

# Disentangling the impact of childhood abuse and neglect on depressive affect in adulthood: A machine learning approach in a general population sample

Linda T. Betz<sup>a,\*</sup>, Marlene Rosen<sup>a</sup>, Raimo K.R. Salokangas<sup>b</sup>, Joseph Kambeitz<sup>a</sup>

<sup>a</sup> Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital of Cologne, University of Cologne, Cologne, Germany

<sup>b</sup> Department of Psychiatry, University of Turku, Turku, Finland

## ARTICLE INFO

### Keywords:

Early adversity  
Childhood trauma  
Childhood maltreatment  
Adverse childhood experiences  
Depression  
Precision medicine

## ABSTRACT

**Background:** Different types of childhood maltreatment (CM) are key risk factors for psychopathology. Specifically, there is evidence for a unique role of emotional abuse in affective psychopathology in children and youth; however, its predictive power for depressive symptomatology in adulthood is still unknown. Additionally, emotional abuse encompasses several facets, but the strength of their individual contribution to depressive affect has not been examined.

**Method:** Here, we used a machine learning (ML) approach based on Random Forests to assess the performance of domain scores and individual items from the Childhood Trauma Questionnaire (CTQ) in predicting self-reported levels of depressive affect in an adult general population sample. Models were generated in a training sample ( $N = 769$ ) and validated in an independent test sample ( $N = 466$ ). Using state-of-the-art methods from interpretable ML, we identified the most predictive domains and facets of CM for adult depressive affect.

**Results:** Models based on individual CM items explained more variance in the independent test sample than models based on CM domain scores ( $R^2 = 7.6\%$  vs.  $6.4\%$ ). Emotional abuse, particularly its more subjective components such as reactions to and appraisal of the abuse, emerged as the strongest predictors of adult depressive affect.

**Limitations:** Assessment of CM was retrospective and lacked information on timing and duration. Moreover, reported rates of CM and depressive affect were comparatively low.

**Conclusions:** Our findings corroborate the strong role of subjective experience in CM-related psychopathology across the lifespan that necessitates greater attention in research, policy, and clinical practice.

## 1. Introduction

A history of childhood maltreatment (CM) has been associated with adverse long-term consequences for mental health, including increased risk for experiencing depressive affect (Hoppen and Chalder, 2018; Humphreys et al., 2020; Kessler et al., 1997; Lindert et al., 2014). CM includes any act of commission or omission by a parent or other caregiver that results in actual, potential or threatened harm to a child, even if unintentional (Gilbert et al., 2009). Generally, five types of childhood maltreatment are distinguished: physical abuse, emotional abuse, sexual abuse, physical neglect, and emotional neglect (for details, see Zeanah and Humphreys, 2018). To the extent that different types of CM represent vastly different social experiences with distinct influences on

development, they have been found to be differentially relevant in shaping pathways to affective symptoms in clinical and non-clinical samples (Betz et al., 2020b; Cecil et al., 2017; de Oliveira et al., 2018; Haidl et al., 2021; Humphreys et al., 2020; LoPilato et al., 2019; McLaughlin et al., 2014; Salokangas et al., 2019; Sheridan and McLaughlin, 2014). Specifically, a growing body of literature points to childhood emotional abuse as an important contributor to the development of depression: Meta-analytic evidence suggests that, though all types of CM are significantly associated with depression symptom scores, emotional abuse demonstrates the strongest link, with smaller effect sizes in clinical than in non-clinical samples (Humphreys et al., 2020). When controlling for other types of CM, there is a unique relation between emotional abuse and internalizing as well as affective

\* Corresponding author at: Department of Psychiatry and Psychotherapy, Kerpener Str. 62, 50937 Cologne, Germany.

E-mail address: [linda.betz@uk-koeln.de](mailto:linda.betz@uk-koeln.de) (L.T. Betz).

<https://doi.org/10.1016/j.jad.2022.07.042>

Received 4 February 2022; Received in revised form 15 July 2022; Accepted 19 July 2022

Available online 23 July 2022

0165-0327/© 2022 Elsevier B.V. All rights reserved.

psychopathology in children, adolescents, and young adults from the community (Brown et al., 2016; Cecil et al., 2017; de Oliveira et al., 2018). Emotional abuse, which is generally considered to be difficult to define, recognize and evidence, has received less attention than other types of maltreatment, and disentangling its effects has significant implications for childcare policy. At this juncture, it is unclear whether specific types of CM, in particular emotional abuse, have a similarly strong and unique role in depressive affect in adult population samples. To address this question, it is necessary to model different, often co-occurring types of CM simultaneously rather than focusing on a single type of CM or total scores (Bernet and Stein, 1999; Cecil et al., 2016, 2017; de Oliveira et al., 2018; McCrory et al., 2017; Salokangas et al., 2019, 2020; Tiemeier, 2020). Additionally, there is a growing awareness that emotional abuse comprises dissimilar aspects, such as behaviorally specific actions (caregiver calling names) and beliefs (thinking that parents wished they were never born), but the strength of their individual contribution to depressive affect has not been examined to date (Cecil et al., 2017; de Oliveira et al., 2018; Meehan et al., 2021). Expanding modeling efforts to include a broader range of specific components of CM allows to better tease out, for example, which aspect drives the observed effects of emotional abuse (e.g., behaviorally specific actions vs. beliefs), which may lead to a better characterization of the processes by which depressive affect, as well as other mental health outcomes, arise (Cecil et al., 2017; Read and Mayne, 2017). These associations are obscured in the evaluation of CM domain scores.

Machine learning (ML) methods are ideal for analyses of complex, potentially non-linear associations among large numbers of different aspects of CM predicting depressive affect. ML, a paradigm with an inherent focus on prediction (Bzdok et al., 2021), integrates well into mental health research that has come to increasingly scrutinize the potential of risk factors and biomarkers for individualized predictions in unseen data (Bzdok et al., 2020; Dwyer et al., 2018; Rosen et al., 2021; Rosenbusch et al., 2021; Yarkoni and Westfall, 2017). This is achieved by testing whether a ML model generated in an initial sample ('training' data set) can accurately predict the outcome in individuals in a different sample ('test' data set). Independent test data is essential due to the risk of 'overfitting': Prediction models do not only capitalize on patterns reflecting true relationships between predictors and outcome, but also on idiosyncratic noise of the particular data set they were trained on (Bzdok et al., 2021; Rosenbusch et al., 2021; Yarkoni and Westfall, 2017). Thus, assessing the predictive power of a prediction model on the sample they were trained on can result in overly optimistic estimates (Bzdok et al., 2021; Yarkoni and Westfall, 2017). Validation, i.e., testing how well a model based on CM can predict depressive affect in out-of-sample data, can provide a more realistic estimate of the role of different types of CM as risk factors for depressive affect (Altman and Royston, 2000; Bzdok et al., 2020, 2021; Danese, 2020; Justice et al., 1999). In summary, the predictive power of CM for depressive outcomes cannot be automatically inferred from strong in-sample associations (Bzdok et al., 2020; Humphreys et al., 2020). As we move toward precision psychiatry, appropriate frameworks, such as ML, are needed to explicitly evaluate how well predictive patterns between CM and depression generalize to new individuals (Bzdok et al., 2021).

To address the existing research gaps, the main goals in the present study were twofold: (1) to assess the how well a ML model based on domain scores and individual items from the Childhood Trauma Questionnaire (CTQ) can predict self-reported levels of depressive affect in adulthood in unobserved data, and (2) to identify the most predictive domains and facets of CM for depressive affect in adulthood, using state-of-the-art methods from interpretable ML (Molnar, 2019). Based on prior research (Brown et al., 2016; Cecil et al., 2017; de Oliveira et al., 2018; Humphreys et al., 2020; Salokangas et al., 2019; Spertus et al., 2003), we hypothesized that emotional abuse would emerge as the strongest predictor of depressive affect in adulthood. In terms of individual facets of emotional abuse, we expected that aspects embedded in all forms of CM, such as beliefs of being worthless or hated, would have

the strongest link to depressive affect (Cecil et al., 2017; de Oliveira et al., 2018; Hart et al., 1997; Hart and Glaser, 2011).

## 2. Methods

### 2.1. Samples

#### 2.1.1. Training sample

The data used as the training sample in this study come from the Midlife Development in the United States II (MIDUS II) Biomarker project, a longitudinal follow-up of the original MIDUS I study, a nationally representative study of health and aging in the noninstitutionalized civilian population of the 48 contiguous United States initiated in 1995 (Brim et al., 2019; Ryff et al., 2019) ( $N = 7108$ ). All living MIDUS I respondents were eligible for participation in the MIDUS biomarker project if their existing health information indicated that they could travel to one of three clinical research centers (University of California at Los Angeles, Georgetown University, University of Wisconsin–Madison) without undue risk to the respondent or project personnel. Members of the Milwaukee sample of African Americans newly recruited at MIDUS II were also part of the recruitment pool. There were no other eligibility criteria (Dienberg Love et al., 2010). The Biomarker project enrolled a total of 1255 respondents aged 35–86 years, including two distinct subsamples: (1) Main survey sample ( $n = 1054$ ) and (2) Milwaukee sample, a sample of African Americans from Milwaukee ( $n = 201$ ). A detailed description of the sampling and participation rates across the MIDUS studies is provided elsewhere (Dienberg Love et al., 2010). Relevant project data, including assessments of CM and depressive affect, were obtained by a standardized protocol between July 2004 and May 2009. We restricted the analytic sample to those aged 60 years and younger at the time of the biomarker project with available outcome data and <25 % missings in the predictor variables ( $N = 769$ ; 'training sample') to ensure that results were not influenced by aging-related depressive affect.

