*Psychological Medicine*

**cambridge.org/psm**

**Original Article**

*These two authors contributed equally to this work.

**Cite this article:** Bertie L-A *et al* (2024). Predicting remission following CBT for childhood anxiety disorders: a machine learning approach. *Psychological Medicine* **54**, 4612–4622. https://doi.org/10.1017/S0033291724002654

Received: 15 March 2024
Revised: 22 August 2024
Accepted: 30 September 2024
First published online: 17 December 2024

**Keywords:**
childhood anxiety; cognitive behavior therapy; machine learning; risk prediction

**Corresponding author:**
Jennifer L. Hudson;
Email: jennie.hudson@blackdog.org.au



# Predicting remission following CBT for childhood anxiety disorders: a machine learning approach

Lizel-Antoinette Bertie[1,2,*], Juan C. Quiroz[3,4,*], Shlomo Berkovsky[4], Kristian Arendt[5], Susan Bögels[6], Jonathan R. I. Coleman[7], Peter Cooper[8], Cathy Creswell[8,9], Thalia C. Eley[7], Catharina Hartman[10], Krister Fjermestadt[11], Tina In-Albon[12], Kristen Lavallee[13], Kathryn J. Lester[14], Heidi J. Lyneham[15], Carla E. Marin[16], Anna McKinnon[15], Lauren F. McLellan[15], Richard Meiser-Stedman[17], Maaike Nauta[10], Ronald M. Rapee[15], Silvia Schneider[18], Carolyn Schniering[15], Wendy K. Silverman[16], Mikael Thastum[5], Kerstin Thirlwall[8], Polly Waite[8,9], Gro Janne Wergeland[19], Viviana Wuthrich[15] and Jennifer L. Hudson[1,2]

[1]Black Dog Institute, University of New South Wales, Sydney, NSW, Australia; [2]School of Psychology, UNSW, Sydney, Australia; [3]Center for Big Data Research, UNSW, Sydney, Australia; [4]Centre for Health Informatics, Australian Institute of Health Innovation, Faculty of Medicine, Health and Human Sciences, Macquarie University, Sydney, Australia; [5]Department of Psychology, University of Aarhus, Denmark; [6]Research Institute Child Development and Education, University of Amsterdam, the Netherlands; [7]Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, & King's College London, UK; [8]School of Psychology and Clinical Language Sciences, University of Reading, UK; [9]Departments of Psychiatry and Experimental Psychology, University of Oxford, UK; [10]Department of Psychiatry, University Medical Centre Groningen, University of Groningen, the Netherlands; [11]Department of Psychology, University of Oslo, Norway; [12]Clinical Child and Adolescent Psychology and Psychotherapy, Department of Psychology, University of Koblenz-Landau, Landau, Germany; [13]Institute of Psychology, University of Basel, Switzerland; [14]School of Psychology, University of Sussex, UK; [15]Department of Psychological Sciences, Centre for Emotional Health, Macquarie University, Sydney, NSW, Australia; [16]Yale University, Child Study Center, New Haven, CT, USA; [17]MRC Cognition and Brian Sciences Unit, Cambridge, UK; [18]Mental Health Research and Treatment Center, Ruhr-Universität Bochum, Germany and [19]Department of Clinical Medicine, Faculty of Medicine, University of Bergen, Norway

## Abstract

**Background.** The identification of predictors of treatment response is crucial for improving treatment outcome for children with anxiety disorders. Machine learning methods provide opportunities to identify combinations of factors that contribute to risk prediction models.
**Methods.** A machine learning approach was applied to predict anxiety disorder remission in a large sample of 2114 anxious youth (5–18 years). Potential predictors included demographic, clinical, parental, and treatment variables with data obtained pre-treatment, post-treatment, and at least one follow-up.
**Results.** All machine learning models performed similarly for remission outcomes, with AUC between 0.67 and 0.69. There was significant alignment between the factors that contributed to the models predicting two target outcomes: remission of all anxiety disorders and the primary anxiety disorder. Children who were older, had multiple anxiety disorders, comorbid depression, comorbid externalising disorders, received group treatment and therapy delivered by a more experienced therapist, and who had a parent with higher anxiety and depression symptoms, were more likely than other children to still meet criteria for anxiety disorders at the completion of therapy. In both models, the absence of a social anxiety disorder and being treated by a therapist with less experience contributed to the model predicting a higher likelihood of remission.
**Conclusions.** These findings underscore the utility of prediction models that may indicate which children are more likely to remit or are more at risk of non-remission following CBT for childhood anxiety.

## Introduction

Anxiety disorders are among the most common mental health problems in children and adolescents and have a range of adverse consequences across several developmental, familial, social, and school domains (Asselmann, Wittchen, Lieb, & Beesdo-Baum, 2018; Sicouri, Perkes, & Hudson, 2022). At present, cognitive behavioral therapy (CBT) is the first line intervention recommended for childhood anxiety disorders, with on average 5 out of 10 children demonstrating remission at post-treatment (James, Reardon, Soler, James, & Creswell, 2020). It has become crucial to identify risk factors that may indicate which children are more or less likely to respond to treatment, that is, identify factors that predict outcome regardless of the treatment received (Kraemer, Wilson, Fairburn, & Agras, 2002). Identifying individual risk is considered an important step towards personalized intervention and may assist clinicians when making treatment decisions (Bertie & Hudson, 2021; Ozomaro, Wahlestedt, & Nemeroff, 2013).

A problem for the field is the lack of consistency in which predictors (potential risk factors) that have been examined across the multiple trials that have been conducted over the last decades. Existing evidence suggests that a diagnosis of social anxiety disorder, comorbid depression, and parent psychopathology are the most robust baseline predictors of poorer treatment outcomes (Hudson et al., 2015). A recent review and meta-analysis of predictors of child anxiety and depression concluded that higher anxiety symptom severity predicted negative CBT outcomes (Kunas, Lautenbacher, Lueken, & Hilbert, 2021). However, questions remain regarding the potential effects of several potential additional predictors, ranging from simple demographics (e.g. gender), pretreatment clinical factors (e.g. comorbid externalizing disorders), and treatment factors (e.g. format) on anxiety treatment outcomes. The inconsistent pattern of results may implicate limitations inherent to searching for a single predictor to identify which children are at risk of poor treatment outcomes and the potential utility of a combination of factors instead (Wehry, Beesdo-Baum, Hennelly, Connolly, & Strawn, 2015).

In addition to focusing on a variety of divergent risk factors, investigations into the diagnostic complexity involved in predicting likelihood of response or remission following CBT have been hindered mainly by small samples, varied methodology, and traditional statistical methods (Beam & Kohane, 2018). This has led to increased incentive for the use of predictive analytics, such as machine learning methods in mental health research (Bertie & Hudson, 2021), which confers several advantages relative to the linear models typically employed (DeRubeis, 2019). For instance, machine learning can integrate larger sets of multidimensional psychological data to build multivariate models that increase predictive ability, can capture complex and non-linear relationships in the data, handle large datasets more effectively, and are better suited for personalized predictions (Coutanche & Hallion, 2020; LeCun, Bengio, & Hinton, 2015; Scott, 2021). A systematic review of 16 studies that predicted adult anxiety outcomes found increased use of several machine learning methods with predictive accuracy ranging from 73% to 93% (Pintelas, Kotsilieris, Livieris, & Pintelas, 2018).

The present study aimed to build on previous predictor work that identified more robust but single predictors of treatment outcome. Using a large, archived dataset from multiple child anxiety treatment trials, the goal of the current exploratory work was to develop machine learning models that predicted remission from anxiety disorders to demonstrate its feasibility for use by clinicians when assessing children and planning treatment for childhood anxiety based on the combination of factors and estimate the likely effectiveness of such a potential tool in clinical practice. The present study took advantage of the large, combined sample size and the inclusion of multiple potential predictors to examine the role and combination of impactful factors. This study used a 'data-driven' approach (Cheng & Phillips, 2014), which relies on all available predictor data for analysis. Therefore, the models included demographic (i.e. age, gender) and clinical predictors (i.e. severity, type of anxiety disorder, and comorbidity). Treatment variables (i.e. modality, therapist experience) were also included in the model. Given the exploratory nature of this study, the combination of specific factors that would predict remission is not hypothesized.

