopathology among victimized children (Moons et al., 2009; Steyerberg and Vergouwe, 2014). We also internally validated our findings using a nested cross-validation approach to determine the classification accuracy of prediction models when applied to independent cases in our sample (Steyerberg, 2009).

## 2. Methods

### 2.1. Participants

Participants were members of the Environmental Risk (E-Risk) Longitudinal Twin Study, which tracks the development of a nationally-representative birth cohort of 2232 twin children born in England and Wales in 1994–1995. Full details about the sample are reported elsewhere (Moffitt and E-Risk Study Team, 2002) and in Supplementary Material. Briefly, the E-Risk sample was constructed in 1999–2000 when 1116 families (93% of those eligible) with same-sex 5-year-old twins participated in home-visit assessments. This sample comprised 56% monozygotic and 44% dizygotic twin pairs; sex was evenly distributed within zygosity (49% male). Families were recruited to represent the UK population of families with newborns in the 1990s, on the basis of residential location throughout England and Wales and mother's age.

Follow-up home-visits were conducted when children were aged 7, 10, 12, and 18 (participation rates were 98%, 96%, 96%, and 93%, respectively). At age 18, 2066 participants were assessed. Average age at time of assessment was 18.4 years ( $SD = 0.36$ ); all interviews were conducted after the 18th birthday. There were no differences between those who did and did not take part at age 18 in terms of socioeconomic status, assessed when the cohort was initially defined ( $\chi^2 = 0.86$ ,  $p = .65$ ), age-5 IQ ( $t = 0.98$ ,  $p = .33$ ), age-5 behavioral ( $t = 0.40$ ,  $p = .69$ ) or emotional ( $t = 0.41$ ,  $p = .68$ ) problems, or childhood poly-victimization ( $z = 0.51$ ,  $p = .61$ ). The Joint South London and Maudsley and Institute of Psychiatry Research Ethics Committee approved each study phase. Parents gave informed consent, and twins gave assent between 5 and 12 years and then informed consent at age 18.

The present study focuses on a subset of the E-Risk Study sample who were exposed to any type of severe victimization during childhood ( $n = 591$ ; 26.5% of sample; 50% male). Information about the victimization measure used to derive this analytic sample is provided below and in Supplementary Material. Based on multivariate logistic regression and odds ratios (ORs), victimized and non-victimized participants did not differ on the distribution of sex (OR = 0.92, 95% CI = 0.76–1.12,  $p = .40$ ) and ethnicity (White vs non-White;

OR = 1.00, 95% CI = 0.72–1.39,  $p = .99$ ); however, victimized young people were more likely to be in the lowest social class tertile compared to their non-victimized counterparts (OR = 2.91, 95% CI = 2.40–3.54,  $p < 0.001$ ).

### 2.2. Measures

#### 2.2.1. Childhood victimization

Prospective measures of victimization utilized in this cohort are described elsewhere (Danese et al., 2017; Fisher et al., 2015), and in Supplementary Material. In brief, lifetime exposure to several types of victimization was assessed repeatedly when children were 5, 7, 10, and 12 years of age. Comprehensive dossiers were compiled for each child with cumulative information about: exposure to domestic violence between mother and partner; frequent bullying by peers; physical abuse by an adult; sexual abuse; emotional abuse and neglect; and physical neglect, all between birth and age 12 years. Dossiers comprised reports from caregivers, recorded narratives of caregiver interviews, recorded debriefings with research workers who had coded any indications of abuse and neglect at any of the successive home visits, and information from clinicians whenever the study team made a child-protection referral. These were reviewed by two independent researchers and rated for the presence and severity (none/mild/severe) of each type of victimization. In the present study, prospectively-measured victimization was dichotomized to represent 'none/mild' (0) vs 'severe' (1) victimization.

#### 2.2.2. Age-18 psychopathology

Past-year psychopathological symptoms were comprehensively assessed through private interviews available at age-18 follow-up. Ten disorder diagnoses were organized into three domains (internalizing, externalizing, and thought disorders), based on a reliable latent factor structure for psychopathology previously identified within the E-Risk Study (Schaefer et al., 2018). Full information on individual diagnoses is available in Supplementary Material. In brief, in order to effectively evaluate the classification accuracy of each prediction model, binary classifications were derived for each domain, denoting the presence of at least one of the constituent disorders based on diagnostic cut-offs. Participants were classified as having 'internalizing disorder' if they met diagnostic criteria for generalized anxiety disorder, major depressive disorder or post-traumatic stress disorder, or presented at least 2 of 5 eating disorder symptoms from an established screening tool, indicating a possible case of anorexia nervosa or bulimia nervosa (Morgan et al., 1999). Participants were classified as having 'externalizing disorder' where they met diagnostic criteria for attention-deficit/hyperactivity disorder, conduct disorder, alcohol dependence, cannabis dependence, or tobacco dependence. Finally, 'thought disorder' classification was based on the definite presence of at least one of seven psychotic symptoms, centered on delusions and hallucinations. The measure has good construct validity, as it was shown to share many of genetic, social, neurodevelopmental, and behavioral risk factors and correlates as adult schizophrenia (Polanczyk et al., 2010). From these three domain-specific classifications, an overall binary outcome for 'any psychiatric disorder' was created, denoting the presence of any internalizing, externalizing, or thought disorder (1), or the absence of all three (0).

The prevalence of each disorder and wider diagnostic domain, within both the overall E-Risk sample and our victimized sub-sample, is presented in Table 1. Overall, out of 553 victimized participants, 60.4% met diagnostic criteria for any of the ten psychiatric disorders, 37.9% met criteria for internalizing disorder, and 43.3% met criteria for externalizing disorder. Regarding comorbidity between internalizing and externalizing domains, among those diagnosed with at least one of their nine underlying disorders, 36.4% ( $n = 118$ ) had both an internalizing and externalizing disorder, 27.2% ( $n = 88$ ) had an internalizing disorder alone, and 36.4% ( $n = 118$ ) had an externalizing disorder alone.

**Table 1**

Prevalence of psychopathology at age 18 among (i) the main E-Risk sample, and (ii) a subsample of E-Risk participants exposed to severe childhood victimization.

