

RESEARCH

Open Access



# Common and differential variables of anxiety and depression in adolescence: a nation-wide smartphone-based survey

Martin Weiß<sup>1,7\*</sup>, Julian Gutzeit<sup>1,2†</sup>, Rüdiger Pryss<sup>3,4</sup>, Marcel Romanos<sup>5</sup>, Lorenz Deserno<sup>5,6</sup> and Grit Hein<sup>1</sup>

## Abstract

**Background** Mental health in adolescence is critical in its own right and a predictor of later symptoms of anxiety and depression. To address these mental health challenges, it is crucial to understand the variables linked to anxiety and depression in adolescence.

**Methods** Here, we analyzed data of 278 adolescents that were collected in a nation-wide survey provided via a smartphone-based application during the COVID-19 pandemic. We used an elastic net regression machine-learning approach to classify individuals with clinically relevant self-reported symptoms of depression or anxiety. We then identified the most important variables with a combination of permutation feature importance calculation and sequential logistic regressions.

**Results** 40.30% of participants reported clinically relevant anxiety symptoms, and 37.69% reported depressive symptoms. Both machine-learning models performed well in classifying participants with depressive (AUROC = 0.77) or anxiety (AUROC = 0.83) symptoms and were significantly better than the no-information rate. Feature importance analyses revealed that anxiety and depression in adolescence are commonly related to sleep disturbances (anxiety OR = 2.12, depression OR = 1.80). Differentiating between symptoms, self-reported depression increased with decreasing life satisfaction (OR = 0.43), whereas self-reported anxiety was related to worries about the health of family and friends (OR = 1.98) as well as impulsivity (OR = 2.01).

**Conclusion** Our results show that app-based self-reports provide information that can classify symptoms of anxiety and depression in adolescence and thus offer new insights into symptom patterns related to adolescent mental health issues. These findings underscore the potentials of health apps in reaching large cohorts of adolescence and optimize diagnostic and treatment.

**Keywords** Adolescence, Anxiety, Depression, Machine learning, e-Health

<sup>†</sup>Martin Weiß and Julian Gutzeit have contributed equally to this work.

\*Correspondence:

Martin Weiß  
Weiss\_M11@ukw.de

<sup>1</sup>Department of Psychiatry, Psychosomatic and Psychotherapy, Center of Mental Health, University Hospital Würzburg, Margarete-Höppel-Platz 1, 97080 Würzburg, Germany

<sup>2</sup>Department of Psychology III, University of Würzburg, Würzburg, Germany

<sup>3</sup>Institute of Clinical Epidemiology and Biometry, University of Würzburg, Würzburg, Germany

<sup>4</sup>Institute of Medical Data Science, University Hospital Würzburg, Würzburg, Germany

<sup>5</sup>Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Center of Mental Health, University Hospital Würzburg, Würzburg, Germany

<sup>6</sup>Department of Psychiatry and Psychotherapy, Technische Universität Dresden, Dresden, Germany

<sup>7</sup>Department of Psychology I, University of Würzburg, Würzburg, Germany



## Introduction

Anxiety disorders and depressive disorders are among the most prevalent mental health disorders in adolescents worldwide. Pre-pandemic estimates for anxiety and depression indicate that about 11.6% [57] and 12.9% [35] respectively, of adolescents were affected. According to a recent meta-analysis including data from 329,159 children and adolescents  $\leq 18$  years between January 2020 and December 2023, however, 26% reported symptoms of anxiety and 23% reported symptoms of depression [2]. For distinct countries, the numbers were even higher, as the prevalence during the COVID-19 outbreak for anxiety and depression in Chinese adolescents were 37.4% and 43.7% respectively [61].

In line with these findings, longitudinal studies including data from pre- and during the COVID-19 pandemic report increases in symptoms of depression and anxiety in 13–16 year-old adolescents (e.g., Magson et al. [36, 48]). However, a meta-analysis of 12 longitudinal studies suggest that only depressive symptoms increased significantly, while anxiety symptoms remained stable [7].