### Transparency and openness

This study involved an analysis of existing data rather than new data collection. We report on how we determined the sample under investigation, eligibility criteria, and all measures in this secondary analysis of existing data. Data were analyzed using Python 3.9. This study's design and its analyses were not preregistered. The code and all used packages are available here: https://github.com/juancq/remission-cbt. Further site and treatment information, as well as details relevant to the assessment of risk of bias of the developed prognostic prediction models (Wolffet al., 2019) are included in online Supplementary material.

## Methods

### Sample

The sample comprised data from 2114 children collected from 11 global sites: Sydney, Australia ($n = 1199$); Reading, UK ($n = 342$); Aarhus, Denmark ($n = 120$); Bergen, Norway ($n = 156$); Bochum, Germany ($n = 82$); Basel, Switzerland ($n = 87$); Groningen, the Netherlands ($n = 41$); Amsterdam, the Netherlands ($n = 7$); Oxford, United Kingdom ($n = 28$); Cambridge, UK ($n = 3$); and Miami, Florida ($n = 49$). Participants met inclusion criteria if (a) they were assigned a primary diagnosis of a DSM-IV anxiety disorder (APA, 2003), assigned at the individual site after a semi-structured diagnostic interview, and (b) received a course of manualized individual or group CBT for anxiety disorders.

### Measures

#### Anxiety outcomes

The two primary outcomes were remission of (1) all anxiety disorders and (2) the primary anxiety disorder. These remission outcomes served as the target variables that the machine learning models predicted. Remission was defined as clinician-assessed sub-clinical anxiety status at any point during the follow-up period, for both the primary anxiety disorder and all anxiety disorder diagnoses Anxiety status was measured by the Anxiety Disorders Interview Schedule Child and Parent Version (ADIS-IV-C/P) (Silverman & Albano, 1996), a semi-structured clinical interview administered to both parents and children that assessed baseline severity and diagnosis based on a composite report. Impairment was operationalized as clinician-rated severity ratings (CSR) and was rated per disorder on a scale of 0–8 and according to DSM criteria, with a diagnosis made when a CSR score of 4 or more was assigned. All sites assigned diagnoses according to the ADIS-IV-C/P except for two (Bochum and Basel), where a diagnostically comparable measure, the Kinder-DIPS for DSM-IV-TR (Diagnostisches Interview bei psychischen Störungen im Kindes-und Jugendalter or Diagnostic interview for mental disorders for children and adolescents) (Schneider, Unnewehr, & Margraf, 2009) was used. Assessments were completed at pre-treatment, immediately post-treatment, and at least once more at three months, six months, or one-year post-treatment (one-year follow-up). Assessments were mainly completed by graduate

https://doi.org/10.1017/S0033291724002654 Published online by Cambridge University Press

assistants or clinical psychologists who were trained in the administration of the relevant instruments. An assessment of remission at any point during follow-up was coded as a binary variable (1 representing remission at any point, 0 representing no remission during follow-up), for each of the target outcome models.

#### Demographic and clinical variables

The models contained several parent-reported demographic and clinical baseline factors including age, gender, and ethnicity for the child, and parental age and family structure. Comorbid mood (depression and dysthymia) and externalizing disorders (oppositional defiant disorder, conduct disorder, or attention-deficit/hyperactivity disorder (ADHD)) were assessed using ADIS-IV-C/P or Kinder/DIPS-C/P at ten sites except for Bergen. In Bergen, the Development and Well-being Assessment (DAWBA: Goodman, Ford, Richards, Gatward, & Meltzer, 2000) was used. Other self-report measures for children and their parents were included which have all shown satisfactory reliability and validity. These included the Spence Children's Anxiety Scale (SCAS: Spence, 1998), child and parent versions (SCAS-C, SCAS-P), which assessed child anxiety symptoms. The Short Mood and Feelings Questionnaire (SMFQ: Messer, Angold, Costello, & Loeber, 1995), child and parent versions, were used to assess child's depressive symptoms. The Strengths and Difficulties Questionnaire (SDQ, Goodman, 1997), addressed the child's externalizing problems and functional impairment. Finally, eight sites used the Depression Anxiety Stress Scale (DASS: Lovibond and Lovibond, 1995), to assess parental self-rated depression, anxiety, and stress.

#### Treatment variables

All participants received CBT categorized as follows: individual CBT (with or without a parent; $n = 1075$, 51%) for individual children or family dyads, or group-based CBT (49%) for children ($n = 263$) or families ($n = 776$). The present study relied on archived data and as such, treatment variables of intensity, parental involvement, and therapist experience were coded according to trial specific information mostly recorded at the general study level. Treatment intensity was coded as low (i.e. guided self-help with therapist intervention), medium (i.e. standard face-to-face intervention), and high (i.e. increased number of sessions or medication intervention). In the present sample parental involvement was coded as low involvement, active involvement with high contingency management (CM) or strategies for transfer of control (TC), and active involvement with low CM/TC (Manassis et al., 2014). Therapist experience was coded as low (i.e. therapy delivered by trainees or students), medium (therapy delivered by novice therapists), and high (experienced therapists). This variable was not coded based on individual therapists but rather the level of experience of therapists delivering the allocated treatment condition. Although treatment delivery formats varied across studies, treatment variables were coded according to the study protocols which were comparable for core elements of CBT (i.e. coping skills, cognitive restructuring, and exposure), and have been reported in prior subgroup studies (Coleman et al., 2016; Jennifer L Hudson et al., 2015; McKinnon et al., 2018).

### Machine learning methodology

Our machine learning approach followed established practice (Dwyer, Falkai, & Koutsouleris, 2018) including feature selection, training a model on a subset of data, evaluating the model on unseen data, and selecting metrics to assess predictive performance and generalizability. Three machine learning algorithms were used to predict remission of all anxiety disorders and the primary anxiety disorder (Fig. 1) from the demographic, clinical, and treatment variables: logistic regression (LR), a standard baseline model for binary classification tasks; gradient boosted trees (LGBM), an efficient algorithm for structured data; and neural oblivious decision ensembles (NODE), a deep learning model that performs well on structured data (Popov, Morozov, & Babenko, 2019). Using multiple models allows for a robust comparison of predictive performance across different algorithmic approaches, each with its own strength. Logistic regression used the implementation from the scikit-learn library for Python (Pedregosa et al., 2011), gradient boosted trees used the LightGBM library for Python (Ke et al., 2017), and NODE used the Pytorch Tabular implementation (Joseph, 2021). As such, three machine learning models were trained for each of the two target outcomes, wherein both outcomes were binary. The total number of variables used for each model was 34. Missing values for continuous variables were imputed using a multi-variate imputer that estimated each variable from all the other variables. Missing values for categorical variables were imputed with the most-frequent value.

Models were evaluated using a train-validation-test split, with this process repeated 30 times (10-fold cross-validation, repeated three times). Metrics from the cross-validation are presented as mean and standard deviation over the 30 runs. The metrics included area under the receiving operating characteristic curve (ROC), area under the precision-recall curve (PR-ROC), and expected calibration error (ECE) (Géron, 2022). ROC and PR-ROC measure the discriminative performance of a predictive model at various classification thresholds. The ROC curve is generated by plotting the sensitivity against the false positive rate. The PR-ROC curve is generated by plotting precision (positive predictive value) against sensitivity, a useful metric when the outcomes of interest in the dataset are imbalanced (i.e. number of remissions much lower than non-remissions). Calibration of a model refers to the reliability of the probability predictions generated by a model, as probabilities of experiencing an outcome are more clinically informative than a binary outcome of remission/non-remission. LR and LGBM were calibrated using a validation set with post-hoc calibration using the isotonic method. NODE probabilities were not recalibrated with post-hoc calibration. Calibration was assessed using ECE and reliability diagrams. Hyperparameters for all models were tuned on the validation set of each fold.