Diagnosis	Main Sample (N = 2050–2066)		Victimized Sample (N = 551–558)	
	N	%	N	%
<b>Any psychiatric disorder</b>	973	47.5	334	60.4
<b>Internalizing disorder</b>	585	28.5	209	37.9
Generalized anxiety disorder	153	7.4	63	11.3
Major depressive disorder	414	20.1	163	29.2
Post-traumatic stress disorder	90	4.4	41	7.4
Eating disorder	204	9.9	71	12.9
<b>Externalizing disorder</b>	656	31.9	239	43.3
Attention-deficit/hyperactivity disorder	171	8.3	64	11.5
Conduct disorder	309	15.1	119	21.5
Alcohol dependence	263	12.8	90	16.1
Cannabis dependence	89	4.3	35	6.3
Tobacco dependence	183	8.9	85	15.3
<b>Thought disorder</b>	59	2.9	29	5.2

### 2.2.3. Childhood predictors

Individual-, family-, and community-level predictors were assessed between ages 5 and 12 years. We utilized a recent systematic review of multi-level predictors of maltreatment outcomes (Meng et al., 2018) and mapped these predictors to variables measured in the E-Risk Study (see Supplementary Tables S1 and S2). This approach is preferable to selecting predictors based on statistically-significant bivariate associations with the outcome in the target sample, as it avoids the circular logic of testing the predictive ability of variables already known to be associated with the outcome within that sample (Fusar-Poli et al., 2018). We identified 22 predictors, summarized in Table 2, with detailed descriptions in the Supplementary Material. With 22 predictors,

**Table 2**

Summary of childhood predictors.

Measure	Age	Informant	Description
<b>Individual</b>			
Sex	5	Mother	1 = Male; 2 = Female
IQ	12	Child	Pro-rated WISC-R score (Matrix Reasoning and Information subtests)
Openness to experience	12	Researcher	Sum of 5 BFI items
Conscientiousness			Sum of 6 BFI items
Extraversion			Sum of 6 BFI items
Agreeableness			Sum of 5 BFI items
Neuroticism			Sum of 5 BFI items
ADHD symptoms	12	Mother/teacher	18 DSM-IV inattentive and impulsive-hyperactive symptoms from CBCL (mothers) and TRF items (teachers), averaged across raters
Conduct disorder symptoms	12	Mother/teacher	14 DSM-IV criteria from CBCL (mothers) and TRF items (teachers), averaged across raters
Anxiety	12	Child	Symptom score from 10 MASC items
Depression	12	Child	Symptom score from 27 CDI items
Self-harm/suicide attempt	12	Mother	Any deliberate self-harm or attempted suicide in previous six months
Psychotic symptoms	12	Child	Presence of at least one definite psychotic symptom
<b>Family</b>			
Maternal warmth	5; 10	Mother	Warmth, enthusiasm, interest in, enjoyment of child during FMSS, summed across time-points
Sibling warmth	7; 10	Mother	6 items each, summed across time-points
Adult involvement	12	Child	13 items assessing presence of supportive adult
Family history of psychopathology	12	Mother	Proportion (0.0–1.0) of family members (parents, grandparents, aunts/uncles) with history of psychiatric disorder
Biological parents in household	10	Mother	Number of biological parents in household up to age 10 (0 = both always absent; 1 = one absent at some point; 2 = both always present)
Socioeconomic status	5	Mother	Tertiles derived from standardized composite of parental income, education, and occupation
<b>Community</b>			
Neighborhood crime victimization	5	Mother	3 items assessing family's experience of violent crime, burglary, or theft in local area
Social cohesion	5	Mother	Sum of 5 items (neighbors close-knit, share values, trust each other, etc.)
Status among peers	12	Child	Self-selected position within 5 peer status 'circles'

**Notes.** For full details of measures with accompanying references, see Supplementary Material. Child's age given in years. ADHD = attention-deficit/hyperactivity disorder; BFI = Big Five Inventory; CBCL = Child Behavior Checklist; CDI = Children's Depression Inventory; DSM-IV = Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; FMSS = Five-Minute Speech Sample; MASC = Multidimensional Anxiety Scale for Children; TRF = Teacher's Report Form; WISC-R = Wechsler Intelligence Scale for Children-Revised.

we achieved a minimum events per variable (EPV) ratio of 10 for our main 'any psychiatric disorder' outcome, mitigating potential model instability due to over-fitting (Pavlou et al., 2015).

### 2.3. Statistical analyses

Analyses were conducted using STATA version 15.0 and R version 3.4.2. First, we tested whether the prevalence of each form of psychopathology significantly differed between victimized and non-victimized participants using logistic regression, correcting for familial clustering using the *cluster* STATA command.

Second, we developed and internally validated separate prediction models for three age-18 psychiatric outcomes using regularized logistic regression in the *glmnet* R package (Friedman et al., 2010). Specifically, we evaluated predictive ability for the overall measure of 'any psychiatric disorder', as well as separately for internalizing and externalizing disorders, as the two dimensions do show distinct features and risk profiles, even over and above a general factor for psychopathology (Lahey et al., 2017). We could not reliably estimate a separate model for thought disorder given the low prevalence of psychotic symptoms. Complete data were available in 91.3%, 91.5% and 91.3% of victimized children with available data for 'any psychiatric', 'externalizing' and 'internalizing' disorders, respectively ( $n = 504$ – $505$ ). Therefore, a complete-cases approach was utilized for model development and internal validation.

A full description of our analyses is available in Supplementary Material. Briefly, we conducted regularized regression using the Least Absolute Shrinkage and Selection Operator (LASSO) to identify subsets of predictors that maximized prediction accuracy for each outcome in unseen cases within the sample. LASSO shrinks coefficients towards zero, thereby reducing the variance of these estimates. Predictors whose coefficients are shrunk to zero are excluded from the model, enabling parsimonious solutions. The degree of shrinkage is determined by a tuning parameter, lambda ( $\lambda$ ). Cross-validation was used to identify the optimal  $\lambda$  for a model (Tibshirani, 1996). To obtain an estimate of

prediction accuracy in new cases from the same underlying population, we internally validated each model using nested 10-fold cross-validation (see Supplementary Fig. S2; Hastie et al., 2009).