Different research reports heterogeneous patterns of predictors that contribute to clinical levels of anxiety and depression in adolescents. For instance, Stewart and colleagues [54] found that in a Scottish sample, self-reported symptoms of anxiety and depression were more likely among older individuals, females, those with current or past mental health support, those needing extra school support, and those experiencing poorer home relationships since the COVID-19 pandemic. In a Jordanian sample, AlAzzam et al. [1] showed that the educational level of both parents, perceived difficulties in online education, sex, and age were related to depression, while only the father's level of education, difficulties in online education, sex, and age were related to anxiety disorders.

As both disorders can lead to substantial impairments in daily functioning [19], earlier identification and intervention are critical to prevent common chronic and comorbid disease patterns. Further, child and adolescent mental health is one critical predictor of adult mental health (e.g., Otto et al. [41] for review, see Johnson et al. [27]). Therefore, the aim of this study is to identify the predictors of anxiety and depression in a sample of adolescents from a nation-wide survey provided via a smartphone-based application (app) in Germany.

Adolescents participated in the current study without targeted recruitment, as the app was publicly available in the app stores. This approach overcomes several limitations typically associated with (adolescent) mental health research and services (e.g., Bantjes [6]). These limitations include scheduling issues (i.e., adolescents rely on their parents' scheduling and mobility to participate in research), adolescents' unwillingness to share personal information with professionals, and their need for

autonomy and independence [43]. Serious apps (should) guarantee anonymity, thus helping to overcome stigmatization, and allow users to decide for themselves when and to what extent they provide data about their mental well-being [24]. The use of mental health apps in research thus helps to increase the heterogeneity of samples (e.g., in resource-limited settings, Lehtimäki et al. [32] and by removing geographic barriers, Bührmann et al. [12] and ideally to reach more of those who would need help on the basis of self-diagnosis.

Leveraging these benefits of data retrieved via mental health apps, we used elastic net regressions to classify individuals with levels above the cut-off for anxiety and depression. By variable selection and regularization, this machine-learning method identifies the most important variables associated with anxiety and depression, respectively, and provides robust classifications.

## Methods

### Participants

We included 278 adolescents that provided data via the Corona Health App (129 female, mean age=15.24, SD=1.57, range=12–17). The Corona Health App was initiated by the Mental Health Research Unit of the Robert Koch Institute (RKI), i.e., the German federal agency for public health responsible for disease control and prevention, and the Universities of Würzburg, Ulm and Regensburg (for details, see Beierle et al. [9]). The goal of Corona Health was to monitor mental and physical well-being during the COVID-19 pandemic. Users could download the Corona Health App for free from the Apple Store and from the Google Play store. Thus, there were no specific criteria for recruitment. All participants gave their consent within the app, and without consent the app was closed. After giving consent, participants could select the study they wanted to participate in (in this case, the “Mental Health for Adolescents” module) and were forwarded to a baseline questionnaire. Participants were then able to allow mobile sensing features and schedule follow-up surveys (which is not part of the present analysis). The median duration to fill out the baseline survey was 8:52 min. The present analyses were based on the cross-sectional data collected between July 2020 and September 2022.

### Measurements

The two questionnaires used to assess the dependent variables were the Patient Health Questionnaire-2 for symptoms of depression (PHQ-2, two items, cutoff  $\geq 3$ ; Löwe et al. [34] and the Brief Spence Children's Anxiety Scale for symptoms of anxiety (SCAS-C-8, eight items based on the DSM-5, Cutoff  $> 6.5$ ; Reardon et al. [46]). The PHQ-2 with a cutoff of  $\geq 3$  has been shown to be a valid screening of clinically relevant depressive symptoms for

13–17 year olds [47]. The SCAS-C-8 was mainly developed for younger samples, but has also been used for adolescents up to the age of 18 (e.g., Orgilés et al. [40]). The reliabilities of both questionnaires in our sample were Cronbach's  $\alpha=0.79$  and  $\alpha=0.84$ , for depression and anxiety, respectively. Apart from quality of life, all other items uniquely assessed an aspect in the dataset (e.g., worries about COVID-19, media usage, going outside, impulsivity, arguments at home/school, prior anxiety and depression disorder, etc.). All items were selected or generated by a multi-professional team of different domain experts (app developers, medical professionals, psychologists, etc.) in a multi-stage process of discussion and several feedback loops [9].