#### 2.1.2. Independent test sample

The data used as the independent test sample in this study come from the Midlife Development in the United States Refresher (MIDUS Refresher) Biomarker project, a longitudinal follow-up of the original MIDUS I Refresher (Weinstein et al., 2019). From 2011 to 2014, the MIDUS Refresher study recruited a national probability sample of 3577 adults, aged 25 to 74, designed to replenish the original MIDUS baseline cohort and paralleling the five decadal age groups of the MIDUS baseline survey. The MIDUS Refresher Biomarker project obtained data from 863 respondents ( $n = 746$  Main sample,  $n = 117$  African Americans from Milwaukee) between October 2012 and August 2016. The MIDUS Refresher Biomarker study employed the same standardized assessments at the same research centers as the original MIDUS sample. Thus, the present approach can be classified as a 'prospective validation' (Altman and Royston, 2000; Justice et al., 1999). Analogous to the training sample, we restricted analyses to participants aged 35–60 with available depressive affect scores and <25 % missings in the predictor variables ( $N = 466$ ; 'test sample'). The test sample was compared statistically to the training sample using appropriate classes of permutation tests (using the R package 'coin', Hothorn et al., 2008).

#### 2.1.3. Ethics approval

All participants gave their written informed consent to participate in the study prior to beginning the study procedures, and each MIDUS research center obtained institutional review board approval for all studies.

### 2.2. Predictors

CM (up to age 18) was assessed with the English version of the Childhood Trauma Questionnaire-Short Form (CTQ-SF, Bernstein et al.,

2003) in the MIDUS Biomarker Projects. The CTQ-SF is a self-report measure of CM, including five items within each of five different domains of CM: (1) physical neglect (failure of a caretaker to provide basic necessities for a child such as food, clothing, shelter); (2) physical abuse (bodily assault on a child posing a risk of or resulting in injury); (3) emotional neglect (failure of caretaker basic emotional and psychological needs for a child, such as love and nurturance); (4) emotional abuse (verbal assaults on a child, such as humiliation); and (5) sexual abuse (unwanted sexual conduct between a child and an adult). Responses for each item are recorded on a 5-point-Likert scale, ranging from 1 (never true) to 5 (very often true). Moreover, we included sex in the prediction models.

### 2.3. Outcome: depressive affect

As the outcome in the prediction models, we used the depressive affect subscale from the Center for Epidemiological Studies Depression Scale (CES–D, Radloff, 1977), a self-report depression scale for research in the general population. Respondents are asked how much they endorsed a given statement during the past two weeks, with responses given on a 4-point-Likert scale for each item, ranging from ‘Rarely or none of the time’ to ‘Most or all of the time’. The depressive affect scale consists of 7 items such as ‘I felt that I could not shake off the blues even with the help of my family and friends’, ‘I felt lonely’, or ‘I had crying spells’, and total scores range from 0 to 21.

### 2.4. Machine learning approach

To implement our analyses, we used the ‘mlr3’ ecosystem, an object-oriented framework for ML in R (Lang et al., 2019; R Development Core Team, 2021). For all analyses, we used R version 4.1.0. The ML algorithm used in all models was a Random Forest (RF), an ensemble learning method based on decision trees (Breiman, 2001), as implemented in the R package ‘ranger’ (Wright and Ziegler, 2015). RFs are powerful ensemble ML models, well-suited to deal with high-dimensional data containing different types of predictors, outliers and

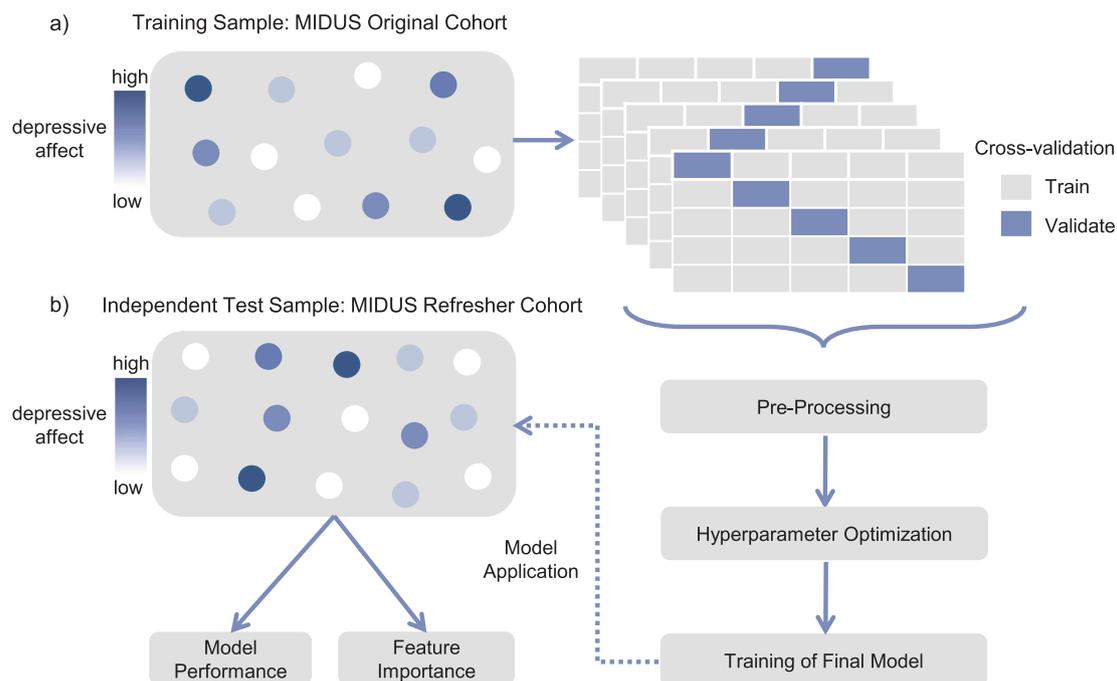
non-linear relationships (Breiman, 2001; Fernandez-Delgado et al., 2014; Touw et al., 2013). Within the RF modeling approach, we constructed an ensemble of 1000 decision trees, and used sampling without replacement as well as regularization (Altmann et al., 2010; Probst et al., 2019; Strobl et al., 2007). Within a 5-fold repeated (5 times) cross-validation resampling scheme, we performed pre-processing, hyperparameter optimization, and the fitting of the RF (details below). Fig. 1 illustrates the analytic workflow. Code to reproduce all analyses and figures is available at [https://github.com/LindaBetz/ML\\_CT\\_Depression](https://github.com/LindaBetz/ML_CT_Depression).

#### 2.4.1. Pre-processing

For pre-processing, we collapsed the rarest factor levels in the training samples to reduce noise associated with very rare factor levels, retaining at least two factor levels. Subsequently, we imputed missing data via the k-Nearest-Neighbor (kNN) algorithm. Overall, an average of 0.2 % of CM domain predictor variables was missing in the training data set per person, and 0.1 % in the test data set. In the training data set, 0.3 % of CM individual item predictor variables were missing on average per person, and 0.1 % in the test data set.

#### 2.4.2. Hyperparameter optimization

Following published tuning strategies (Probst et al., 2019), we optimized three hyperparameters that reflect the degree of randomness in a RF: the number of drawn candidate variables in each split (‘mtry’), the minimum number of observations in a terminal node (‘min.node.size’), and the fraction of observations to be used in each tree (‘sample.fraction’). Moreover, we tuned the amount of regularization in the RF (‘regularization.factor’). Additionally, we tuned several characteristics of our preprocessing pipeline, including the prevalence level above which no collapsing of factor levels was applied, as well as the numbers of neighbors used in kNN imputation of missing values. We optimized all hyperparameters with respect to percent of explained variance ( $R^2$ ) and used a random search with 100 evaluations to optimize the hyperparameter configuration (Bergstra and Bengio, 2012). For ranges in which the hyperparameters were tuned, see Supplementary Table 1.



**Fig. 1.** Analytic workflow. a) First, a random forest model was trained on data from the training sample, the Midlife Development in the United States (MIDUS) Biomarker sample (Brim et al., 2019; Ryff et al., 2019). This involved preprocessing and hyperparameter optimization within a five-times repeated five-fold cross-validation scheme, as well as subsequent training of a final model. b) The final model was then applied to the independent test sample, the MIDUS Refresher Biomarker sample (Weinstein et al., 2019) and estimates of model performance and feature importance were derived.

### 2.4.3. Application to independent test sample

The final model, which resulted from retraining a RF on the whole training sample with optimal hyperparameter settings derived from cross-validation, was used to predict levels of depressive affect in the independent test sample. Thus, there was a complete separation between training and testing of the model. Model performance was assessed with  $R^2$ .