Performance measures were based on each child's predicted probability when held out from model selection and estimation as an unseen 'test case'. Discrimination (the model's ability to accurately classify those with and without psychopathology) was visualized using the receiver operator characteristic (ROC) curve, and quantified using the area-under-the-curve (AUC; Steyerberg, 2009). AUC ranges from 0.5 (chance-level) to 1 (perfect discrimination), with the following proposed benchmarks: <0.7 (poor); 0.7–0.8 (acceptable); 0.8–0.9 (excellent); 0.9–1.0 (outstanding; Hosmer et al., 2013). Calibration was assessed by plotting predicted probabilities (grouped into equal-interval bins) against observed outcomes, where a 45° line indicates perfect agreement. A chi-square test of unreliability (*U*) determined whether each calibration plot's intercept ('calibration-in-the-large') and slope significantly differed from the ideal line (Fenlon et al., 2018). Overall performance was assessed by the proportion of explained deviance ('pseudo- $R^2$ '), and the Brier score, or mean squared difference between predicted probabilities and actual binary outcomes (Brier, 1950). We scaled Brier scores by their maximum possible values, which vary based on the incidence of the outcome, from 0% (non-informative) to 100% (perfect; Steyerberg, 2009).

Finally, we carried out two sensitivity analyses. First, we tested whether model performance and/or interpretability would improve if we allowed inclusion of a set of correlated data in our model. To do this, we re-estimated each model using elastic net regularization. Unlike LASSO, this regularization method selects or excludes sets of correlated variables, potentially allowing more predictors to be retained. Second, to test whether predictive ability was inflated by the presence of non-independent twin observations, we re-ran analyses in ten subsamples, each consisting of one twin per pair.

### 3. Results

#### 3.1. Descriptive statistics

Victimized children were significantly more likely than their non-victimized peers to present with any psychiatric (OR = 2.05, 95% CI = 1.63–2.57,  $p < .001$ ), internalizing (OR = 1.82, 95% CI = 1.44–2.30,  $p < .001$ ), externalizing (OR = 1.99, 95% CI = 1.58–2.49,  $p < .001$ ), or thought disorder at age 18 (OR = 2.70, 95% CI = 1.57–4.62,  $p < .001$ ). However, a substantial proportion of victimized children did not have any psychiatric (39.6%), internalizing (62.1%), externalizing (56.7%), or thought (94.8%) disorder at age 18 (see Supplementary Fig. S1). Therefore, we developed prediction models to accurately identify individualized risk for psychopathology among victimized children.

#### 3.2. Prediction modeling

Unstandardized coefficients for the predictors retained in each of the three models following LASSO variable selection (see Table 3) were obtained by re-estimating the cross-validated model using the average optimal  $\lambda$  tuning parameter, identified from ten repeated model development loops. Using each set of coefficients, we derived regression equations to estimate a child's individual risk for an outcome based on his/her values on these retained predictors (see Supplementary Table S3). Additionally, the classification accuracy achieved by each model across various thresholds of predicted risk are presented in Supplementary Tables S4–S6. Model composition and performance are described below for each outcome.

##### 3.2.1. Any psychiatric disorder

17 of 22 predictors were retained in the final LASSO model (see Table 3, column A), based on an average  $\lambda$  tuning parameter of 0.0109

**Table 3**  
Unstandardized coefficients for prediction models.

Predictor	A	B	C
	Any psychiatric disorder ( <i>n</i> = 505) <i>B</i>	Internalizing disorder ( <i>n</i> = 504) <i>B</i>	Externalizing disorder ( <i>n</i> = 505) <i>B</i>
Intercept	2.186	−1.643	2.558
<b>Individual</b>			
Sex (female)	−0.291	.329	−0.811
IQ	−0.001	−	−0.008
Openness to experience	.038	.037	−
Conscientiousness	−0.018	−0.018	−
Extraversion	.107	.090	.048
Agreeableness	−0.166	−0.059	−
Neuroticism	−	.088	−
ADHD symptoms	−	−	.005
CD symptoms	.250	.063	.267
Anxiety symptoms	.029	.057	−
Depression symptoms	.002	.005	.005
Self-harm/suicide attempt	−	−0.416	−
Psychotic symptoms	1.044	1.112	.551
<b>Family</b>			
Maternal warmth	.033	.101	−
Sibling warmth	−0.027	−	−0.051
Adult involvement	−0.011	−0.031	−0.019
Family history of psychopathology	.323	.709	.038
Biological parents in household	−0.310	−0.106	−0.317
Socioeconomic status	−0.164	−0.194	−0.004
<b>Community</b>			
Neighborhood crime victimization	−	−0.060	.070
Social cohesion	−	−	−0.023
Status among peers	−0.047	−0.091	.067
Deviance explained (Pseudo- $R^2$ )	13.5%	11.1%	15.4%

Notes. For sex, positive values imply a greater number of females with that outcome, while negative values suggest greater prevalence among males. Regression equations for individual risk calculation, derived from these coefficients, are presented in Supplementary Table S3. B = unstandardized regression coefficient; ADHD = attention-deficit/hyperactivity disorder; CD = conduct disorder.

(range: 0.0075–0.0151). Frequency distributions of predicted risks for victimized children with and without psychopathology at age 18, obtained using nested 10-fold cross-validation, are presented in Fig. 1a. This showed that, beginning at a predicted risk of 0.70, there was a higher proportion of victimized children with versus without any psychiatric disorder in each successive risk class. The ROC curve following internal validation yielded an AUC of 0.69 (95% CI = 0.64–.73; Fig. 2, solid line), representing a 69% probability that a randomly-selected victimized child with any psychiatric disorder will be classified as such by the model compared to a victimized child without psychopathology. The calibration plot (Fig. 3a) showed high agreement between predicted probabilities and observations when the predicted probability was 0.30 and higher; only two participants, or 0.4% of all cases, fell below this point, which likely explains the poorer calibration at this end of the distribution. Calibration-in-the-large (i.e., intercept;  $\alpha = 0.02$ ) and slope ( $\beta = 0.96$ ) did not significantly differ from expected values of a perfectly-calibrated model ( $U = -0.004$ ,  $\chi^2(2) = 0.06$ ,  $p = .968$ ). Regarding overall performance, the model explained 13.5% of binomial deviance, while the scaled Brier score indicated that the predictors explained 11% of the mean squared error between predicted probabilities and observed outcomes compared to a non-informative model.