### Analysis

We used elastic net logistic regression, a regularized regression method, to classify adolescents' above vs. below the cutoffs for clinically relevant symptoms of depression and anxiety. With the elastic net, the number of selected features can be even larger than the sample size, achieving a sparse model [13, 63]. We chose to classify cases and controls (i.e., individuals above and below the cutoff of clinically relevant levels of anxiety and depression) as this might be more helpful for practitioners in clinical practice. For the remaining 60 items, we applied an iterative imputation method based on a random forest using the "missForest" package in *R* (R Core Team 2021) to obtain a data set without missing values [53]. In this way, 25 missing values were imputed in the training data set and 9 missing values in the test data set. These items were then included in our analyses (z-standardized metric or dummy-coded in case of factors). This means that in the depression model, the 60 items plus the anxiety cut-off item were used and in the anxiety model, the 60 items plus the depression cut-off item were used. A list of all included items can be found in the Supplementary Materials. We randomly selected 70% of the data as the training dataset and applied 5-fold repeated cross-validation to train and tune our model over a grid of  $\alpha$  and  $\lambda$  hyperparameters, using the *R* package "caret" [29]. First, we refitted the model on the training dataset with the best performing hyperparameters to calculate the final penalized  $\beta$  coefficients. Second, we applied the model to the remaining 30% of the sample, i.e., the test dataset, to estimate model performance (accuracy, area under the receiver operating characteristic curve [AUROC], sensitivity, specificity, positive and negative prediction value, balanced accuracy, and Cohen's kappa). With this split we aimed for a large enough train data set for adequately tuning hyperparameters while still holding out a large enough test set to achieve a robust model performance evaluation accounting for the imbalanced dataset (see Table 1). We conducted separate variable

**Table 1** Prevalence of clinically significant symptoms of depression (PHQ2  $\geq 3$ ) and anxiety (SCAS-8  $> 6.5$ ) and gender

	Male	Female	Total
No symptoms of depression	95 (68.35%)	72 (55.81%)	167 (62.31%)
Symptoms of depression	44 (31.65%)	57 (44.19%)	101 (37.69%)
No symptoms of anxiety	96 (69.04%)	64 (49.61%)	160 (59.70%)
Symptoms of Anxiety	43 (30.94%)	65 (50.39%)	108 (40.30%)

Absolute numbers and relative numbers for each gender group in parentheses

importance analyses for depression and anxiety models using permutation importance on the test data set. This technique involves repeatedly permuting (100 permutations in our case) features and evaluating model performance with each variable in a non-informative context. The importance of each feature is quantified by calculating the decrease in model performance (AUROC) resulting from the lack of information provided by the permuted feature [3]. To identify the most important variables we computed multiple logistic regressions sequentially including all predictors in order of the calculated importance determined by the elastic net regression starting with only one (the most important) predictor and ending with all predictors (including the least important predictor last). We then selected the model with the lowest Bayesian information criterion (BIC), a measure indicating goodness of fit while simultaneously penalizing complexity of the model and thus addressing overfitting. We then tested the best performing models with the reduced feature space on the test data set to investigate the generalizability and robustness of our findings.

## Results

### Sample

Classification of participants with clinically relevant depressive or anxiety symptoms by gender are described in Table 1. The comorbidity of depressive and anxiety symptoms in our study was 22.8%. The reliabilities (Cronbach's  $\alpha$ ) for PHQ2 and SCAS-8 are 0.79 and 0.83, respectively. The training and test data set did not differ significantly regarding demographic characteristics (see Supplementary material, Table S2).

### Depression

In summary, the classification model for symptoms of depression had an overall accuracy of 71.60% with a 95% confidence interval of [60.50%, 81.07%] and an AUROC of 0.77. The accuracy was significantly better than the no-information rate (NIR=61.73%,  $p=0.041$ ). It performed well in classifying healthy participants (specificity=90.00%) but showed a low sensitivity of 41.94%. Despite that, it was able to classify individuals with clinically relevant symptoms of depression (positive prediction value=72.22%) and healthy participants (negative prediction value=71.43%). Overall, the model had a

kappa value of 0.347 (considered “fair” according to Landis and Koch [31] and a balanced accuracy of 65.97%.