### Model interpretation and utility

We used SHapley Additive exPlanations (SHAP) to identify the contribution of each variable to the remission predictions made by the trained machine learning models (Lundberg & Lee, 2017). The results are visualized with SHAP summary plots showing the magnitude and direction of variable attributions to the prediction of anxiety outcomes. Each dot represents an individual in the dataset and the density represents multiple dots at the same position. The position of the dot on the x-axis represents the impact of a particular variable value on the model predicting remission, that is, negative values on the x-axis contribute to a lower likelihood of remission prediction and positive values on the x-axis contribute to a higher likelihood of remission prediction. The variables are ordered by mean impact on the predictive ability of the model and not by individual variable significance. The colors represent the variable values, with red indicating high and blue indicating

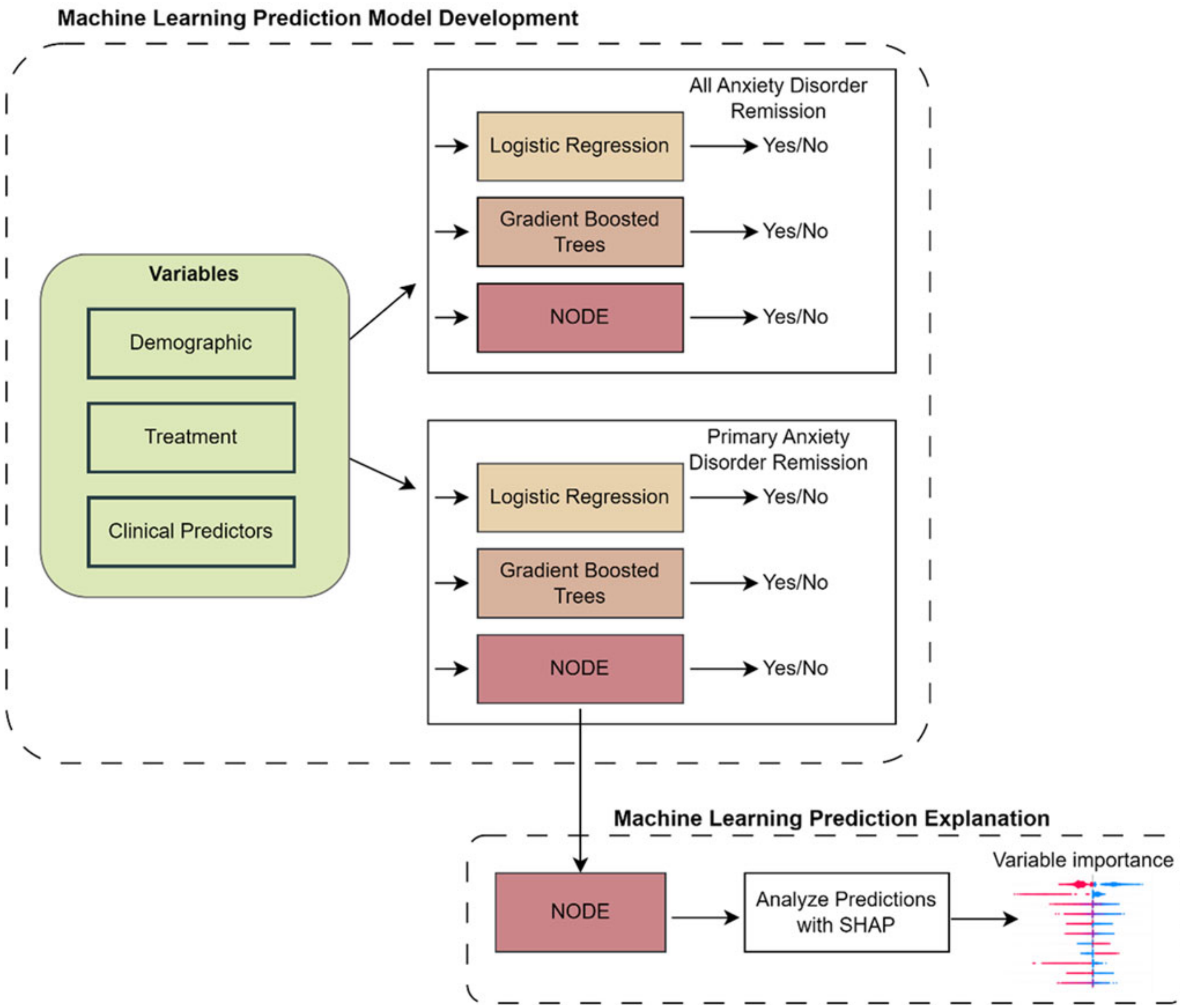


**Figure 1.** Machine learning prediction model development and explanation.

low. Explanation was restricted to the 20 variables with the highest mean impact for readability purposes. Results will be discussed by describing variables in terms of demographic, clinical, or treatment variables having an impact on model prediction.

## Results

### Sample characteristics

Participants were 5–18 years of age ($M = 10$, S.D. = 2.4), balanced by gender (female = 52%, male = 48%), and overall contained more children than adolescents (5–12 years = 84%). The demographic characteristics of the sample are presented in Table 1, and the treatment details are presented in Table 2.

### Predicting remission

#### Model performance

Table 3 shows the performance of the prognostic models predicting all anxiety disorder remission and remission of the primary anxiety disorder. The NODE model achieved the highest discriminative performance, with a balanced precision and recall demonstrated by the high PR-ROC. The performance of LR and LGBM was similar, but NODE also had the lowest calibration error.

Figure 2 shows the calibration achieved by the NODE model for both primary and all anxiety disorder remission predictions. The model probabilities predicting remission from all anxiety disorders are well calibrated. For primary anxiety disorder remission, the model is not as well calibrated because smaller predictions under-predict the actual likelihood of remission, whereas the larger predictions show good calibration (probabilities > = 0.4 show good calibration, but probabilities < 0.4 are not as trustworthy).

#### Remission from all anxiety disorders

Post-hoc explanations of the NODE model predicting remission of all anxiety disorders are presented in Fig. 3. The SHAP summary plot indicates the factors contributing to the model predicting remission from all anxiety disorders. Considering the direction and magnitude of each feature contribution, child age was one demographic factor that contributed to the model

https://doi.org/10.1017/S0033291724002654 Published online by Cambridge University Press

**Table 1.** Total sample demographic and clinical characteristics

| Factor | | | |
|---|---|---|---|
| Child demographic | | | |
| Age *M* (s.d.) | | 10.0 | (2.4) |
| Children *n* (%) | | 1779 | 85% |
| Adolescents *n* (%) | | 316 | 15% |
| Gender female (%) | | 1092 | 52% |
| Ethnicity | | | |
| White European | | 1330 | 63% |
| Other | | 169 | 8% |
| Undisclosed | | 615 | 29% |
| Family demographic | | | |
| Mother age *M* (s.d.) | | 41.6 | (5.0) |
| Father age *M* (s.d.) | | 43.6 | (6.4) |
| Clinical (ADIS-IV-C/P; Kinder-DIPS-C/P; DAWBA) | | | |
| Primary diagnosis CSR *M* (s.d.) | | 6.2 | (1.0) |
| Primary diagnosis *n* (%) | | | |
| Generalized anxiety | | 813 | 38% |
| Separation anxiety | | 480 | 23% |
| Social anxiety | | 470 | 22% |
| Specific phobia | | 212 | 10% |
| Other disorders | | 139 | 7% |
| Symptomatology | | | |
| Anxiety (SCAS) *M* (s.d.) | | | |
| Mother reported | *n* = 1859 | 35.9 | (14.6) |
| Father reported | *n* = 1267 | 31.2 | (13.5) |
| Child reported | *n* = 1871 | 35.9 | (17.4) |
| Mood (SMFQ) *M* (s.d.) | | | |
| Mother reported | *n* = 1468 | 7.4 | (5.6) |
| Father reported | *n* = 740 | 6.1 | (4.9) |
| Child reported | *n* = 1554 | 7.0 | (5.5) |
| Externalizing (SDQ) *M* (s.d.) | | | |
| Mother reported | *n* = 1689 | 15.1 | (6.0) |
| Father reported | *n* = 1178 | 14.1 | (6.1) |
| Child reported | *n* = 1640 | 14.1 | (5.8) |
| Contextual factors | | | |
| Parental psychopathology (DASS) *M* (s.d.) | | | |
| Mother | *n* = 1820 | 25.3 | (18.7) |
| Father | *n* = 1290 | 21.7 | (16.7) |