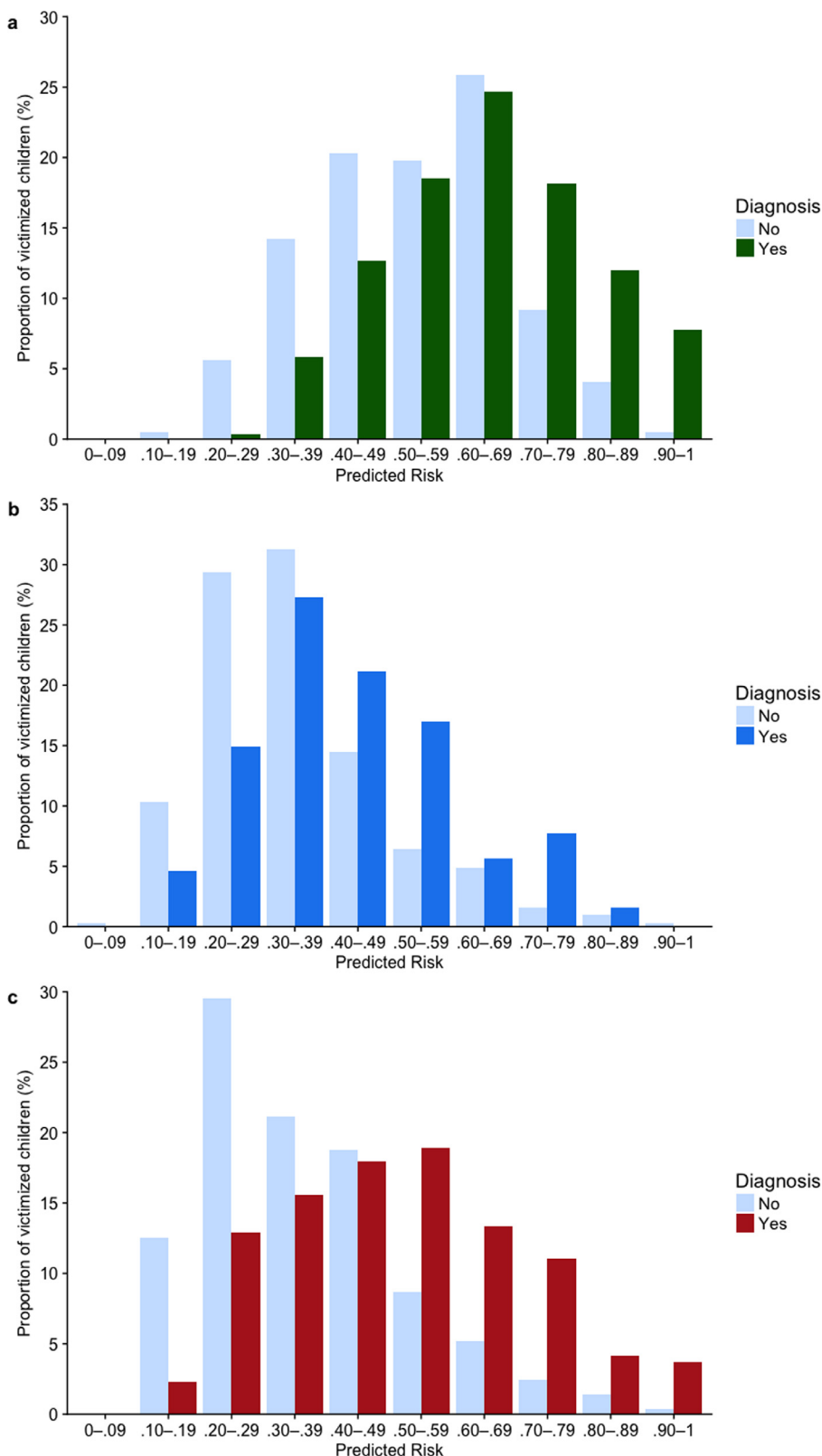
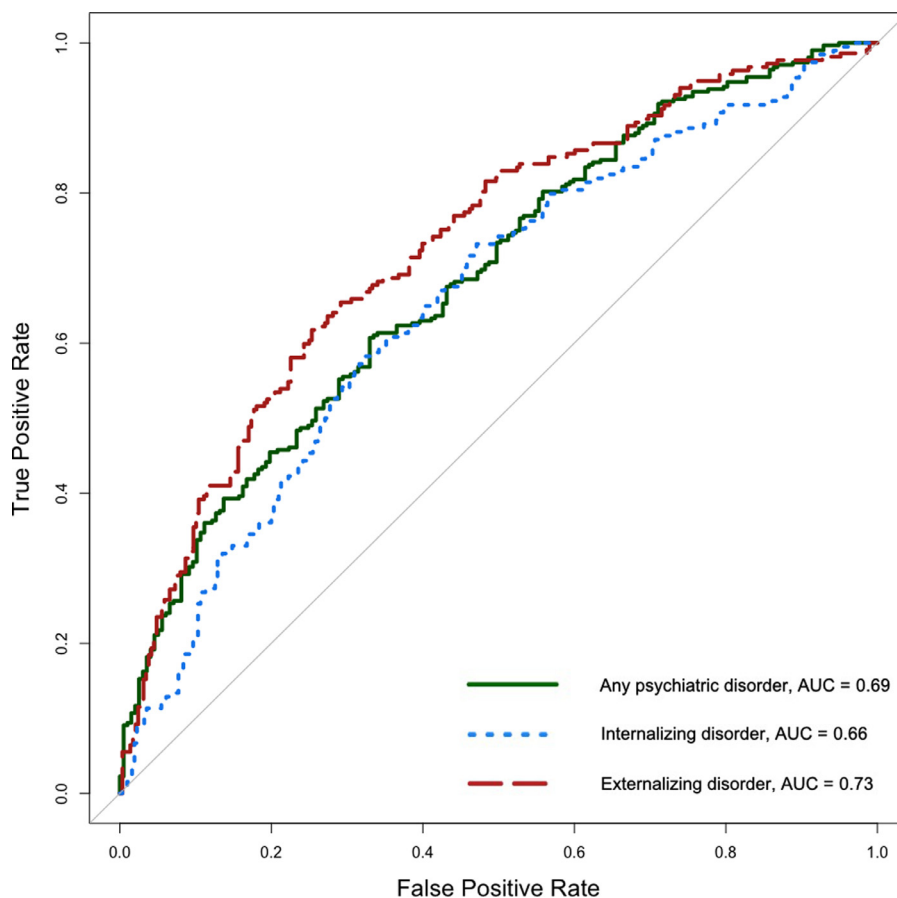