The top two predictors identified by the BIC for the classification of depressive symptoms were *sleep disturbances* and *life satisfaction* (Fig. 1, upper panel).

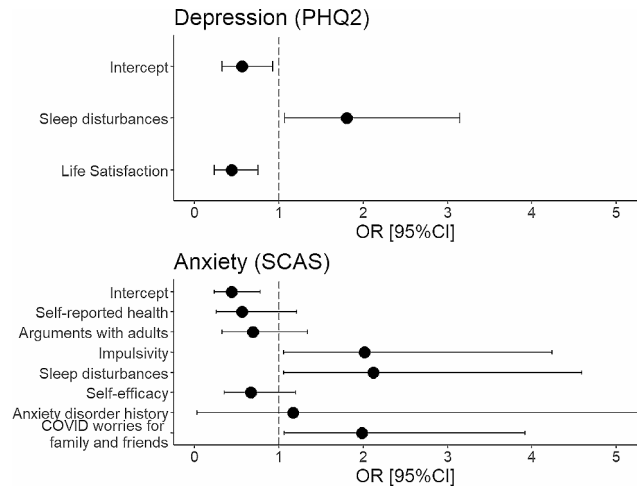
The logistic regression with these two predictors as independent variables and depression scores as dependent variable on the test data set was significant. ( $\chi^2 = 19.68, p < 0.001$ , Nagelkerke  $R^2 = 0.293$ ) and both predictors had a significant effect on the odds of being depressed (see Table 2).

### Anxiety

The classification model for symptoms of anxiety showed an overall accuracy of 76.54% with a 95% confidence interval of [65.82%, 82.52%] and an AUROC of 0.83. The difference in information rate compared to the no-information rate was significant (NIR: 64.20%,  $p = 0.012$ ). The model had a sensitivity of 65.52% and a specificity of 82.69%. It successfully classified between healthy individuals and individuals with anxiety (negative prediction value = 81.13%; positive prediction value = 67.86% indicating a higher false positive rate than false negative rate). Overall, the model had a moderate performance, as indicated by its kappa value of 0.486 and its balanced accuracy of 74.10%.

The top seven predictors selected with BIC for the classification of anxiety were *self-reported health*, *arguments with adults*, *impulsivity*, *sleep disturbances*, *self-efficacy*, *anxiety disorder history* and *COVID worries for family and friends* (Fig. 1, lower panel).

The logistic regression with these seven predictors as independent variables and anxiety as dependent variable on the test data set was significant ( $\chi^2 = 30.42, p < 0.001$ , Nagelkerke  $R^2 = 0.430$ ). However, only predictors *impulsivity*, *sleep disturbances*, and *COVID worries for family and friends* had a significant effect on the odds of suffering from anxiety. All coefficients can be seen in Table 3.



**Fig. 1** Odds ratios and 95% CIs of logistic regression models with the most important predictors determined by BIC

**Table 2** Coefficients and odds-ratios of the logistic regression on depression with top two predictors of the elastic net regression

Predictor	Estimate	SE	z	p	OR [95%CI]
<b>(Intercept)</b>	<b>-0.58</b>	<b>0.26</b>	<b>-2.21</b>	<b>0.027</b>	<b>0.55 [0.33; 0.93]</b>
<b>Sleep disturbances</b>	<b>0.59</b>	<b>0.27</b>	<b>2.16</b>	<b>0.031</b>	<b>1.80 [1.07; 3.14]</b>
<b>Life satisfaction</b>	<b>-0.83</b>	<b>0.29</b>	<b>-2.83</b>	<b>0.005</b>	<b>0.43 [0.24; 0.75]</b>

Significant predictors are shown in bold

### Discussion

In this study, we analyzed self-reported data of adolescents collected in their everyday lives via the RKI Corona Health App between July 2020 and September 2022 to find associations between relevant variables and symptoms of depression and anxiety. We computed two elastic net regression models to classify healthy individuals and individuals with symptoms of depression or symptoms of anxiety, respectively. Both models achieved significantly higher accuracies in classifying cases than the no-information rates and classified most cases correctly. We conducted a variable importance analysis, identifying significant predictors for depression and anxiety symptoms with common and distinct profiles.