### 2.4.4. Feature importance

We calculated feature importance of the CM predictor variables using a permutation-based approach (Fisher et al., 2019; Molnar, 2019; Strobl et al., 2007). The importance of a feature is quantified by calculating the increase in the prediction error of the model after permuting the feature, i.e., after ‘breaking’ any meaningful association between the feature and the outcome. The stronger the increase in model error due to permutation, the more important the feature, because the model relies on the feature for the prediction (Fisher et al., 2019; Molnar, 2019). Thus, permutation feature importance provides a highly condensed, global insight into the behaviour of a model. As recommended, we used the independent test sample, instead of the training sample, to assess permutation feature importance (Molnar, 2019). To obtain further context from the RF models, Accumulated Local Effects (ALE) (Apley and Zhu, 2016) were computed on the test sample to understand associations between important features and the outcome, depressive affect. ALE plots vary a feature across its range to consider its association with the outcome: ALE can be interpreted as the effect of the feature at a certain value compared to the average prediction of the data (Apley and Zhu, 2016; Molnar, 2019). ALE are unbiased in the presence of correlated features (Molnar, 2019), which makes them particularly useful for data from the mental health domain.

### 2.5. Assessment of the role of potential confounds

We assessed the impact of potential confounds (i.e., age, data collection site) post-hoc on the level of ML predictions, as proposed recently (Dinga et al., 2020). In brief, this method partitions the variance in the outcome into the parts attributable to the predictions from the model, the potential confounds, and the shared variance between the model predictions and the confounds. The lower the shared variance in the outcome, the higher the contribution of the model above and beyond confounds, and the lower the effect of confounds on the predictions of the model. Operating solely at the level of the outcome of ML models, this method avoids the problems of confound adjustment at the level of input variables and can be used for arbitrarily complex ML models (Dinga et al., 2020).

## 3. Results

### 3.1. Samples

Table 1 presents the sample characteristics of the training and test sample. The training sample ( $n = 769$ ) and the test sample ( $n = 466$ ) were comparable in terms of sex distribution, CM domain scores, and depressive affect scores. Participants in the test sample were slightly younger than in the training sample (48.1 vs. 49.8 years), were more likely to have been diagnosed with depression in their lifetime by a physician, and were overall less likely to have smoked regularly in their lifetime. Moreover, less data was collected at the University of Washington in the test compared to the training sample.

### 3.2. Prediction of depressive affect based on CM

#### 3.2.1. CM domain scores model

The model based on CM domain scores explained 6.4 % of variance in depressive affect in the independent test sample. Permutation feature importance (Fig. 2) showed that emotional abuse was the most

**Table 1**

Demographic and clinical characteristics in the training and test sample. Mean (SD) unless stated otherwise.

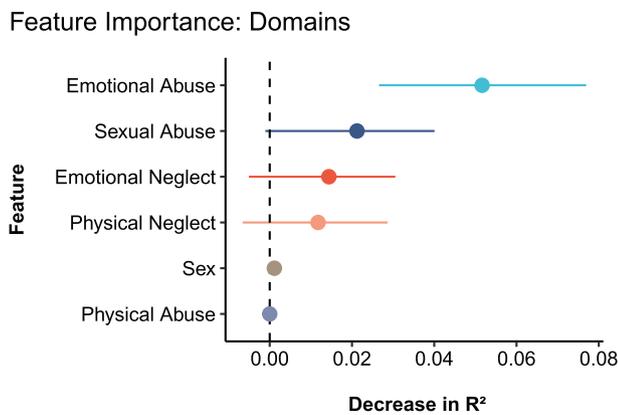
Variable	Training sample ( $n = 769$ )	Test sample ( $n = 466$ )	Comparison (test statistic, $p$ -value)
Age	49.8 (6.4)	48.1 (7.1)	$Z = 4.49, p < .001$
Sex (% female)	57.2	57.9	$\chi^2 = 0.06, p = .812$
Married or living with a partner (% yes)	70.7	65.7	$\chi^2 = 3.6139, p = .064$
Currently employed (% yes)	84.0	84.1	$\chi^2 = 0.02, p = .936$
Data collection site (%)	UCLA (32.4), UW (45.8) Georgetown (21.8)	UCLA (33.5), UW (43.1), Georgetown (23.4)	$\chi^2 = 282.2, p < .001$
Childhood trauma scores (CTQ)			
Emotional neglect	10.1 (4.6)	10.2 (4.7)	$Z = -0.32, p = .754$
Physical neglect	7.0 (2.9)	7.0 (2.9)	$Z = -0.14, p = .896$
Emotional abuse	8.7 (4.7)	8.6 (4.4)	$Z = 0.44, p = .651$
Physical abuse	7.4 (3.4)	7.4 (3.7)	$Z = 0.04, p = .969$
Sexual abuse	6.9 (4.4)	7.1 (4.7)	$Z = -0.71, p = .482$
Depressive affect score (CES-D)	2.3 (3.5)	2.3 (3.0)	$Z = 0.08, p = .939$
Ever had depression diagnosed by a physician (% yes)	21.3	27.7	$\chi^2 = 6.21, p = .013$
Number of chronic health conditions ever <sup>a</sup> (median, IQR)	3 (1–5)	3 (2–4)	$Z = -1.73, p = .086$
BMI	29.9 (7.1)	30.3 (8.0)	$Z = -0.78, p = .425$
Ever smoked regularly (% yes)	45.3	36.9	$\chi^2 = 8.29, p = .005$

Abbreviations: BMI: Body Mass Index; CES—D: Center for Epidemiologic Studies Depression Scale; CTQ: Childhood Trauma Questionnaire; UCLA: University of California, Los Angeles; UW: University of Washington.

<sup>a</sup> Note: Respondents were asked to indicate (yes/no) whether they were ever diagnosed with each of the following conditions: asthma, cancer, heart disease, tuberculosis, other lung problems, joint diseases, backache, skin trouble, thyroid disease, hay fever, stomach trouble, urinary/bladder problems, constipation, gallbladder problems, foot trouble, varicose veins, HIV, autoimmune diseases, mouth problems, high blood pressure, emotional disorders, alcohol/drug problems, migraines, diabetes, neurological disorders, stroke, ulcer, hernia, hemorrhoids, and swallowing problems.

important feature to contribute robustly to accurate prediction in the test sample, with a median reduction in explained variance in depressive affect of 5.2 % when the feature was permuted. Second most important was sexual abuse (median reduction: 2.1 %), yet the contribution of this feature and the remaining ones can be described as less robust, as indicated by the fact that 0 was contained within the 90 % confidence limits of the null distribution of decrease in  $R^2$ . Note that the sum of the feature importances is not equal to the total explained variation in the model, but larger. This is due to interaction effects, which are also ‘broken’ by permutation and included in the importances of more than one feature (Molnar, 2019).

The ALE plot reveals that higher scores on all CM domains, except physical abuse, were associated with greater predicted depressive affect (Fig. 3). Particularly for physical neglect and physical abuse, the ALE plot indicates saturation effects: above certain thresholds, increases in self-reported CM are associated with only minimal increases in predicted depressive affect. For emotional and sexual abuse, and, to some degree,



**Fig. 2.** Feature importance in the model based on domain scores from the Childhood Trauma Questionnaire (CTQ; Bernstein et al., 2003). Feature importance was computed in the test sample with a permutation-based approach, which measures the decrease in predictive accuracy ( $R^2$  score) that results from permuting a single feature. For each feature, we plotted the median decrease of  $R^2$  across 1000 permutations, along with the corresponding 90% confidence interval (ranging from the 5th to the 95th quantile). Greater decreases reflect greater importance of the feature. Decreases close to 0 suggest that the corresponding feature is irrelevant for predicting depressive affect in the test sample.

emotional neglect, increases in predicted depressive affect are observable also at higher levels of CM.

### 3.2.2. CM individual items model

The model based on individual facets of CM, measured by individual

items from the CTQ, explained 7.6 % of variance in depressive affect in the test sample. Permutation feature importance (Fig. 4) showed that two facets reflecting emotional abuse contributed robustly to accurate prediction in the test sample: ‘Felt that family member hated me’ (median reduction in explained variance: 2.6 %), and ‘Believe I was emotionally abused’ (median reduction in explained variance: 1.1 %). The predictive contribution of the other features was overall less robust, as indicated by the fact that 0 was contained within the 90 % confidence limits of the null distribution of decrease in  $R^2$ . Similar to the model based on CM domain scores, feature importance values do not add up due to presence of interaction effects (Molnar, 2019).

The six most predictive features all showed positive associations with depressive affect, i.e., higher levels of CM went along with higher levels of depressive affect (Fig. 5).