*Note:* Sample characteristics based on raw data. ADIS-IV-C/P, Anxiety Disorder Interview Schedule Child and Parent Version; KINDER-DIPS, Diagnostic interview for mental disorders for children and adolescents; DAWBA, Development and Well-being Assessment; SCAS, Spence child anxiety scale; SMFQ, Short mood and feelings questionnaire; SDQ, Strengths and difficulties questionnaire. DASS, Depression, anxiety, and stress scale. 'Other disorders' included agoraphobia, panic disorder (with and without agoraphobia), obsessive-compulsive disorder, post-traumatic stress disorder, selective mutism, or unspecified anxiety disorder.

**Table 2.** Treatment characteristics

| Factor | *n* | % |
|---|---|---|
| CBT treatment type *n* (%) | | |
| Individual | 1075 | 51% |
| Group | 263 | 12% |
| Family | 776 | 37% |
| CBT treatment intensity *n* (%) | | |
| Low | 289 | 14% |
| Medium | 1752 | 83% |
| High | 73 | 3% |
| Parental involvement | | |
| Low | 408 | 19% |
| Active + low CM/TC | 94 | 4% |
| Active + high CM/TC | 1612 | 76% |
| Therapist experience | | |
| Low: Students/Trainees/Novices | 538 | 25% |
| Medium: mixed experience | 951 | 45% |
| High: experienced | 625 | 30% |

*Note:* CBT, Cognitive Behavioral Therapy; CM/TC, Contingency management and transfer of control.

predicting that older ages was related to a lower likelihood of remission. Clinical factors contributing to a lower likelihood of remission included greater number of anxiety disorders, the presence of a mood disorder, higher physical injury anxiety (mother and child report), higher internalizing symptoms (mother report), higher externalizing symptoms (mother and child report), as well as higher maternal anxiety, depression, and stress. Among the treatment factors, receiving group CBT delivery, treatment with low treatment intensity, and treatment with low parental involvement also contributed to the model predicting lower likelihood of remission. In contrast, the absence of a social anxiety diagnoses and receiving therapy from a therapist with less experience contributed to the model predicting higher probability of remission from all anxiety disorders.

**Table 3.** Cross validation and standard deviation model results by anxiety outcome

| Model | ROC | PR-ROC | ECE |
|---|---|---|---|
| All anxiety disorder remission | | | |
| LR | 0.686 (0.037) | 0.726 (0.038) | 0.066 (0.030) |
| LGBM | 0.683 (0.050) | 0.718 (0.043) | 0.076 (0.045) |
| NODE | 0.694 (0.039) | 0.724 (0.031) | 0.048 (0.015) |
| Primary anxiety disorder remission | | | |
| LR | 0.670 (0.028) | 0.808 (0.033) | 0.063 (0.027) |
| LGBM | 0.660 (0.041) | 0.815 (0.026) | 0.062 (0.026) |
| NODE | 0.680 (0.033) | 0.812 (0.029) | 0.056 (0.021) |

*Note*: ROC, receiving operating characteristic curve; PR-ROC, area under the precision recall curve; ECE, expected calibration error; LR, logistic regression; LGBM, logistic gradient boosted model; NODE, neural oblivious decision ensembles.

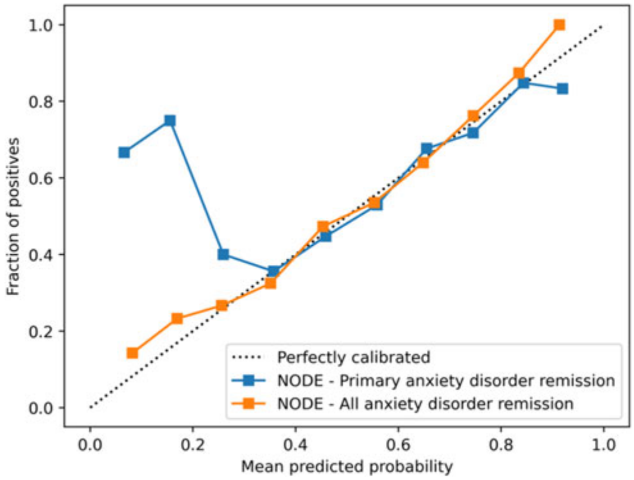


**Figure 2.** Calibration achieved by the NODE model for both primary and all anxiety disorder remission model prediction.

### *Remission from the primary anxiety disorder*

Post-hoc explanations of the NODE model predicting remission of the primary anxiety disorder are presented in Fig. 4. Similar to the previous model, older child age contributed to the model predicting lower likelihood of remission. Clinical factors contributing to a lower likelihood of remission in this model included the presence of both a mood disorder and an externalizing disorder, as well as higher social anxiety (mother report), higher physical injury anxiety (mother and child report), higher internalizing symptoms (mother report), and higher externalizing symptoms (mother and child report). Among the treatment factors in this model, group CBT delivery and low treatment intensity contributed to the model predicting lower likelihood of remission. Like the previous model, the absence of a social anxiety diagnoses and receiving therapy from a therapist with less experience contributed to the model predicting higher probability of remission from the primary anxiety disorder.

## Discussion

### *Key findings – model*

In this exploratory study, we applied machine learning techniques to develop prognostic models for predicting childhood anxiety disorder remission after receiving CBT. The NODE deep learning model had the best discriminative performance and the best calibration amongst the three evaluated algorithms. It was anticipated that in the present study, an informative model would have an ROC of 0.5 for both target outcomes and a PR-ROC of 0.55 for all anxiety disorder remission and 0.69 for primary anxiety disorder remission. The results of ROC being less than 0.70 reflects