A clear advantage of this analysis is the assessment of mental health via smartphone app. this gives many

**Table 3** Coefficients and odds-ratios of the logistic regression on anxiety with top seven predictors of the elastic net regression

Predictor	Estimate	SE	z	p	OR [95%CI]
<b>(Intercept)</b>	<b>-0.83</b>	<b>0.30</b>	<b>-2.71</b>	<b>0.007</b>	<b>0.44 [0.23; 0.78]</b>
Self-reported health	-0.57	0.39	-1.47	0.143	0.56 [0.26; 1.21]
Arguments with adults	-0.37	0.36	-1.03	0.303	0.69 [0.33; 1.34]
<b>Impulsivity</b>	<b>0.70</b>	<b>0.35</b>	<b>2.01</b>	<b>0.045</b>	<b>2.01 [1.06; 4.24]</b>
<b>Sleep disturbances</b>	<b>0.75</b>	<b>0.37</b>	<b>2.03</b>	<b>0.042</b>	<b>2.12 [1.06; 4.59]</b>
Self-efficacy	-0.41	0.31	-1.33	0.184	0.67 [0.35; 1.20]
Anxiety disorder history	0.15	1.67	0.09	0.927	1.16 [0.03; 42.25]
<b>COVID worries for family and friends</b>	<b>0.68</b>	<b>0.33</b>	<b>2.08</b>	<b>0.037</b>	<b>1.98 [1.06; 3.91]</b>

Significant predictors are shown in bold

people the opportunity to take part in such a survey who might otherwise have experienced certain geographical, scheduling or motivational barriers to participating in a conventional survey. As a matter of fact, we found comparably high prevalence of symptoms of anxiety and depression in our sample. 37.69% of all participants were classified with clinically significant symptoms of depression and 40.30% with clinically significant symptoms of anxiety. These prevalences are noticeably higher than reported in recent literature (e.g., 23.3% for anxiety disorders in German adolescents, Niermann et al. [39] 11.5% for depression in German adolescents Scheiner et al. [50]). This could imply that more affected people were reached by the survey via the app. However, direct comparisons between our prevalence rates and these reported rates must be interpreted cautiously. In our study, anxiety and depressive symptoms were assessed using self-report screenings in a self-selected sample. In both reported studies, the samples were larger and randomly drawn from the population. Mental disorders were assessed either with multiple, longer scales than ours [50] or with a standardized interview [39]. Thus, due to the psychometric properties of our measurements, we might have over-detected clinically relevant symptoms in our sample compared to those in these studies (see [Limitations](#)).

In both symptom constellations sleep disturbances had the largest odds ratio, i.e., the largest effect size, in the reduced models. This could indicate various things. First, there could be a non-causal relation between sleep-disturbances, symptoms of anxiety, and symptoms of depression in our sample explained by a fourth variable. For instance, family disorganization (i.e., lack of structure and routine within the home) has been associated with sleep disturbance, anxiety disorders [58], and an increased risk for depression [22]. During the pandemic, families were confronted with increased strain and commotion due to several COVID-19 related stressors, such as physical and mental health concerns, economic stress, intensified relationships, and conflicts [59]. However, symptoms of anxiety, symptoms of depression, and sleep disturbances could also be more directly related to each other. The finding that sleep disturbances in adolescents are associated with both symptoms of depression and anxiety is well documented [21]. Sleep disturbances are a common symptom of both depression and generalized anxiety disorder according to the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; DSM-5; APA, [4]). There is a high comorbidity of depression and anxiety disorders, ranging up to 50% in community samples [19] and 22.8% symptom comorbidity in our study. Since the present study only examines cross-sectional data, it is not possible to say clearly whether sleep disorders are the cause or consequence of symptoms of anxiety or depression, or whether they arise primarily in the comorbidity