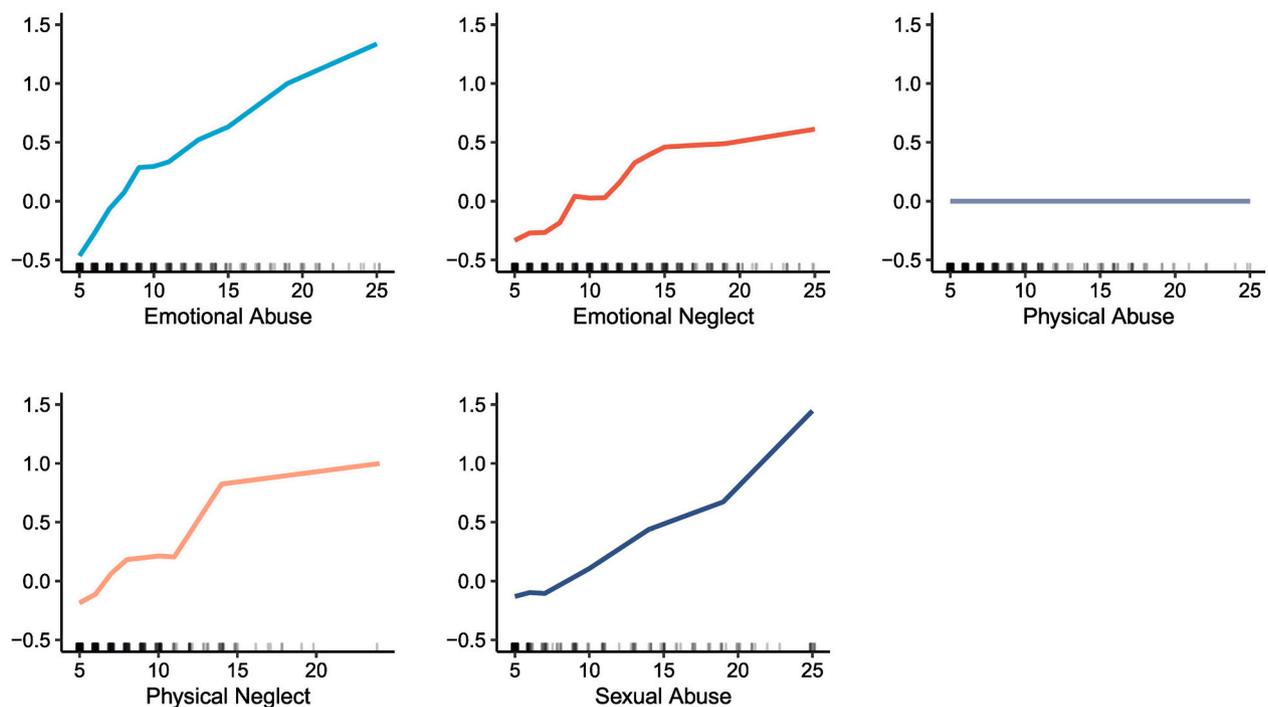
### 3.3. Assessment of the role of potential confounds

The variance in depressive affect due to potential confounds age and site ( $R^2_{\text{confounds}} = 0.015$ ) showed virtually no overlap with the variance explained by the predictions from the CM domain scores model ( $R^2_{\text{shared}} = -0.0004$ ) or the CM individual items model ( $R^2_{\text{shared}} = -0.0009$ ). These results suggest that the effect of these confounds on the ML models are negligible (Dinga et al., 2020). ML predictions and the predictions from the assessed confounds seem to capture different, independent aspects of depressive affect in the present sample.

## 4. Discussion

In this study, we (1) provided a realistic quantification of the predictive power of childhood abuse and neglect for depressive affect in adulthood with a ML and (2) identified the most predictive domains and

### Accumulated Local Effects for Depressive Affect: Domains



**Fig. 3.** Effect of features on prediction of depressive affect in the model based on domain scores from the Childhood Trauma Questionnaire (CTQ; Bernstein et al., 2003), as assessed by Accumulated Local Effect (ALE) plots in the test sample. The value of the ALE can be interpreted as the main effect of the feature at a certain value compared to the average prediction of the data. Specifically, positive ALEs indicate an increase in predicted depressive affect compared to the average prediction, while negative ALEs reflect a decrease in predicted depressive affect.

Feature Importance: Items

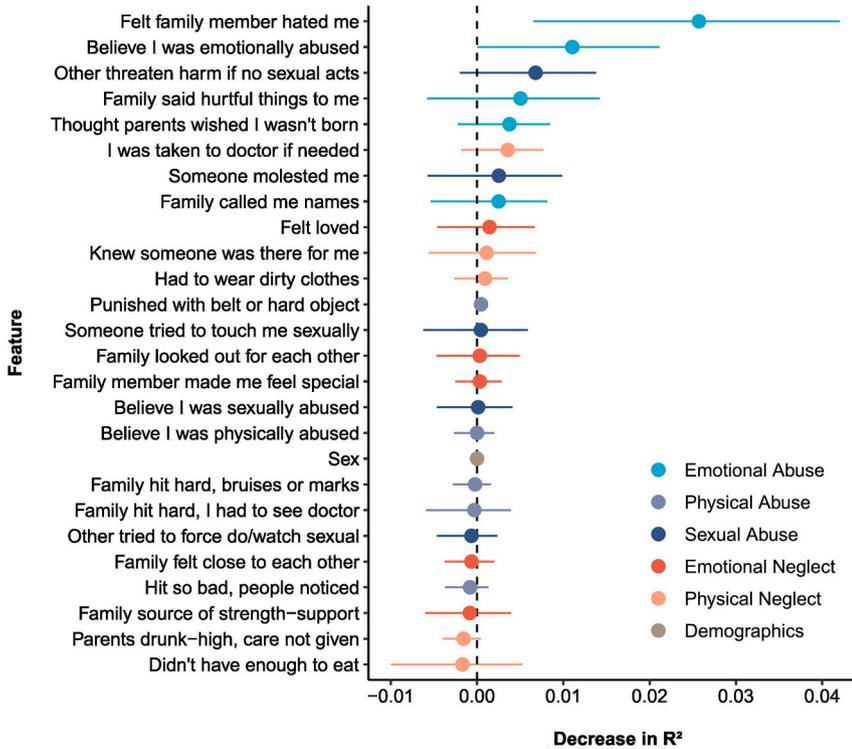


Fig. 4. Feature importance in the model based on individual items from the Childhood Trauma Questionnaire (CTQ; Bernstein et al., 2003). Feature importance was computed in the test sample with a permutation-based approach, which measures the decrease in predictive accuracy ( $R^2$  score) that results from permuting a single feature. For each feature, we plotted the median decrease of  $R^2$  across 1000 permutations, along with the corresponding 90% confidence interval (ranging from the 5th to the 95th quantile). Greater decreases reflect greater importance of the feature. Decreases close to 0 suggest that the corresponding feature is irrelevant for predicting depressive affect in the test sample. Wording of the items from the CTQ was shortened for visualization purposes.

Accumulated Local Effects for Depressive Affect: Items

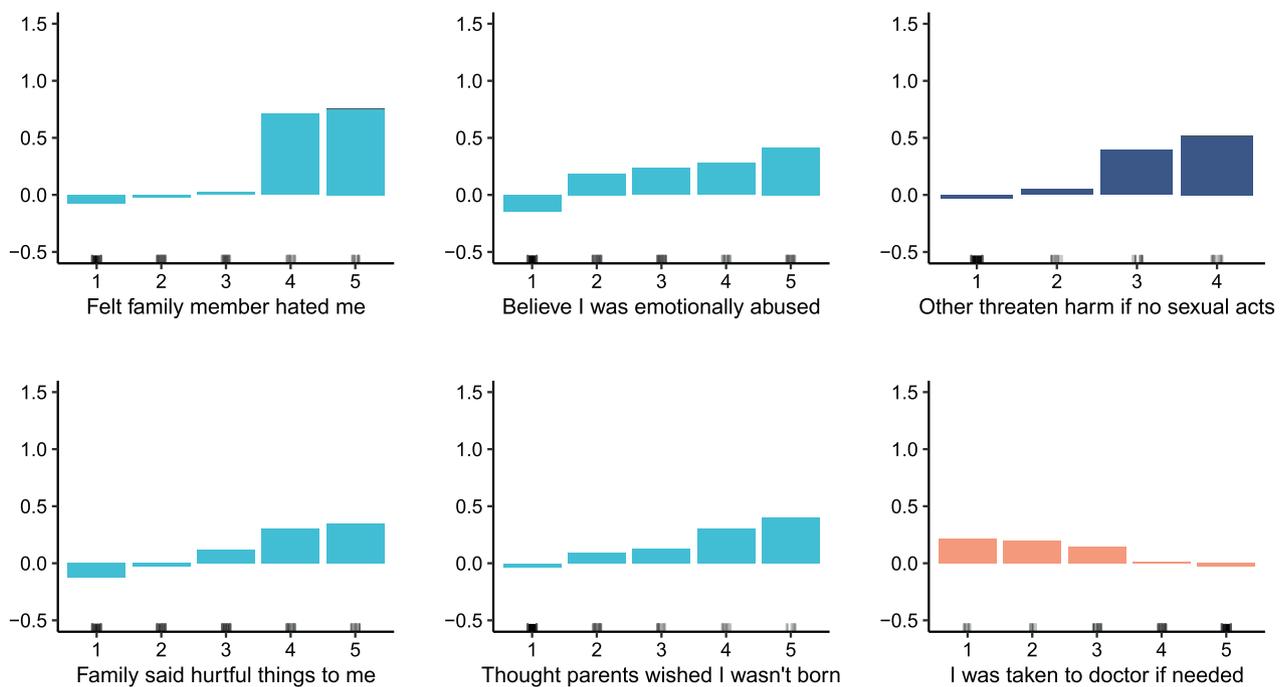


Fig. 5. The six most important features' effects on prediction of depressive affect in the model based on individual items from the Childhood Trauma Questionnaire (CTQ; Bernstein et al., 2003), as assessed by Accumulated Local Effect (ALE) plots in the test sample. The value of the ALE can be interpreted as the main effect of the feature at a certain value compared to the average prediction of the data. Specifically, positive ALEs indicate an increase in predicted depressive affect compared to the average prediction, while negative ALEs reflect a decrease in predicted depressive affect. Wording of the items from the CTQ was shortened for visualization purposes.

items of the CTQ using methods from interpretable ML. Models based on domains and individual items explained 6.4 % and 7.6 % of the variance in depressive affect in an independent test sample, respectively. Extending findings in child and youth cohorts (Brown et al., 2016; Cecil et al., 2017; de Oliveira et al., 2018), the strongest and most robust predictor of depressive affect in adulthood was emotional abuse. Moreover, present findings provide empirical support for the idea that particularly subjective components of emotional abuse – feeling hated and the subjective appraisal of emotional abuse – may underlie its unique impact on mental health outcomes.