Fig. 1. Frequency distributions of predicted risk among victimized children with and without (a) any psychiatric; (b) internalizing; and (c) externalizing disorder.

3.2.2. Internalizing disorder

18 predictors were retained here (see Table 3, column B), based on a mean  $\lambda$  of 0.0089 (range: 0.0057–0.0118). There was a higher proportion of victimized children with versus without internalizing disorder between predicted risk thresholds of 0.50 and 0.89 (see Fig. 1b).

The internally-validated model yielded an AUC of 0.66 (95% CI = 0.61–.71; Fig. 2, dotted line). Again, calibration-in-the large ( $\alpha = -0.10$ ) and slope values ( $\beta = 0.75$ ) did not diverge significantly from a perfectly-calibrated model ( $U = 0.002, \chi^2(2) = 3.08, p = .214$ ; see Fig. 3b). Indeed, the model was well-calibrated between



**Fig. 2.** Receiver operating characteristic (ROC) curves for the three prediction models among victimized children. ROC curves plot the true positive rate (sensitivity; the proportion of actual positive outcomes correctly identified as such) against the false positive rate (1–specificity; the proportion of incorrectly-classified positive outcomes). The solid diagonal line denotes chance-level discrimination.

predicted probabilities of 25% and 60%, within which 89% of cases fell (median = 0.35; interquartile range = 0.27–0.47). The model explained 11.1% of binomial deviance and, based on the scaled Brier score, accounted for 5.8% of mean squared prediction error over a non-informative model.

### 3.2.3. Externalizing disorder

LASSO regularization selected 15 predictors (see Table 3, column C) from a mean  $\lambda$  of 0.0146 (range: 0.0105–0.0178). Based on frequency distributions for predicted risk, there was a higher proportion of victimized children with versus without externalizing disorder from a predicted risk of 0.50 and upwards (see Fig. 1c). Discrimination was denoted by an AUC of 0.73 (95% CI = 0.69–.77; Fig. 2, dashed line), while calibration-in-the-large ( $\alpha = -0.01$ ) and slope ( $\beta = 1$ ) values did not diverge significantly from those of a perfectly-calibrated model ( $U = -0.004$ ,  $\chi^2(2) = 0.003$ ,  $p = .999$ ; see Fig. 3c). Overall, the model explained 15.4% of the deviance, with the scaled Brier score indicating a 15.3% reduction in mean squared error compared to a non-informative model.

## 3.3. Sensitivity analyses

### 3.3.1. Regularization penalty

For all three outcomes, model performance statistics (see Supplementary Table S7) and corresponding plots (see Supplementary Figs S3 and S4) obtained from nested 10-fold cross-validation using less restrictive ‘elastic net’ regularization resembled those generated using LASSO, implying that predictive ability was not significantly worsened by applying the most parsimonious form of regularization to model coefficients.

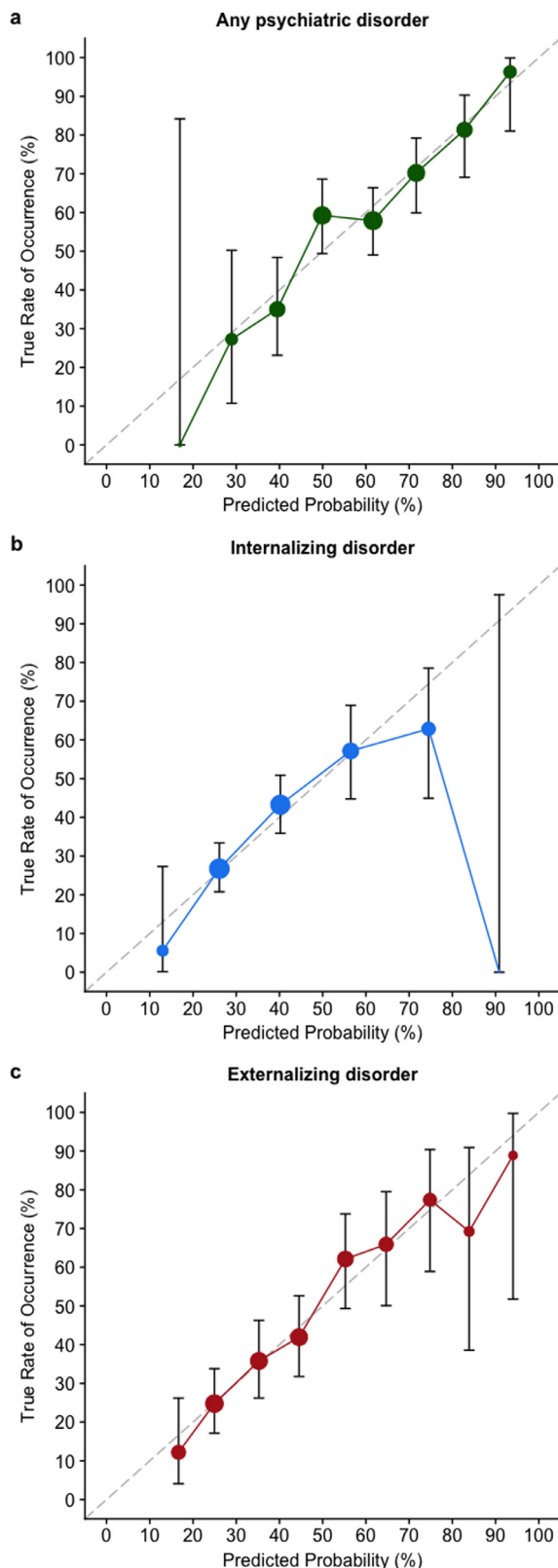
### 3.3.2. Non-independence of twins

Average cross-validated performance statistics for each outcome across 10 random single-twin subsamples ( $n = 304$ –305) resembled those obtained using the full sample (see Supplementary Tables S8–S10), suggesting that the inclusion of non-independent observations did not significantly bias predictive accuracy.