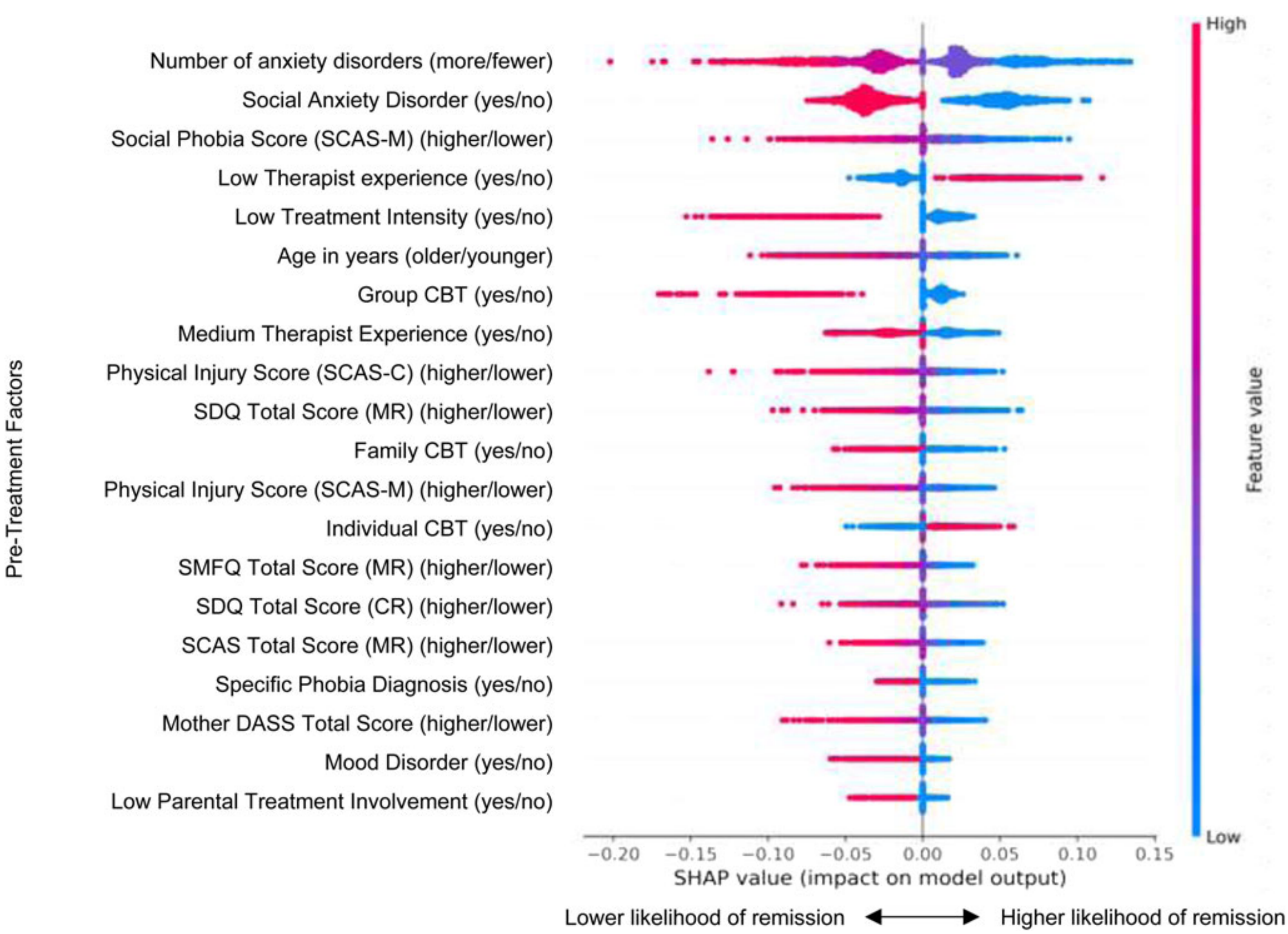


**Figure 3.** Summary plot describing the relationship between the value of the pre-treatment factors and the impact on the predictive model for remission of all anxiety disorders (outcome 1). The top twenty factors were displayed.

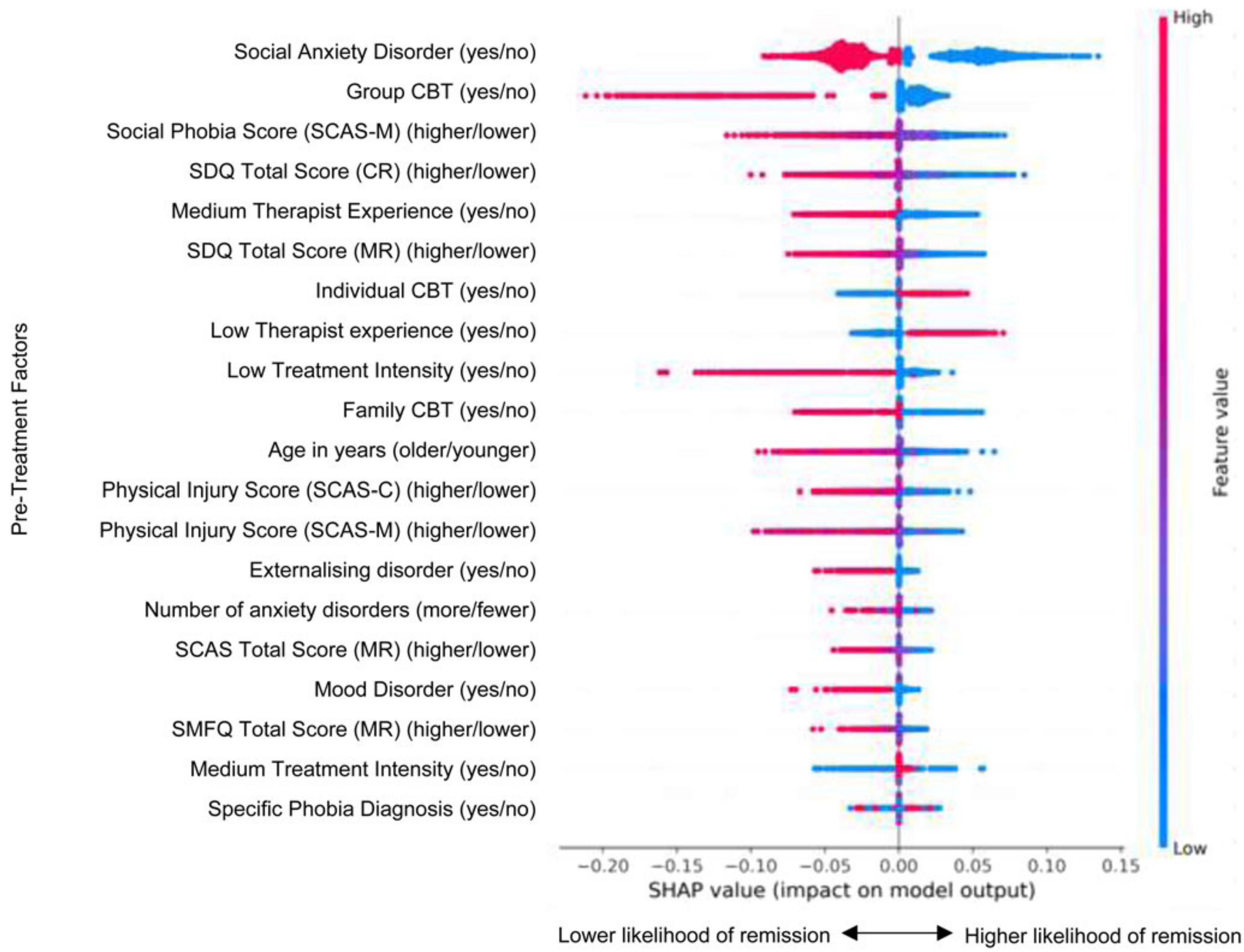


**Figure 4.** Summary plot describing the relationship between the value of the pre-treatment factors and the impact on the remission of primary anxiety disorders (outcome 2). The top twenty factors were displayed.

the complexity of predicting remission after receiving CBT, and is often mischaracterized as representing bad discrimination (Perlis, 2011). Indicative of the range of ROC metrics in machine learning studies, a systematic review of prediction models in child and adolescent mental health reported that the ROC of included studies ranged from 0.57 to 0.99 (Senior, Fanshawe, Fazel, & Fazel, 2021). The ROC of 0.70 and PR-ROC of close to 0.70 demonstrate that the predictions of the models are useful and suggests that additional variables beyond those currently captured in the dataset are necessary to improve prediction performance. Future research should therefore explore the use of different clinical or biological variables for predicting remission.

### *Key findings – predictors*

There was significant alignment between the two predictive models regarding the contributing factors present, with the models sharing 17 of the top 20 contributing factors. Of the demographic factors, older age was associated with model predictions of a lower probability of remission of all anxiety disorders and the primary anxiety disorder. This finding is inconsistent with a previous individual patient data meta-analysis (IPDMA) that reported treatment benefits were comparable between adolescents and younger children (Bennett et al., 2013). However, the IPDMA was based on a smaller sample ($n = 1171$) and our analysis has substantially increased power to detect effects not only because of the larger sample size but also because we examined age as a continuous variable rather than categorical. However, the present study sample contained a small percentage of adolescents ($n = 316$, 15%), the effect of age on treatment outcome predictions requires further investigation into whether delivering interventions during childhood is more effective than during adolescence.