In line with a recent study that showed only fair discriminative performance of cumulative exposure to early adversity for depressed mood (Meehan et al., 2021), the observed predictive power of CM was small in absolute numbers also in the present study, which can be expected given the highly multifactorial nature of adult depressive symptoms (Fried and Nesse, 2015; Kendler, 2019; Kendler and Aggen, 2017). Collectively, these results show that information on CM alone is insufficient to predict individual risk for depressive affect in adulthood. Still, present results highlight the relative relevance of early adversity: CM explained more variance in adult depressive symptomatology than polygenic risk scores typically do when applied to independent population samples, for instance (Demirkan et al., 2011; Musliner et al., 2015; Thorp et al., 2021). Most likely, information on early adversity will only reach full clinical significance within a multimodal clinical workflow to identify individuals at the greatest risk of psychopathology and with the greatest need for intervention (Danese, 2020; Koutsouleris et al., 2020; Lewis et al., 2019; Meehan et al., 2021).

As hypothesized, emotional abuse emerged as the strongest and uniquely robust predictor of depressive affect in adulthood. Taken together with previous findings (Brown et al., 2016; Cecil et al., 2017; de Oliveira et al., 2018; Edwards et al., 2003), this suggests that emotional abuse plays a unique role in internalizing and affective psychopathology across the whole lifespan, over and above other types of CM. This specificity collectively demonstrates the need to model different types of CM concurrently to disentangle their effects on mental health outcomes (Bernet and Stein, 1999; Betz et al., 2020a; Betz et al., 2020b; Cecil et al., 2017; Danese, 2020; de Oliveira et al., 2018; McCrory et al., 2017; Salokangas et al., 2020; Sheridan and McLaughlin, 2014; Tiemeier, 2020). Findings from the present study suggest that for a holistic picture, it may be necessary to disaggregate CM domain scores even further. Individual CM items explained a greater share of variance in depressive affect than domain scores did, indicating that some degree of predictive specificity is forfeited when aggregating information from different facets of CM to domain scores. From emotional abuse, the most predictive domain, two facets were particularly predictive of depressed affect: feeling hated by a family member, as well as the subjective appraisal of emotional abuse. These results provide empirical support for the idea that emotional abuse has a particularly strong impact on mental health because emotional reactions, such as feeling hated, are embedded in many forms of CM and adversity (Cecil et al., 2017; Hart et al., 1997; Hart and Glaser, 2011). Overall, present results corroborate the key role of subjective experience in CM (Berthelot et al., 2022; Danese and Widom, 2020): The effects of emotional abuse on depressive affect seem to be primarily driven by more subjective appraisals and reactions to the abuse as opposed to objective behaviors, such as being called names. These more subjective components of emotional abuse may underlie unhelpful cognitions about the self and others, which are central to the development and maintenance of depressive affect (Beck and Bredemeier, 2016; Danese and Widom, 2020; McCrory et al., 2017; Salokangas et al., 2018). Our findings are also consistent with recent research suggesting that how people process trauma and adverse relationships may be more important to affective responses than exposure to adversity per se (Berthelot et al., 2022). Ensuing maladaptive cognitions may also contribute to the transdiagnostic relevance of CM; for instance, in that they facilitate threatening or negative appraisals of daily life events, which may in turn feed into the formation of psychotic

psychopathology in some individuals (Betz et al., 2020a; Garety et al., 2001, 2007; McLaughlin et al., 2020). Thus, one viable strategy for the development of more effective interventions for CM-related psychopathology may consist in modifying the subjective appraisal and memory of emotional abuse (Alameda et al., 2020; Danese and Widom, 2020; Hart and Glaser, 2011). Despite bearing potential to tailor pathways in care based on the patient's needs, information on childhood abuse and neglect is typically either ignored or not acquired at all in clinical settings (Read et al., 2018a, 2018b). To improve clinical practice, a shift toward routine enquiry of information on trauma and adversity and adoption of trauma-informed care practices is therefore essential (Read et al., 2018b). In terms of policy, our results add to an accumulating body of research that highlights the detrimental effects of emotional abuse on mental health outcomes (Brown et al., 2016; Cecil et al., 2017; de Oliveira et al., 2018; Humphreys et al., 2020). Collectively, these findings point to the necessity to expand focus of childcare policy from physical and sexual abuse to emotional abuse, which is often overlooked (Baker and Brassard, 2019; de Oliveira et al., 2018).

An interesting insight facilitated by the present ML approach pertains to the relationship of the reported level exposure to different types of CM and depressive affect. For all CM domains except physical abuse, predicted depressive affect increased with increasing levels of reported CM. Notably, there were saturation effects for physical and emotional neglect, indicating that exposure to these CM domains ceases to have additional impact on depressive affect once a certain level of low to moderate exposure is reached. Emotional and sexual abuse followed a different pattern, where increases in CM were related to increases in predicted depressive affect even for higher exposures. Given the potential implications for research and practice, it will be important to examine if these diverging patterns across CM domains reflect a consistent phenomenon across different samples and age groups. From a methodological perspective, the observed saturation effects, which are inherently nonlinear, underscore the need to move to appropriate modeling approaches, such as RFs, to address the complexity of the relationship between early adversity and depressive affect. Such strategies allow for unique insights into the role of CM in psychopathology and may thereby open avenues for optimizing prevention and intervention on this transdiagnostic risk factor (Danese, 2020).

#### 4.1. Limitations and future directions

The present findings should be interpreted in light of several limitations that bear implications for future research. First, the retrospective measure of CM, the CTQ, and the measure of depressive affect, the CES–D, are self-report measures preferentially used in large general population samples due to their resource-efficiency. Recent research suggests that risk of psychopathology linked to subjective reports of CM is high, independent of whether the reports align with objective, court-documented evidence of CM (Danese and Widom, 2020). However, self-reported and register-based measures of CM identify largely independent groups (Danese and Widom, 2020), and due to a lack of objective information on CM in the present sample, a comparison of results based on subjective and objective accounts of CM was not possible. A related issue pertains to the CES–D, which, as a self-report measure, may overestimate the impact of CM on depressive affect in adulthood (Reuben et al., 2016). It could, however, also be argued that self-report measures of psychopathology are more sensitive in detecting depressive affect than observer ratings, particularly in the low to intermediate range (Cecil et al., 2017). In any case, extension of the present modeling approach to observer ratings of psychopathology, and potentially individual symptoms (Fried and Nesse, 2015), is desirable and may help to better characterize the different processes by which CM increases risk for depressive affect. While it has also been argued that retrospective reports of CM, such as assessed with the CTQ, could be biased by current depressive affect (e.g., Hardt and Rutter, 2004), a recent study suggests only little evidence for such a ‘depressive recall bias’: CTQ-scores were

found to be very stable across time, and altered only marginally in initially healthy participants after they were diagnosed with a depressive disorder (Goltermann et al., 2021). Second, the CTQ, while being one of the most widely used measures of CM in research (Humphreys et al., 2020), covers only a subset of relevant components of emotional abuse, and provides no information on the duration and timing of exposure (Cecil et al., 2017; Danese, 2020; Hart et al., 1997). The latter aspect may be particularly relevant in the context of the present results: The older children are when they are exposed to CM, the more they understand what is happening to them, which may impact their subjective appraisal of CM (Widom, 2020). Additionally, it has been argued that some aspects classified as emotional abuse, such as having family members calling them names, may either be common experiences in an unkind family environment, or be embedded within a demeaning pattern of communication (McCrory et al., 2017; Widom, 2020). With the CTQ alone, it is impossible to disentangle these possibilities. Third, the overall level of depressive affect, as well as exposure to high levels of certain facets of CM, such as physical abuse and neglect, was rather low in the present general population sample. This likely contributed to the observed reduced importance of these features for depressive affect. Generalization of results to clinical samples, where greater levels of depressive affect and CM can be observed, needs to be examined (McCrory et al., 2017). Related, testing the accuracy of predictions in data at different sites can shed light on the geographic transportability of findings (Justice et al., 1999). Fourth, the training and test sample differed slightly with respect to some characteristics; specifically, more people in the test set had been diagnosed with depression than in the training sample. Depressive symptoms assessed with the CES-D, however, did not differ significantly between the samples. One hypothesis consistent with this pattern of results is that participants in the test sample, studied approximately eight years after the training sample, were more likely to seek professional help for depressive symptoms, reflecting increasing mental health care utilization over time (Olsson et al., 2014). Finally, proxy diagnostic information was available for depression but not for other mental disorders, which would have allowed better characterization of the sample.