## 4. Discussion

We found that factors known to be associated with psychopathology among victimized children can be statistically modelled to predict individual risk for psychopathology during the transition to adulthood. All three models were well-calibrated, and discrimination, while only adequate (AUC = 0.66–0.73), was within the range of established risk calculators for medical conditions (Damen et al., 2016; Wang et al., 2018), as well as recently-developed prediction models in psychiatry (Cannon et al., 2016; Fusar-Poli et al., 2017; Hafeman et al., 2017). Moreover, because there is currently no evidence base for evaluating individual risk among victimized children, these data-driven approaches are likely to improve on current practice owing to their objectivity and consistency. Our results also suggest a need to consider prediction of internalizing and externalizing psychopathology separately, given observed variations in predictive performance and model configuration; for example, while all five personality dimensions were retained in the prediction of internalizing disorder, only extraversion contributed to the prediction of externalizing disorder.

Our findings should be interpreted in light of several limitations. First, as the E-Risk Study comprises twins, we cannot be certain that results generalize to singletons. However, the prevalence of victimization in our sample is comparable to recent UK general population estimates (Radford et al., 2013), while rates of psychopathology are generally comparable between twins and singletons (Gjone and



**Fig. 3.** Calibration plots for (a) any psychiatric, (b) internalizing, and (c) externalizing disorder in victimized children, showing overall agreement between model-predicted risks and observed outcomes. The broken diagonal line represents a perfectly-calibrated model. Error bars represent 95% confidence intervals. Each data-point denotes a probability bin, while its size indicates the relative number of cases within that bin.

Novik, 1995). Sensitivity analyses using one twin from each family also suggested that the presence of twins did not inflate predictive accuracy estimates. Second, as findings were based on a community sample of British twins, they may not generalize to other samples, for example children in foster care. External validation is needed to test the relevance of our findings to different contexts before clinical implementation. Third, unlike conventional regression approaches, our prediction models do not permit interpretation of individual coefficients. Specifically, in applying a penalty to reduce over-fitting to the data, LASSO regularization introduces bias into the regression estimates, such that coefficients are no longer reflective of true population-level associations with the outcome. Moreover, LASSO models perform inbuilt variable selection based around the ability of a risk factor to accurately predict variation in the outcome. In this way, we maximize predictive performance in unseen cases but limit model interpretability, as coefficients cannot be used to infer a causal relationship with the outcome; indeed, many variables that are important for prediction may only exert small effects when evaluated based on statistical significance or effect size (Goldstein et al., 2017; Yarkoni and Westfall, 2017). Instead, these coefficients should be viewed as components of the multifactorial risk profile represented by each model's specific combination of predictors.

Despite these limitations, the paper makes important methodological contributions to promote use of prediction modeling techniques in the context of childhood victimization. Compared to previous research, our prospective design ensured appropriate temporal ordering of effects from birth to early adulthood, the developmental point by which most psychiatric disorders have emerged (Kim-Cohen et al., 2003), while a rich characterization of individual-, family-, and community-level characteristics within our sample allowed for comprehensive multivariate models that incorporated potential determinants of both risk and resilience. Moving beyond the group-based average effects estimated in conventional regression analyses, our prediction modeling approach yielded an individualized risk estimate for each victimized child, quantifying their unique likelihood of developing psychopathology. Moreover, our internal validation procedure (i.e., nested cross-validation) ensured that each participant's risk score was calculated when they were excluded from model development, thus reducing over-fitting and increasing the generalizability of our models to unseen cases (Steyerberg, 2009). Finally, automatic variable selection by LASSO regularization suggested that predictors with established univariate-significant associations may not offer an independent contribution to predictive accuracy when considered in a more systemic context.

Overall, these initial findings provide proof-of-principle evidence that a range of individual, family, and community factors can be combined to derive individualized risk scores for psychopathology among victimized young people. With sufficient external validation, prediction modeling has the potential to enhance evidence-based clinical decision-making for this vulnerable population in social work and child and adolescent psychiatry settings and, in turn, inform more rational allocation of limited resources. However, although such risk calculators could introduce greater objectivity and consistency to clinical practice, they should be used to support, rather than replace, professional judgments. Crucially, good predictive performance, even in external samples, does not guarantee clinical usefulness. For example, evaluating overall predictive accuracy across a range of decision thresholds for absolute risk classification (i.e., 'low-risk' versus 'high-risk'), we found that cut-points of 60%, 40%, and 40% provided an optimal balance of sensitivity and specificity in predicting any psychiatric disorder, internalizing disorder, and externalizing disorder, respectively (see Supplementary Tables S4–S6). However, the appropriate risk threshold is likely to vary depending on the cost-benefit ratio of the corresponding clinical decision. For example, clinicians could opt for a lower cut-off during initial screening, where the benefit of early detection of true-positive cases may outweigh the cost of higher false-

positives, but may prefer a higher threshold for decisions around treatment allocation to ensure that places are allocated to those with the greatest need. Furthermore, classification is a forced choice, and individual predicted probabilities of risk should be presented to both clinicians and patients for an informed choice when all information is available. A decision should not simply rely on a prediction model based only on data included in the prediction model. Therefore, the clinical context, net-benefit and cost-effectiveness of any risk calculator should be evaluated in consultation with key stakeholders (Steyerberg and Vergouwe, 2014).

## 5. Conclusion

The present study provides initial evidence in support of the use of multivariable prediction modeling to derive individualized risk estimates for psychopathology among victimized children. Although we acknowledge a need for external validation and evaluation of clinical usefulness before these models can be confidently integrated into decision-making, they nevertheless offer the potential to supplement current practice, ultimately providing more personalized care for some of the most vulnerable children in society.