Regarding clinical diagnostic factors, a diagnosis of Social Anxiety Disorder anywhere in the child's profile had the greatest average impact on both model predictions. The absence of a diagnosis of Social Anxiety Disorder contributed to the models predicting a greater likelihood of remission in this sample. This finding aligns with earlier evidence that Social Anxiety Disorder is a consistent and robust predictor of poorer treatment outcome (Hudson et al., 2015). In addition, a recent systematic review and meta-analysis showed that children with a primary disorder of Social Anxiety had a significantly lower post-CBT recovery rate than children diagnosed with other primary anxiety disorders (Evans, Clark, & Leigh, 2021).

As expected, a greater number of diagnosed anxiety disorders predicted a greater likelihood that children would not remit all

https://doi.org/10.1017/S0033291724002654 Published online by Cambridge University Press

their anxiety disorders following treatment, in only the model for remission from all anxiety disorders. There is considerable evidence supporting the notion that targeting a primary anxiety disorder will have a positive effect on other comorbid anxiety disorders (Rapee et al., 2013). Despite this, when child experiences multiple fears and worries, treatment involving a fixed number of sessions reduces the available time allocated to each set of fears or worries. Although generalization across worries is expected to occur, it is likely that children with many feared situations may struggle to tackle all of these fears during treatment. Children with a greater number of anxiety disorders may require additional treatment sessions to ensure that each set of fears and worries has received adequate attention.

The presence of a mood disorder predicted a lower likelihood of remission in both models. This finding is consistent with studies examining treatment outcome based on remission or end state functioning with larger samples (Hudson et al., 2015). Sample size and youth age are both important issues when examining the effects of depression comorbidity on treatment outcome, because children are much less likely to experience depression (Spoelma, Serafimovska, & Parker, 2023) than adolescents and the base rates of comorbid depression in clinical trials of anxiety (sometimes due to exclusion criteria) can be low (e.g.15.2%, see Rapee et al., 2013).

The presence of an externalizing disorder indicated a lower likelihood of remission, but only for the primary anxiety model. Studies have shown that children diagnosed with comorbid ADHD fared worse following CBT compared to children without such comorbidity (Gould, Porter, Lyneham, & Hudson, 2018; Halldorsdottir & Ollendick, 2016). However, in Halldorsdottir et al., a comorbid diagnosis of oppositional defiant disorder showed no such negative impact on treatment outcome. This finding may suggest that the effect of externalizing diagnoses on treatment outcome may depend on the specific disorder under investigation. Further, it may also indicate that the immediate effect of these comorbid diagnoses is related to the most interfering anxiety disorder and does not necessarily impact overall improvement following treatment.

Secondary anxiety measures (SCAS) also contributed to the prediction of remission. More specifically, mother report of child social anxiety contributed to the primary anxiety disorder remission model, whereas both mother and child rated fear of physical injury was a contributing factor in primary and all anxiety remission models. In all instances, these factors indicated that higher anxiety symptom scores predicted lower likelihood of remission. Internalizing symptomatology also contributed to remission prediction in both models, but only for mother ratings. Higher scores for depressive symptoms (SMFQ) were associated with a lower likelihood of remission. Similarly, both mother and child ratings of externalizing symptomatology contributed to remission prediction. In both models, measures of externalizing problems and functional impairment (SDQ) indicated that higher total scores had a strong negative impact on the probability of remission from anxiety disorders.

Regarding parental factors, higher mother self-reported depression, anxiety, and stress (DASS) scores contributed to model predictions of lower likelihood of remission from all anxiety disorders. This finding is supported by a recent systematic review and meta-analysis conclusion that parental psychopathology was a significant predictor of worse CBT outcomes for anxious youth (Kunas et al., 2021). Father psychopathology was not included in our analysis due to the low number of father participants across the trials.

Several treatment factors contributed to remission prediction. In both prediction models, treatment modality made the models more likely to predict a lower likelihood of remission when youth received group child and family CBT, with the mean impact of group child CBT on the model output being greater than family CBT. Conversely in both models, receiving individual CBT contributed to predicting higher likelihood of remission. These predictions are consistent with findings from a meta-analysis showing youth who received group delivery showed worse outcomes than those who received individual therapy (Reynolds, Wilson, Austin, & Hooper, 2012). In contrast, findings from a recent meta-analysis (Guo et al., 2021), concluded that therapy delivered in a group format did not significantly differ from individual delivery for children based on remission outcomes. However, the same study showed individual therapy was more effective than group therapy for adolescents. Despite model prediction differences in treatment modality, these findings do not reveal which children would respond to which treatment modality and future research should focus on identifying significant interactions that may indicate optimal treatment modality for certain subgroups of children.

Low treatment intensity predicted a lower likelihood of remission in both models. In the present sample, low intensity treatments also referred to delivery that required little therapist intervention and was mostly self-directed (i.e. online programs or bibliotherapy). This is consistent with remission rates following digital interventions being lower than traditional face to face therapies (Garrido et al., 2019). Nevertheless, low intensity interventions have important value, particularly when shortages in the mental health workforce reduce access to evidence-based care (Grist, Croker, Denne, & Stallard, 2019).

With regards to therapist experience, across the two prediction models, lower therapist experience was associated with a higher likelihood of remission. Despite this, high level of therapist experience was not a top 20 contributing factor in either predictive model. This is a somewhat surprising finding. In previous explorations of therapist factors within school settings, external clinicians (that is, those with more experience treating mental disorders) deliver better outcomes for student mental health than school staff (Werner-Seidler et al., 2021; Zhang, Wang, & Neitzel, 2023). In line with this, evidence from clinical settings also indicates greater therapist experience predicts improved CBT outcome (Podell et al., 2013), yet this has not always been the case (Cecilione, McLeod, Southam-Gerow, Weisz, & Chorpita, 2021). There are a number of possible explanations for this seemingly surprising result. First, although the current analyses are conducted within a much larger sample size than previous studies, there were only three studies (two sites) that utilized therapists with low experience. It is also possible that student therapists with less experience receive more intensive supervision and have received more recent training. Also, the measurement of therapist experience in this study was coded retrospectively and based on the average therapist delivering therapy within the study condition, rather than on ratings of individual therapists allocated to each child. This effect warrants further investigation.

Finally, low parental involvement contributed to the model predicting lower likelihood of remission from all anxiety disorders but not for primary disorder remission. Earlier research suggests treatment that involve parents and used contingency management (CM) strategies for transfer of control (TC) showed better outcomes than other types or levels of involvement, especially in the longer term (Manassis et al., 2014).

### *Strengths, limitations, and future directions*

A major strength of the present study is the large, combined dataset used to facilitate the innovative machine learning approach to prediction modeling. Although the wide age range is a positive characteristic of the sample because it provides greater heterogeneity, the sample contained a relatively small percentage of adolescents. This restricts the generalization of findings primarily to children. Methodological strengths and limitations are also present. The explanations of the models included the most important variables used by the best performing model (NODE). This does not mean that all other variables had no impact. The SHAP plot post-hoc explanation serves as a proxy for understanding how the model makes predictions given the variables included in the model. It is important to bear in mind that the prognostic models in this framework exploit correlations between variables when calculating relative contribution by single factors and should not be interpreted as causal effects. We also explained the models based on the 20 factors with the highest mean impact on predicted ability of the model, although other factors not listed also impacted remission prediction. Finally, for settings where a fully transparent model is the highest priority for a clinical deployment, our results show that logistic regression achieves competitive results with the other algorithms. If obtaining the most accurate and calibrated algorithm is the top priority, then investing in building and training a deep learning model (such as the NODE algorithm) has the most potential. The definition of treatment success is an important consideration for understanding the differing effects across studies of treatment predictors. As the present study focused on models that predict remission from anxiety disorders, future research could investigate the impact of the top variables on outcomes that include treatment response such as improvement or change following therapy.