#### 4.2. Conclusion

The present study highlights that emotional abuse is the strongest predictor of depressive affect in adulthood, over and above other types of CM, and that especially subjective components of emotional abuse seem to underlie its unique impact. These results corroborate the contribution of subjective experience in CM-related psychopathology that necessitates greater attention in research, policy, and clinical practice.

#### Funding

The MIDUS study has been funded by the following: John D. and Catherine T. MacArthur Foundation Research Network, National Institute on Aging (P01-AG020166), National Institute on Aging (U19-AG051426). MIDUS Biomarker data collection was further supported by the National Institutes of Health National Center for Advancing Translational Sciences (NCATS) Clinical and Translational Science Award (CTSA) program as follows: UL1TR001409 (Georgetown), UL1TR001881 (UCLA) and 1UL1RR025011 (UW). The funding source had no involvement in the analysis and interpretation of data in this report.

#### CRediT authorship contribution statement

**Linda T. Betz:** Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Visualization. **Marlene Rosen:** Writing – review & editing. **Raimo K.R. Salokangas:** Writing – review & editing. **Joseph Kambeitz:** Methodology, Formal analysis,

Supervision, Writing – review & editing.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgment

The authors would like to thank all the participants and research staff of the MIDUS Study. The authors are also grateful for the opportunity to use these publicly available data.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jad.2022.07.042>.

#### References

- Alameda, L., Rodriguez, V., Carr, E., Aas, M., Trotta, G., Marino, P., Vorontsova, N., Herane-Vives, A., Gadelrab, R., Spinazzola, E., Di Forti, M., Morgan, C., Murray, R. M., 2020. A systematic review on mediators between adversity and psychosis: potential targets for treatment. *Psychol. Med.* 50 (12), 1966–1976. <https://doi.org/10.1017/S0033291720002421>.
- Altman, D.G., Royston, P., 2000. What do we mean by validating a prognostic model? *Stat. Med.* 19, 453 <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.610.4279&rep=rep1&type=pdf>. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.610.4279&rep=rep1&type=pdf>.
- Altman, A., Tološi, L., Sander, O., Lengauer, T., 2010. Permutation importance: a corrected feature importance measure. *Bioinformatics* 26 (10), 1340–1347. <https://doi.org/10.1093/bioinformatics/btq134>.
- Apley, D.W., Zhu, J., 2016. Visualizing the effects of predictor variables in black box supervised learning models. In *arXiv [stat.ME]*. arXiv. <http://arxiv.org/abs/1612.08468>.
- Baker, A.J.L., Brassard, M., 2019. Predictors of variation in self-reported rates of psychological maltreatment: a survey of statutes and a call for change. *Child Abuse Negl.* 96, 104102 <https://doi.org/10.1016/j.chiabu.2019.104102>.
- Beck, A.T., Bredemeier, K., 2016. A unified model of depression: integrating clinical, cognitive, biological, and evolutionary perspectives. *Clin. Psychol. Sci.* 4 (4), 596–619. <https://doi.org/10.1177/2167702616628523>.
- Bergstra, J., Bengio, Y., 2012. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* 13, 281–305 <https://jmlr.csail.mit.edu/papers/volume13/bergstra12a/bergstra12a.pdf>. <https://jmlr.csail.mit.edu/papers/volume13/bergstra12a/bergstra12a.pdf>.
- Bernet, C.Z., Stein, M.B., 1999. Relationship of childhood maltreatment to the onset and course of major depression in adulthood. *Depression Anxiety* 9 (4), 169–174. [https://doi.org/10.1002/\(sici\)1520-6394\(1999\)9:4<169::aid-da4>3.0.co;2-2](https://doi.org/10.1002/(sici)1520-6394(1999)9:4<169::aid-da4>3.0.co;2-2).
- Bernstein, D.P., Stein, J.A., Newcomb, M.D., Walker, E., Pogge, D., Ahluvalia, T., Stokes, J., Handelsman, L., Medrano, M., Desmond, D., Zule, W., 2003. Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse Negl.* 27 (2), 169–190. [https://doi.org/10.1016/s0145-2134\(02\)00541-0](https://doi.org/10.1016/s0145-2134(02)00541-0).
- Berthelot, N., Savard, C., Lemieux, R., Garon-Bissonnette, J., Ensink, K., Godbout, N., 2022. Development and validation of a self-report measure assessing failures in the mentalization of trauma and adverse relationships. *Child Abuse Negl.* 128, 105017 <https://doi.org/10.1016/j.chiabu.2021.105017>.
- Betz, L.T., Penzel, N., Kambeitz-Ilankovic, L., Rosen, M., Chisholm, K., Stainton, A., Haidl, T.K., Wenzel, J., Bertolino, A., Borgwardt, S., Brambilla, P., Lencer, R., Meisenzahl, E., Ruhrmann, S., Salokangas, R.K.R., Schultze-Lutter, F., Wood, S.J., Uthegrove, R., Koutsouleris, N., PRONIA consortium, 2020. General psychopathology links burden of recent life events and psychotic symptoms in a network approach. *NPJ Schizophrenia* 6 (1), 40. <https://doi.org/10.1038/s41537-020-00129-w>.
- Betz, L.T., Penzel, N., Rosen, M., Kambeitz, J., 2020. Relationships between childhood trauma and perceived stress in the general population: a network perspective. *Psychol. Med.* 51 (15), 2696–2706. <https://doi.org/10.1017/S003329172000135X>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45 (1), 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Brim, O.G., Ryff, C.D., Kessler, R.C., 2019. How healthy are we?: a national study of well-being at midlife. University of Chicago Press. <http://www.midus.wisc.edu/findings/pdfs/3.pdf>.
- Brown, S., Fite, P.J., Stone, K., Bortolato, M., 2016. Accounting for the associations between child maltreatment and internalizing problems: the role of alexithymia. *Child Abuse Negl.* 52, 20–28. <https://doi.org/10.1016/j.chiabu.2015.12.008>.
- Bzdok, D., Engemann, D., Thirion, B., 2020. Inference and prediction diverge in biomedicine. *Patterns (New York, N.Y.)* 1 (8), 100119. <https://doi.org/10.1016/j.patter.2020.100119>.