## CRedit authorship contribution statement

**Alan J. Meehan:** Formal analysis, Investigation, Writing - original draft. **Rachel M. Latham:** Writing - review & editing. **Louise Arseneault:** Data curation, Funding acquisition, Writing - review & editing. **Daniel Stahl:** Methodology, Writing - review & editing. **Helen L. Fisher:** Conceptualization, Funding acquisition, Writing - review & editing. **Andrea Danese:** Conceptualization, Funding acquisition, Writing - original draft, Writing - review & editing.

## Declaration of Competing Interest

The authors have no conflicts of interest to declare.

## Role of the funding source

The E-Risk Study is funded by the UK Medical Research Council (grant number G1002190). Additional support was provided by the US National Institute of Child Health and Human Development (grant number HD077482); the Jacobs Foundation; and a research grant from the National Society for Prevention of Cruelty to Children (NSPCC) and Economic and Social Research Council (ESRC). HLF was supported by an MQ Fellows Award (grant number MQ14F40). LA is the Mental Health Leadership Fellow for the ESRC. This paper represents independent research part-funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London to AD. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

## Acknowledgments

We are grateful to the Study mothers and fathers, the twins, and the twins' teachers and neighbors for their participation. Our thanks to Professors Terrie Moffitt and Avshalom Caspi who founded the E-Risk Study, CACI, Inc., the UK Ministry of Justice, and the E-Risk team for their dedication, hard work, and insights.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.jad.2019.10.034](https://doi.org/10.1016/j.jad.2019.10.034).

## References

- Braga, T., Gonçalves, L.C., Basto-Pereira, M., Maia, Â., 2017. Unraveling the link between maltreatment and juvenile antisocial behavior: a meta-analysis of prospective longitudinal studies. *Aggress. Behav.* 33, 37–50. <https://doi.org/10.1016/j.avb.2017.01.006>.
- Brier, G.W., 1950. Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* 78, 1–3.
- Cannon, T.D., Yu, C., Addington, J., Bearden, C.E., Cadenhead, K.S., Cornblatt, B.A., Heinsen, R., Jeffries, C.D., Mathalon, D.H., McGlashan, T.H., Perkins, D.O., Seidman, L.J., Tsuang, M.T., Walker, E.F., Woods, S.W., Kattan, M.W., 2016. An individualized risk calculator for research in prodromal psychosis. *Am. J. Psychiatry* 173, 980–988. <https://doi.org/10.1176/appi.ajp.2016.15070890>.
- Capusan, A.J., Kuja-Halkola, R., Bendtsen, P., Viding, E., McCrory, E., Marteinsdottir, I., Larsson, H., 2016. Childhood maltreatment and attention deficit hyperactivity disorder symptoms in adults: a large twin study. *Psychol. Med.* 46, 2637–2646. <https://doi.org/10.1017/S0033291716001021>.
- Cicchetti, D., 2013. Annual research review: resilient functioning in maltreated children - past, present, and future perspectives. *J. Child Psychol. Psychiatry* 54, 402–422. <https://doi.org/10.1111/j.1469-7610.2012.02608.x>.
- Damen, J.A.A.G., Hooft, L., Schuit, E., Debray, T.P.A., Collins, G.S., Tzoulaki, I., Lassale, C.M., Siontis, G.C.M., Chiochia, V., Roberts, C., Schlüssel, M.M., Gerry, S., Black, J.A., Heus, P., van der Schouw, Y.T., Peelen, L.M., Moons, K.G.M., 2016. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ* 353. <https://doi.org/10.1136/bmj.i2416>.
- Danese, A., Moffitt, T.E., Arseneault, L., Bleiberg, B.A., Dinardo, P.B., Gandelman, S.B., Houts, R., Ambler, A., Fisher, H.L., Poulton, R., Caspi, A., 2017. The origins of cognitive deficits in victimized children: implications for neuroscientists and clinicians. *Am. J. Psychiatry* 174, 349–361. <https://doi.org/10.1176/appi.ajp.2016.16030333>.
- Fenlon, C., O'Grady, L., Doherty, M.L., Dunnion, J., 2018. A discussion of calibration techniques for evaluating binary and categorical predictive models. *Prev. Vet. Med.* 149, 107–114. <https://doi.org/10.1016/j.prevetmed.2017.11.018>.
- Fisher, H.L., Caspi, A., Moffitt, T.E., Wertz, J., Gray, R., Newbury, J., Ambler, A., Zavos, H., Danese, A., Mill, J., Odgers, C.L., Pariante, C., Wong, C.C., Arseneault, L., 2015. Measuring adolescents' exposure to victimization: the environmental risk (E-Risk) longitudinal twin study. *Dev. Psychopathol.* 27, 1399–1416. <https://doi.org/10.1017/S0954579415000838>.
- Fisher, H.L., Schreier, A., Zammit, S., Maughan, B., Munafo, M.R., Lewis, G., Wolke, D., 2013. Pathways between childhood victimization and psychosis-like symptoms in the ALSPAC birth cohort. *Schizophr. Bull.* 39, 1045–1055. <https://doi.org/10.1093/schbul/sbs088>.
- Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33, 1–22.
- Fritz, J., de Graaff, A.M., Caisley, H., van Harmelen, A.-L., Wilkinson, P.O., 2018. A systematic review of amenable resilience factors that moderate and/or mediate the relationship between childhood adversity and mental health in young people. *Front. Psychiatry* 9, 230. <https://doi.org/10.3389/fpsy.2018.00230>.
- Fusar-Poli, P., Hijazi, Z., Stahl, D., Steyerberg, E.W., 2018. The science of prognosis in psychiatry: a review. *JAMA Psychiatry* 75, 1289–1297. <https://doi.org/10.1001/jamapsychiatry.2018.2530>.
- Fusar-Poli, P., Rutigliano, G., Stahl, D., Davies, C., Bonoldi, I., Reilly, T., McGuire, P., 2017. Development and validation of a clinically based risk calculator for the transdiagnostic prediction of psychosis. *JAMA Psychiatry* 74, 493–500. <https://doi.org/10.1001/jamapsychiatry.2017.0284>.
- Gjone, H., Novik, T.S., 1995. Parental ratings of behaviour problems: a twin and general population comparison. *J. Child Psychol. Psychiatry* 36, 1213–1224.
- Goldstein, B.A., Navar, A.M., Carter, R.E., 2017. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *Eur. Heart J.* 38, 1805–1814. <https://doi.org/10.1093/eurheartj/ehw302>.
- Hafeman, D.M., Merranko, J., Goldstein, T.R., Axelson, D., Goldstein, B.I., Monk, K., Hickey, M.B., Sakolsky, D., Diler, R., Iyengar, S., Brent, D.A., Kupfer, D.J., Kattan, M.W., Birmaher, B., 2017. Assessment of a person-level risk calculator to predict new-onset bipolar spectrum disorder in youth at familial risk. *JAMA Psychiatry* 74, 841–847. <https://doi.org/10.1001/jamapsychiatry.2017.1763>.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, New York, NY.
- Hippisley-Cox, J., Coupland, C., 2015. Development and validation of risk prediction algorithms to estimate future risk of common cancers in men and women: prospective cohort study. *BMJ Open* 5, e007825. <https://doi.org/10.1136/bmjopen-2015-007825>.
- Hosmer, D.W., Lemeshow, S., Sturdivant, R.X., 2013. *Applied Logistic Regression*, 3rd ed. John Wiley & Sons, New York, NY.
- Kim-Cohen, J., Caspi, A., Moffitt, T.E., Harrington, H., Milne, B.J., Poulton, R., 2003. Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. *Arch. Gen. Psychiatry* 60, 709–717. <https://doi.org/10.1001/archpsyc.60.7.709>.
- Lahey, B.B., Krueger, R.F., Rathouz, P.J., Waldman, I.D., Zald, D.H., 2017. A hierarchical causal taxonomy of psychopathology across the life span. *Psychol. Bull.* 143, 142–186. <https://doi.org/10.1037/bul0000069>.
- Meng, X., Fleury, M.-J., Xiang, Y.-T., Li, M., D'Arcy, C., 2018. Resilience and protective factors among people with a history of child maltreatment: a systematic review. *Soc. Psychiatry Psychiatr. Epidemiol.* 53, 453–475. <https://doi.org/10.1007/s00127-018-1485-2>.
- Moffitt, T.E., E-Risk Study Team, 2002. Teen-aged mothers in contemporary Britain. *J. Child Psychol. Psychiatry* 43, 727–742.