### *Clinical implications*

The findings of these models have implications for clinical practice. Compared to earlier research, this study provides a performance evaluation of machine learning models that can predict remission after a standard course of CBT, and which could function as an application used by clinicians when assessing children. The explanation of the variables most impactful on the predictions made by the models can also assist clinicians to determine which children are less likely to respond to a standard course of CBT: older children, children whose mothers have elevated depression and anxiety, children diagnosed with SoAD, as well as a greater number of comorbid anxiety disorders, and the presence of a comorbid mood or externalizing disorder. Rather than determining 'average' effects, to increase clinical utility, understanding the combination of factors in determining idiographic patterns of outcome will be critical to improving the precision of individualized care for children and their families.

Although this takes us one step closer to improving personalization of mental health care for children undergoing treatment for anxiety disorders, limitations to the clinical utility of these findings should be noted. First, as models become more complex and incorporate a greater number of variables, it becomes increasingly difficult to interpret how they arrive at their predictions. Educating and training clinicians on how machine learning models generate predictions, the limitations of current interpretability methods, and understanding model calibration (a measure of uncertainty around a prediction) will be crucial for successful deployment in clinical practice (Hassija et al., 2024). Second, implementation and integration into clinical workflows will require significant resources, including changes to standards of practice and ongoing monitoring and updates to the models to prevent performance degradation. To mitigate algorithmic bias and maintain relevance over time, models should be trained on continuously updated, representative data (Vollmer et al., 2020). Overcoming these limitations will require ongoing evaluation, refinement, and close collaboration between data scientists, clinicians, and key stakeholders to ensure the models are both effective and practical in real-world settings.

## Conclusion

We employed machine learning models using youth demographic, clinical, and treatment factors to predict the likelihood of remitting from a primary or all anxiety disorders. Twenty factors were reported based on their individual predictive contribution to the remission models, with striking overlap in 17 of those variables. These findings may be useful as they consolidate understanding regarding which factors contribute to lower likelihood of remission following treatment for childhood anxiety. Future research is needed to validate the usefulness of these predictive models in clinical practice, as well as their ability to predict more complex idiographic patterns of response.

**Supplementary material.** The supplementary material for this article can be found at https://doi.org/10.1017/S0033291724002654.

**Funding statement.** This work was supported by Australian Research Council grant DP0878609 (J.L.H., Jenny Donald, PhD, R.M.R., T.C.E.); Australian National Health and Medical Research Council grants (1027556: R.M.R., J.L.H., H.J.L., Cathy Mihalopolous, BSc[Hons], PhD, V.W), (488505: H.J.L., J.L.H., R.M.R.), (382008: J.L.H. and R.M.R.), and (1103611 L.F.M., V.W); TrygFonden grant (7-10-1391: M.T. and Esben Hougaard, PhD); Edith og Godtfred Kirk Christiansens Fond grant (21- 5675: M.T.); Swiss National Science Foundation grant (105314-116517: S.S.); Western Norway Regional Health Authority grants (911253: Odd E. Havik, PhD) and (911366: E.H.); National Institute of Mental Health R01MH079943 (W.K.S.); UK MRC Clinical Fellowship (G0802821: R.M.-S.); UK National Institute for Health Research (NIHR) grants (PB-PG-0110-21190: C.C.) and (PB-PG-0107-12042: P.C.); UK MRC grants (09-800-17: P.C. and C.C.; G0802326: K.T., P.C., and C.C.; G1002011: P.W., C.C., and P.C.; and G0601874; C.C.). Grant 09-800-17 was managed by NIHR on behalf of the MRC-NIHR partnership. Combined study supported by UK MRC grant G0901874/1 (T.C.E.); Combined study presents independent research partly funded by the NIHR Biomedical Research Centre at South London and Maudsley, NHS Foundation Trust, and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

**Competing interests.** The authors declare that there were no conflicts of interest with respect to the authorship or the publication of this article.

**Ethical standards.** The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

**Supporting information.** All supporting information is either included in this article or presented in the Online Supplementary Material.

## References

Asselmann, E., Wittchen, H. U., Lieb, R., & Beesdo-Baum, K. (2018). Sociodemographic, clinical, and functional long-term outcomes in adolescents and young adults with mental disorders. *Acta Psychiatrica Scandinavica*, *137*(1), 6–17. doi:10.1111/acps.12792

Association, A. P. (2003). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association.

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *JAMA*, *319*(13), 1317–1318. doi:10.1001/jama.2017.18391

Bennett, K., Manassis, K., Walter, S. D., Cheung, A., Wilansky-Traynor, P., Diaz-Granados, N., … Wood, J. J. (2013). Cognitive behavioral therapy age effects in child and adolescent anxiety: An individual patient data meta-analysis. *Depression and Anxiety*, *30*(9), 829–841. doi:10.1002/da.22099

Bertie, L.-A., & Hudson, J. L. (2021). CBT for childhood anxiety: Reviewing the state of personalised intervention research. *Frontiers in Psychology*, *12*, 722546. doi:10.3389/fpsyg.2021.722546

Cecilione, J. L., McLeod, B. D., Southam-Gerow, M. A., Weisz, J. R., & Chorpita, B. F. (2021). Examining the relation between technical and global competence in two treatments for youth anxiety. *Behavior Therapy*, *52*(6), 1395–1407. doi:10.1016/j.beth.2021.03.009

Cheng, H., & Phillips, M. R. (2014). Secondary analysis of existing data: Opportunities and implementation. *Shanghai Archives of Psychiatry*, *26*(6), 371–375. doi:10.11919/j.issn.1002-0829.214171

Coleman, J. R., Lester, K. J., Keers, R., Roberts, S., Curtis, C., Arendt, K., … Eley, T. C. (2016). Genome-wide association study of response to cognitive behavioural therapy in children with anxiety disorders. *British Journal of Psychiatry*, *209*(3), 236–243. doi:10.1192/bjp.bp.115.168229

Coutanche, M. N., & Hallion, L. S. (2020). Machine learning for clinical psychology and clinical neuroscience. In A. Wright & M. Hallquist (Eds.), *Research methods in clinical psychology* (pp. 467–482). Cambridge, UK: Cambridge University Press.

DeRubeis, R. J. (2019). The history, current status, and possible future of precision mental health. *Behaviour Research and Therapy*, *123*, 103506. doi:10.1016/j.brat.2019.103506

Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine learning approaches for clinical psychology and psychiatry. *Annual Review of Clinical Psychology*, *14*(1), 91–118. doi:10.1146/annurev-clinpsy-032816-045037

Evans, R., Clark, D. M., & Leigh, E. (2021). Are young people with primary social anxiety disorder less likely to recover following generic CBT compared to young people with other primary anxiety disorders? A systematic review and meta-analysis. *Behavioural and Cognitive Psychotherapy*, *49*(3), 352–369. doi:10.1017/S135246582000079

Garrido, S., Millington, C., Cheers, D., Boydell, K., Schubert, E., Meade, T., & Nguyen, Q. V. (2019). What works and what doesn't work? A systematic review of digital mental health interventions for depression and anxiety in young people. *Frontiers in Psychiatry*, *10*, 759. doi:10.3389/fpsyt.2019.00759

Géron, A. (2022). *Hands-on machine learning with scikit-learn, keras, and TensorFlow*. O'Reilly Media, Inc.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586.

Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, *41*(5), 645–655. https://doi.org/10.1111/j.1469-7610.2000.tb02345.x

Gould, K. L., Porter, M., Lyneham, H. J., & Hudson, J. L. (2018). Cognitive behavioral therapy for children with anxiety and comorbid attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(7), 481–490. e482. doi:10.1016/j.jaac.2018.03.021

Grist, R., Croker, A., Denne, M., & Stallard, P. (2019). Technology delivered interventions for depression and anxiety in children and adolescents: A systematic review and meta-analysis. *Clinical Child and Family Psychology Review*, *22*, 147–171. doi:10.1007/s10567-018-0271-8

Guo, T., Su, J., Hu, J., Aalberg, M., Zhu, Y., Teng, T., & Zhou, X. (2021). Individual vs. group cognitive behavior therapy for anxiety disorder in children and adolescents: A meta-analysis of randomized controlled trials. *Frontiers in Psychiatry*, *12*, 674267. doi:10.3389/fpsyt.2021.674267

Halldorsdottir, T., & Ollendick, T. H. (2016). Long-term outcomes of brief, intensive CBT for specific phobias: The negative impact of ADHD symptoms. *Journal of Consulting and Clinical Psychology*, *84*(5), 465. doi:10.1037/ccp0000088

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., … Hussain, A. (2024). Interpreting black-box models: A review on explainable artificial intelligence. *Cognitive Computation*, *16*(1), 45–74.

Hudson, J. L., Keers, R., Roberts, S., Coleman, J. R. I., Breen, G., Arendt, K., … Eley, T. C. (2015). Clinical predictors of response to cognitive-behavioral therapy in pediatric anxiety disorders: The genes for treatment (GxT) study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*(6), 454–463. doi:10.1016/j.jaac.2015.03.018

James, A. C., Reardon, T., Soler, A., James, G., & Creswell, C. (2020). Cognitive behavioural therapy for anxiety disorders in children and adolescents. *Cochrane Database of Systematic Reviews* (11), 1465–1858. doi:10.1002/14651858.CD013162.pub2

Joseph, M. (2021). Pytorch tabular: A framework for deep learning with tabular data. arXiv preprint arXiv:2104.13638. doi:10.48550/arXiv.2104.13638

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., … Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, *30*. doi:10.1001/archpsyc.59.10.877

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*, *59*(10), 877–883. doi:10.1001/archpsyc.59.10.877

Kunas, S. L., Lautenbacher, L. M., Lueken, U., & Hilbert, K. (2021). Psychological predictors of cognitive-behavioral therapy outcomes for anxiety and depressive disorders in children and adolescents: A systematic review and meta-analysis. *Journal of Affective Disorders*, *278*, 614–626. doi:10.1016/j.jad.2020.09.092

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. doi:10.1038/nature14539

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the beck depression and anxiety inventories. *Behaviour Research and Therapy*, *33*(3), 335–343. doi:10.1016/0005-7967(94)00075-u

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Paper presented at the Advances in neural information processing systems, California.

Manassis, K., Changgun Le, T., Bennett, K., Zhao, X. Y., Mendlowitz, S., Duda, S., … Wood, J. J. (2014). Types of parental involvement in CBT with anxious youth: A preliminary meta-analysis. *Journal of Consulting and Clinical Psychology*, *82*(6), 1163–1172. doi:10.1037/a0036969

McKinnon, A., Keers, R., Coleman, J. R. I., Lester, K. J., Roberts, S., Arendt, K., … Hudson, J. L. (2018). The impact of treatment delivery format on response to cognitive behaviour therapy for preadolescent children with anxiety disorders. *Journal of Child Psychology and Psychiatry*, *59*(7), 763–772. doi:10.1111/jcpp.12872

Messer, S. C., Angold, A., Costello, E. J., & Loeber, R. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, *5*(4), 251–262.

Ozomaro, U., Wahlestedt, C., & Nemeroff, C. B. (2013). Personalized medicine in psychiatry: Problems and promises. *BMC Medicine*, *11*(1), 132. doi:10.1186/1741-7015-11-132

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Dubourg, V. (2011). Scikit-learn: Machine learning in python. *the Journal of Machine Learning Research*, *12*, 2825–2830.

Perlis, R. (2011). Translating biomarkers to clinical practice. *Molecular Psychiatry*, *16*(11), 1076–1087. doi:10.1038/mp.2011.63

Pintelas, E. G., Kotsilieris, T., Livieris, I. E., & Pintelas, P. (2018). A review of machine learning prediction methods for anxiety disorders. Paper presented at the Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion.

Podell, J. L., Kendall, P. C., Gosch, E. A., Compton, S. N., March, J. S., Albano, A. M., … Piacentini, J. C. (2013). Therapist factors and outcomes in CBT for anxiety in youth. *Professional Psychology: Research and Practice*, *44*(2), 89–98. doi:10.1037/a0031700

Popov, S., Morozov, S., & Babenko, A. (2019). Neural oblivious decision ensembles for deep learning on tabular data. arXiv preprint arXiv:1909.06312. doi:10.48550/arXiv.1909.06312

Rapee, R. M., Lyneham, H. J., Hudson, J. L., Kangas, M., Wuthrich, V. M., & Schniering, C. A. (2013). Effect of comorbidity on treatment of anxious children and adolescents: Results from a large, combined sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, *52*(1), 47–56. doi:10.1016/j.jaac.2012.10.002

Reynolds, S., Wilson, C., Austin, J., & Hooper, L. (2012). Effects of psychotherapy for anxiety in children and adolescents: A meta-analytic review. *Clinical Psychology Review*, *32*(4), 251–262. doi:10.1016/j.cpr.2012.01.005

Schneider, S., Unnewehr, S., & Margraf, J. (2009). Kinder-DIPS. Diagnostisches Interview bei psychischen Störungen im Kindes-und Jugendalter. Retrieved from https://www.scholar.google.com

Scott, I. A. (2021). Demystifying machine learning: A primer for physicians. *Internal Medicine Journal*, *51*(9), 1388–1400. doi:10.1111/imj.15200

Senior, M., Fanshawe, T., Fazel, M., & Fazel, S. (2021). Prediction models for child and adolescent mental health: A systematic review of methodology and reporting in recent research. *JCPP Advances*, *1*(3), e12034. doi:10.1002/jcv2.12034

Sicouri, G., Perkes, I., & Hudson, J. L. (2022). Anxiety disorders in children and adolescents. *Medicine Today*, *23*(1-2), 14–21.

Silverman, W., & Albano, A. (1996). *Anxiety disorders interview schedule (ADIS-IV) child and parent interview schedules: 5 of each*. New York, NY: Graywind Publications.

Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behaviour Research and Therapy*, *36*(5), 545–566.

Spoelma, M. J., Serafimovska, A., & Parker, G. (2023). Differentiating melancholic and non-melancholic depression via biological markers: A review. *The World Journal of Biological Psychiatry*, *24*(9), 761–810. doi:10.1080/15622975.2023.2219725

Vollmer, S., Mateen, B. A., Bohner, G., Király, F. J., Ghani, R., Jonsson, P., … Myles, P. (2020). Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. *BMJ*, *368*, l6927.

Wehry, A. M., Beesdo-Baum, K., Hennelly, M. M., Connolly, S. D., & Strawn, J. R. (2015). Assessment and treatment of anxiety disorders in children and adolescents. *Current Psychiatry Reports*, *17*(7), 52. doi:10.1007/s11920-015-0591-z

Werner-Seidler, A., Spanos, S., Calear, A. L., Perry, Y., Torok, M., O'Dea, B., … Newby, J. M. (2021). School-based depression and anxiety prevention programs: An updated systematic review and meta-analysis. *Clinical Psychology Review*, *89*, 102079. doi:10.1016/j.cpr.2021.102079

Wolff, R. F., Moons, K. G., Riley, R. D., Whiting, P. F., Westwood, M., & Collins, G. S., … Group†, P. (2019). PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine*, *170*(1), 51–58. doi:10.7326/M18-1376

Zhang, Q., Wang, J., & Neitzel, A. (2023). School-based mental health interventions targeting depression or anxiety: A meta-analysis of rigorous randomized controlled trials for school-aged children and adolescents. *Journal of Youth and Adolescence*, *52*(1), 195–217. doi:10.1007/s10964-022-01684-4