- Bzdok, D., Varoquaux, G., Steyerberg, E.W., 2021. Prediction, not association, paves the road to precision medicine. *JAMA Psychiatry* 78 (2), 127–128. <https://doi.org/10.1001/jamapsychiatry.2020.2549>.
- Cecil, C.A.M., Smith, R.G., Walton, E., Mill, J., McCrory, E.J., Viding, E., 2016. Epigenetic signatures of childhood abuse and neglect: implications for psychiatric vulnerability. *J. Psychiatr. Res.* 83, 184–194. <https://doi.org/10.1016/j.jpsychires.2016.09.010>.
- Cecil, C.A.M., Viding, E., Fearon, P., Glaser, D., McCrory, E.J., 2017. Disentangling the mental health impact of childhood abuse and neglect. *Child Abuse Negl.* 63, 106–119. <https://doi.org/10.1016/j.chiabu.2016.11.024>.
- Danese, A., 2020. Annual research review: rethinking childhood trauma—new research directions for measurement, study design and analytical strategies. *J. Child Psychol. Psychiatry Allied Discip.* 61 (3), 236–250. <https://doi.org/10.1111/jcpp.13160>.
- Danese, A., Widom, C.S., 2020. Objective and subjective experiences of child maltreatment and their relationships with psychopathology. *Nat. Hum. Behav.* 4 (8), 811–818. <https://doi.org/10.1038/s41562-020-0880-3>.
- de Oliveira, I.R., Matos-Ragazzo, A.C., Zhang, Y., Vasconcelos, N.M., Velasquez, M.L., Reis, D., Ribeiro, M.G., da Rocha, M.M., Rosario, M.C., Stallard, P., Cecil, C.A.M., 2018. Disentangling the mental health impact of childhood abuse and neglect: a replication and extension study in a Brazilian sample of high-risk youth. *Child Abuse Negl.* 80, 312–323. <https://doi.org/10.1016/j.chiabu.2018.03.021>.
- Demirkan, A., Penninx, B.W.J.H., Hek, K., Wray, N.R., Amin, N., Aulchenko, Y.S., van Dyck, R., de Geus, E.J.C., Hofman, A., Uitterlinden, A.G., Hottenga, J.-J., Nolen, W. A., Oostra, B.A., Sullivan, P.F., Willemsen, G., Zitman, F.G., Tiemeier, H., Janssens, A.C.J.W., Boomsma, D.I., Middeldorp, C.M., 2011. Genetic risk profiles for depression and anxiety in adult and elderly cohorts. *Mol. Psychiatry* 16 (7), 773–783. <https://doi.org/10.1038/mp.2010.65>.
- Dienberg Love, G., Seeman, T.E., Weinstein, M., Ryff, C.D., 2010. Bioindicators in the MIDUS national study: protocol, measures, sample, and comparative context. *J. Aging Health* 22 (8), 1059–1080. <https://doi.org/10.1177/0898264310374355>.
- Dinga, R., Schmaal, L., Penninx, B.W.J., Veltman, D.J., Marquand, A.F., 2020. Controlling for effects of confounding variables on machine learning predictions. In: *bioRxiv*. <https://doi.org/10.1101/2020.08.17.255034> (p. 2020.08.17.255034).
- Dwyer, D.B., Falkai, P., Koutsouleris, N., 2018. Machine learning approaches for clinical psychology and psychiatry. *Annu. Rev. Clin. Psychol.* 14, 91–118. <https://doi.org/10.1146/annurev-clinpsy-032816-045037>.
- Edwards, V.J., Holden, G.W., Felitti, V.J., Anda, R.F., 2003. Relationship between multiple forms of childhood maltreatment and adult mental health in community respondents: results from the adverse childhood experiences study. *Am. J. Psychiatry* 160 (8), 1453–1460. <https://doi.org/10.1176/appi.ajp.160.8.1453>.
- Fernandez-Delgado, M., Cernadas, E., Barro, S., 2014. Do we need hundreds of classifiers to solve real world classification problems? *J. Mach. Learn. Res.* 15 (90), 3133–3181 <https://jmlr.org/papers/volume15/delgado14a/delgado14a.pdf>. <https://jmlr.org/papers/volume15/delgado14a/delgado14a.pdf>.
- Fisher, A., Rudin, C., Dominici, F., 2019. All models are wrong, but many are useful: learning a variable's importance by studying an entire class of prediction models simultaneously. *J. Mach. Learn. Res.* 20 (177), 1–81 <http://jmlr.org/papers/v20/18-760.html>. <http://jmlr.org/papers/v20/18-760.html>.
- Fried, E.I., Nesse, R.M., 2015. Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR\*D study. *J. Affect. Disord.* 172, 96–102. <https://doi.org/10.1016/j.jad.2014.10.010>.
- Garety, P.A., Kuipers, E., Fowler, D., Freeman, D., Bebbington, P.E., 2001. A cognitive model of the positive symptoms of psychosis. *Psychol. Med.* 31 (2), 189–195. <https://doi.org/10.1017/s0033291701003312>.
- Garety, P.A., Bebbington, P., Fowler, D., Freeman, D., Kuipers, E., 2007. Implications for neurobiological research of cognitive models of psychosis: a theoretical paper. *Psychol. Med.* 37 (10), 1377–1391. <https://doi.org/10.1017/S003329170700013X>.
- Gilbert, R., Widom, C.S., Browne, K., Fergusson, D., Webb, E., Janson, S., 2009. Burden and consequences of child maltreatment in high-income countries. *Lancet* 373 (9657), 68–81. [https://doi.org/10.1016/S0140-6736\(08\)61706-7](https://doi.org/10.1016/S0140-6736(08)61706-7).
- Goltermann, J., Meinert, S., Hülsmann, C., Dohm, K., Grotegerd, D., Redlich, R., Waltemate, L., Lemke, H., Thiel, K., Mehler, D.M.A., Enneking, V., Borgers, T., Repple, J., Gruber, M., Winter, N., Hahn, T., Brosch, K., Meller, T., Ringwald, K.G., Dannlowski, U., 2021. Temporal stability and state-dependence of retrospective self-reports of childhood maltreatment in major depression: a two-year longitudinal analysis of the childhood trauma questionnaire. In: *bioRxiv*. <https://doi.org/10.1101/2021.08.31.21262884>.
- Haidl, T.K., Hedderich, D.M., Rosen, M., Kaiser, N., Seves, M., Lichtenstein, T., Penzel, N., Wenzel, J., Kambeitz-Ilanovic, L., Ruef, A., Popovic, D., Schultze-Lutter, F., Chisholm, K., Upthegrove, R., Salokangas, R.K.R., Pantelis, C., Meisenzahl, E., Wood, S.J., Brambilla, P., Koutsouleris, N., 2021. The non-specific nature of mental health and structural brain outcomes following childhood trauma. *Psychol. Med.* 1–10 <https://doi.org/10.1017/S0033291721002439>.
- Hardt, J., Rutter, M., 2004. Validity of adult retrospective reports of adverse childhood experiences: review of the evidence. *J. Child Psychol. Psychiatry Allied Discip.* 45 (2), 260–273. <https://doi.org/10.1111/j.1469-7610.2004.00218.x>.
- Hart, S.N., Glaser, D., 2011. Psychological maltreatment: maltreatment of the mind: a catalyst for advancing child protection toward proactive primary prevention and promotion of personal well-being. *Child Abuse Negl.* 35 (10), 758–766. <https://doi.org/10.1016/j.chiabu.2011.06.002>.
- Hart, S.N., Bingle, N.J., Brassard, M.R., 1997. Evidence for the effects of psychological maltreatment. *J. Emot. Abus.* 1 (1), 27–58. [https://doi.org/10.1300/J135v01n01\\_03](https://doi.org/10.1300/J135v01n01_03).
- Hoppen, T.H., Chalder, T., 2018. Childhood adversity as a transdiagnostic risk factor for affective disorders in adulthood: A systematic review focusing on biopsychosocial moderating and mediating variables. *Clin. Psych. Rev.* 65, 81–151. <https://doi.org/10.1016/j.cpr.2018.08.002>.
- Hothorn, T., Hornik, K., van de Wiel, M., Zeileis, A., 2008. Implementing a class of permutation tests: the coin package. *J. Stat. Softw.* 28 (8), 1–23. <https://doi.org/10.18637/jss.v028.i08>.
- Humphreys, K.L., LeMoult, J., Wear, J.G., Piersiak, H.A., Lee, A., Gotlib, I.H., 2020. Child maltreatment and depression: a meta-analysis of studies using the Childhood Trauma Questionnaire. *Child Abuse Negl.* 102, 104361 <https://doi.org/10.1016/j.chiabu.2020.104361>.
- Justice, A.C., Covinsky, K.E., Berlin, J.A., 1999. Assessing the generalizability of prognostic information. *Ann. Intern. Med.* 130 (6), 515–524. <https://doi.org/10.7326/0003-4819-130-6-199903160-00009>.
- Kendler, K.S., 2019. From many to one to many—the search for causes of psychiatric illness. *JAMA Psychiatry* 76 (10), 1085–1091. <https://doi.org/10.1001/jamapsychiatry.2019.1200>.
- Kendler, K.S., Aggen, S.H., 2017. Symptoms of major depression: their stability, familiarity, and prediction by genetic, temperamental, and childhood environmental risk factors. *Depression Anxiety* 34 (2), 171–177. <https://doi.org/10.1002/da.22591>.
- Kessler, R.C., Davis, C.G., Kendler, K.S., 1997. Childhood adversity and adult psychiatric disorder in the US National Comorbidity Survey. *Psychol. Med.* 27 (5), 1101–1119. <https://doi.org/10.1017/S0033291797005588>.
- Koutsouleris, N., Dwyer, D.B., Degenhardt, F., Maj, C., Urquijo-Castro, M.F., Sanfelici, R., Popovic, D., Oeztuerk, O., Haas, S.S., Weiske, J., Ruef, A., Kambeitz-Ilanovic, L., Antonucci, L.A., Neufang, S., Schmidt-Kraepelin, C., Ruhrmann, S., Penzel, N., Kambeitz, J., Haidl, T.K., Writing Group for the PRONIA Consortium, 2020. Multimodal machine learning workflows for prediction of psychosis in patients with clinical high-risk syndromes and recent-onset depression. *JAMA Psychiatry*. <https://doi.org/10.1001/jamapsychiatry.2020.3604>.
- Lang, M., Binder, M., Richter, J., Schratz, P., Pfisterer, F., Coors, S., Au, Q., Casalicchio, G., Kotthoff, L., Bischl, B., 2019. mlr3: a modern object-oriented machine learning framework in R. *J. Open Source Softw.* 4 (44), 1903. <https://doi.org/10.21105/joss.01903>.
- Lewis, S.J., Arseneault, L., Caspi, A., Fisher, H.L., Matthews, T., Moffitt, T.E., Odgers, C. L., Stahl, D., Teng, J.Y., Danese, A., 2019. The epidemiology of trauma and post-traumatic stress disorder in a representative cohort of young people in England and Wales. *Lancet Psychiatry* 6 (3), 247–256. [https://doi.org/10.1016/S2215-0366\(19\)30031-8](https://doi.org/10.1016/S2215-0366(19)30031-8).
- Lindert, J., von Ehrenstein, O.S., Grashow, R., Gal, G., Braehler, E., Weisskopf, M.G., 2014. Sexual and physical abuse in childhood is associated with depression and anxiety over the life course: Systematic review and meta-analysis. *Int. J. Public Health* 59 (2), 359–372. <https://doi.org/10.1007/s00038-013-0519-5>.
- LoPilato, A.M., Addington, J., Bearden, C.E., Cadenhead, K.S., Cannon, T.D., Cornblatt, B.A., Mathalon, D.H., McGlashan, T.H., Perkins, D.O., Tsuang, M.T., Woods, S.W., Walker, E.F., 2019. Stress perception following childhood adversity: unique associations with adversity type and sex. *Dev. Psychopathol.* 1–14 <https://doi.org/10.1017/S0954579419000130>.
- McCrory, E.J., Gerin, M.I., Viding, E., 2017. Annual research review: childhood maltreatment, latent vulnerability and the shift to preventative psychiatry - the contribution of functional brain imaging. *J. Child Psychol. Psychiatry Allied Discip.* 58 (4), 338–357. <https://doi.org/10.1111/jcpp.12713>.
- McLaughlin, K.A., Sheridan, M.A., Lambert, H.K., 2014. Childhood adversity and neural development: deprivation and threat as distinct dimensions of early experience. *Neurosci. Biobehav. Rev.* 47, 578–591. <https://doi.org/10.1016/j.neubiorev.2014.10.012>.
- McLaughlin, K.A., Colich, N.L., Rodman, A.M., Weissman, D.G., 2020. Mechanisms linking childhood trauma exposure and psychopathology: a transdiagnostic model of risk and resilience. *BMC Med.* 18 (1), 96. <https://doi.org/10.1186/s12916-020-01561-6>.
- Meehan, A.J., Baldwin, J.R., Lewis, S.J., MacLeod, J.G., Danese, A., 2021. Poor individual risk classification from adverse childhood experiences screening. *Am. J. Prev. Med.* <https://doi.org/10.1016/j.amepre.2021.08.008>.
- Molnar, C., 2019. Interpretable machine learning. In: *A Guide for Making Black Box Models Explainable* <https://christophm.github.io/interpretable-ml-book/>. <https://christophm.github.io/interpretable-ml-book/>.
- Musliner, K.L., Seifuddin, F., Judy, J.A., Pirooznia, M., Goes, F.S., Zandi, P.P., 2015. Polygenic risk, stressful life events and depressive symptoms in older adults: a polygenic score analysis. *Psychol. Med.* 45 (8), 1709–1720. <https://doi.org/10.1017/S0033291714002839>.
- Olson, M., Blanco, C., Wang, S., Laje, G., Correll, C.U., 2014. National trends in the mental health care of children, adolescents, and adults by office-based physicians. *JAMA Psychiatry* 71 (1), 81–90. <https://doi.org/10.1001/jamapsychiatry.2013.3074>.
- Probst, P., Wright, M.N., Boulesteix, A.-L., 2019. Hyperparameters and tuning strategies for random forest. *WIREs Data Min. Knowl. Discovery* 9 (3), 1–15 <https://onlinelibrary.wiley.com/doi/abs/10.1002/widm.1301>.
- R Development Core Team, 2021. R: A Language And Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria <https://www.r-project.org/>.
- Radloff, L.S., 1977. The CES-D scale: a self-report depression scale for research in the general population. *Appl. Psychol. Meas.* 1 (3), 385–401. <https://doi.org/10.1177/014662167700100306>.
- Read, J., Mayne, R., 2017. Understanding the long-term effects of childhood adversities: beyond diagnosis and abuse. *J. Child Adolesc. Trauma* 10 (3), 289–297. <https://doi.org/10.1007/s40653-017-0137-0>.