- Moons, K.G., Royston, P., Vergouwe, Y., Grobbee, D.E., Altman, D.G., 2009. Prognosis and prognostic research: what, why, and how? *BMJ* 338, b375. <https://doi.org/10.1136/bmj.b375>.
- Morgan, J.F., Reid, F., Lacey, J.H., 1999. The Scoff questionnaire: assessment of a new screening tool for eating disorders. *BMJ* 319, 1467–1468.
- Nanni, V., Uher, R., Danese, A., 2012. Childhood maltreatment predicts unfavorable course of illness and treatment outcome in depression: a meta-analysis. *Am. J. Psychiatry* 169, 141–151. <https://doi.org/10.1176/appi.ajp.2011.11020335>.
- Pavlou, M., Ambler, G., Seaman, S.R., Guttman, O., Elliott, P., King, M., Omar, R.Z., 2015. How to develop a more accurate risk prediction model when there are few events. *BMJ* 351, h3868. <https://doi.org/10.1136/bmj.h3868>.
- Polanczyk, G., Moffitt, T.E., Arseneault, L., Cannon, M., Ambler, A., Keefe, R.S.E., Houts, R., Odgers, C.L., Caspi, A., 2010. Etiological and clinical features of childhood psychotic symptoms: results from a birth cohort. *Arch. Gen. Psychiatry* 67, 328–338. <https://doi.org/10.1001/archgenpsychiatry.2010.14>.
- Radford, L., Corral, S., Bradley, C., Fisher, H.L., 2013. The prevalence and impact of child maltreatment and other types of victimization in the UK: findings from a population survey of caregivers, children and young people and young adults. *Child Abuse Negl.* 37, 801–813. <https://doi.org/10.1016/j.chiabu.2013.02.004>.
- Rutter, M., 2013. Annual research review: resilience – clinical implications. *J. Child Psychol. Psychiatry* 54, 474–487. <https://doi.org/10.1111/j.1469-7610.2012.02615.x>.
- Schaefer, J.D., Moffitt, T.E., Arseneault, L., Danese, A., Fisher, H.L., Houts, R., Sheridan, M.A., Wertz, J., Caspi, A., 2018. Adolescent victimization and early-adult psychopathology: approaching causal inference using a longitudinal twin study to rule out noncausal explanations. *Clin. Psychol. Sci.* 6, 352–371. <https://doi.org/10.1177/2167702617741381>.
- Steyerberg, E.W., 2009. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. Springer, New York, NY.
- Steyerberg, E.W., Vergouwe, Y., 2014. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur. Heart J.* 35, 1925–1931. <https://doi.org/10.1093/eurheartj/ehu207>.
- Takizawa, R., Maughan, B., Arseneault, L., 2014. Adult health outcomes of childhood bullying victimization: evidence from a five-decade longitudinal British birth cohort. *Am. J. Psychiatry* 171, 777–784. <https://doi.org/10.1176/appi.ajp.2014.13101401>.
- Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Series B Stat. Methodol.* 58, 267–288.
- Wang, X., Huang, Y., Li, L., Dai, H., Song, F., Chen, K., 2018. Assessment of performance of the Gail model for predicting breast cancer risk: a systematic review and meta-analysis with trial sequential analysis. *Breast Cancer Res* 20, 18. <https://doi.org/10.1186/s13058-018-0947-5>.
- Yarkoni, T., Westfall, J., 2017. Choosing prediction over explanation in psychology: lessons from machine learning. *Perspect. Psychol. Sci.* 12, 1100–1122. <https://doi.org/10.1177/1745691617693393>.