- Read, J., Harper, D., Tucker, I., Kennedy, A., 2018a. Do adult mental health services identify child abuse and neglect? A systematic review. *Int. J. Ment. Health Nurs.* 27 (1), 7–19. <https://doi.org/10.1111/inm.12369>.
- Read, J., Harper, D., Tucker, I., Kennedy, A., 2018b. How do mental health services respond when child abuse or neglect become known? A literature review. *Int. J. Ment. Health Nurs.* 27 (6), 1606–1617. <https://doi.org/10.1111/inm.12498>.
- Reuben, A., Moffitt, T.E., Caspi, A., Belsky, D.W., Harrington, H., Schroeder, F., Hogan, S., Ramrakha, S., Poulton, R., Danese, A., 2016. Lest we forget: comparing retrospective and prospective assessments of adverse childhood experiences in the prediction of adult health. *J. Child Psychol. Psychiatry Allied Discip.* 57 (10), 1103–1112. <https://doi.org/10.1111/jcpp.12621>.
- Rosen, M., Betz, L.T., Schultze-Lutter, F., Chisholm, K., Haidl, T.K., Kambeitz-Ilankovic, L., Bertolino, A., Borgwardt, S., Brambilla, P., Lencer, R., Meisenzahl, E., Ruhrmann, S., Salokangas, R.K.R., Upthegrove, R., Wood, S.J., Koutsouleris, N., Kambeitz, J., PRONIA Consortium, 2021. Towards clinical application of prediction models for transition to psychosis: a systematic review and external validation study in the PRONIA sample. *Neurosci. Biobehav. Rev.* 125, 478–492. <https://doi.org/10.1016/j.neubiorev.2021.02.032>.
- Rosenbusch, H., Soldner, F., Evans, A.M., Zeelenberg, M., 2021. Supervised machine learning methods in psychology: a practical introduction with annotated R code. *Soc. Personal. Psychol. Compass* 15 (2). <https://doi.org/10.1111/spc3.12579>.
- Ryff, C.D., Seeman, T., Weinstein, M., 2019. Midlife in the United States (MIDUS 2): Biomarker Project, 2004–2009 [Data set]. Inter-university Consortium for Political and Social Research [distributor], Ann Arbor, MI. <https://doi.org/10.3886/ICPSR29282.v9>.
- Salokangas, R.K.R., From, T., Luutonen, S., Hietala, J., 2018. Adverse childhood experiences leads to perceived negative attitude of others and the effect of adverse childhood experiences on depression in adulthood is mediated via negative attitude of others. *Eur. Psychiatry* 54, 27–34. <https://doi.org/10.1016/j.eurpsy.2018.06.011>.
- Salokangas, R.K.R., Luutonen, S., Heinimaa, M., From, T., Hietala, J., 2019. A study on the association of psychiatric diagnoses and childhood adversities with suicide risk. *Nord. J. Psychiatry* 73 (2), 125–131. <https://doi.org/10.1080/08039488.2018.1493748>.
- Salokangas, R.K.R., Schultze-Lutter, F., Schmidt, S.J., Pesonen, H., Luutonen, S., Patterson, P., Graf von Reventlow, H., Heinimaa, M., From, T., Hietala, J., 2020. Childhood physical abuse and emotional neglect are specifically associated with adult mental disorders. *J. Ment. Health* 29 (4), 376–384. <https://doi.org/10.1080/09638237.2018.1521940>.
- Sheridan, M.A., McLaughlin, K.A., 2014. Dimensions of early experience and neural development: deprivation and threat. *Trends Cogn. Sci.* 18 (11), 580–585. <https://doi.org/10.1016/j.tics.2014.09.001>.
- Spertus, I.L., Yehuda, R., Wong, C.M., Halligan, S., Seremetis, S.V., 2003. Childhood emotional abuse and neglect as predictors of psychological and physical symptoms in women presenting to a primary care practice. *Child Abuse Negl.* 27 (11), 1247–1258. <https://doi.org/10.1016/j.chiabu.2003.05.001>.
- Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinforma.* 8, 25. <https://doi.org/10.1186/1471-2105-8-25>.
- Thorp, J.G., Campos, A.I., Grotzinger, A.D., Gerring, Z.F., An, J., Ong, J.-S., Wang, W., Shringarpure, S., Byrne, E.M., MacGregor, S., Martin, N.G., Medland, S.E., Middeldorp, C.M., Derks, E.M., 2021. Symptom-level modelling unravels the shared genetic architecture of anxiety and depression. *Nat. Hum. Behav.* 1–11. <https://doi.org/10.1038/s41562-021-01094-9>.
- Tiemeier, H., 2020. Commentary: Splitting and lumping brain and childhood adversity measures - a commentary on Gheorghie, Li, Gallacher, and Bauermeister (2020). *J. Child Psychol. Psychiatry Allied Discip.* <https://doi.org/10.1111/jcpp.13346>.
- Touw, W.G., Bayjanov, J.R., Overmars, L., Backus, L., Boekhorst, J., Wels, M., van Hijum, S.A.F.T., 2013. Data mining in the Life Sciences with Random Forest: a walk in the park or lost in the jungle? *Brief. Bioinform.* 14 (3), 315–326. <https://doi.org/10.1093/bib/bbs034>.
- Weinstein, M., Ryff, C.D., Seeman, T.E., 2019. Midlife in the United States (MIDUS Refresher): Biomarker Project, 2012–2016 [Data set]. Inter-university Consortium for Political and Social Research [distributor], Ann Arbor, MI. <https://doi.org/10.3886/ICPSR36901.v6>.
- Widom, C.S., 2020. Commentary: A challenge for a higher bar in research on childhood trauma - reflections on Danese (2020) [Review of Commentary: A challenge for a higher bar in research on childhood trauma - reflections on Danese (2020)]. *J. Child Psychol. Psychiatry Allied Discip.* 61 (3), 251–254. <https://doi.org/10.1111/jcpp.13195>.
- Wright, M.N., Ziegler, A., 2015. ranger: a fast implementation of random forests for high dimensional data in C++ and R. In: arXiv [stat.ML]. arXiv. <https://doi.org/10.18637/jss.v077.i01> <https://onlinelibrary.wiley.com/doi/abs/10.1002/widm.1301>.
- Yarkoni, T., Westfall, J., 2017. Choosing prediction over explanation in psychology: lessons from machine learning. *Perspect. Psychol. Sci.* 12 (6), 1100–1122. <https://doi.org/10.1177/1745691617693393>.
- Zeanah, C.H., Humphreys, K.L., 2018. Child abuse and neglect. *J. Am. Acad. Child Adolesc. Psychiatry* 57 (9), 637–644. <https://doi.org/10.1016/j.jaac.2018.06.007>.