of these symptom clusters. However, a large body of research has shown that preceding sleep disturbances are an important risk factor for both disorders (e.g., [10, 14, 17, 21]). Potential mechanisms could be a decrease of anterior white matter tracts and connectivity in the fronto-limbic network caused by sleep disturbance [26]. These regions are associated with the control and regulation of negative emotions. Impaired processing in these regions might lead to psychological processes associated both with anxiety and affective disorders [10]. In conclusion, while it is difficult to safely determine the direction of the triadic association link between sleep disturbance, anxiety symptoms, and depressive symptoms from this cross-sectional data, we conclude that sleep disturbances have a high importance for detecting or diagnosing symptoms of anxiety and depression.

Apart from the predictor overlap of sleep disturbances, the models differ noticeably regarding their most important variables. In the reduced depression model, only life satisfaction was a significant predictor besides sleep disturbances, showing an inverse relation with depressive symptoms. There is evidence that higher life satisfaction is an indicator of an adaptive and positive psychosocial functioning which is a protective factor against emotional deficits, such as symptoms of depression [20, 44]. Although we do not want to draw any causal conclusions on the relationship between life satisfaction and depression, the directionality (i.e., life satisfaction as regressor for depression as criterion) and the inverse relationship is in line with previous research (e.g., Tang et al. [55]). Except for sleep disturbances and (reduced) life satisfaction, there were no more significant predictors for depressive symptoms in the reduced model.

Interestingly, impulsivity was the predictor with the second largest effect for self-reported anxiety symptoms. Past research has reported an inverse relation between anxiety disorders and impulsivity [5, 8]. However, more recent research has also shown positive associations between impulsivity and anxiety disorders among adolescents (e.g., Moustafa et al., [37]). This relationship could arise indirectly through the presence of another mental disorder. For instance, almost every second child diagnosed with ADHD—a mental disorder that is associated with increased impulsivity—suffers from a comorbid mood disorder [62]. This assumption is supported by evidence using cross-sectional and longitudinal psychometric networks showing that ADHD is reciprocally associated with internalizing symptoms via potential bridging symptoms primarily related to anxiety symptoms [52]. There is also evidence that Cyclothymia might also be a common factor of increased impulsivity and anxiety [42]. Last, while anxiety in mood disorders might increase impulsive behaviors related to suicide [16, 51], we doubt that this is the primary driver of the observed

relationship in our dataset. However, it is conceivable that heightened anxiety may specifically increase impulsivity in areas of daily life associated with anxiety.

The third significant aspect of symptoms of anxiety is COVID worries for family and friends. In the different phases of the pandemic, adolescents experienced health-related worry regarding their own health as well as the health of their families and peers—but also worries related to their social lives (for review, see Guessoum et al. [23]). It has been shown that such worries had a direct effect on general anxiety [38]. Independent of the pandemic context, worry was found to be a stable construct that has been linked to anxiety disorders throughout adolescence [45]. Worrying is also bidirectionally related to sleep disturbances [56], another important predictor for anxiety in our model (see above).

More than one in five participants showed comorbid depressive and anxiety symptoms. We did not train a third model to identify this subgroup. However, we would expect sleep disturbances to be one of the most important predictors, as it had high predictive value in both separate models. We would also expect life satisfaction to be strongly decreased for these patients, as it was strongly associated with depressive symptoms and has been shown to be the case for patients suffering from comorbid anxiety and depression [15].

In conclusion, the models classifying depressive and anxiety symptoms show a clear overlap regarding sleep disturbances but differ noticeably in other predictors, namely life satisfaction, impulsivity, and COVID-worries. These findings make an important contribution to the differential diagnosis of both disorders. Our analyses cannot clearly determine whether the significant predictors are risk factors, symptoms, or other relevant variables of the respective syndromes. However, due to the predictor overlap of sleep disorders of both symptom clusters, one could conclude that sleep quality is of great importance for the general prevention of mental illness in adolescents [60].

### Limitations

This study focused solely on the baseline measurement of our app-based survey, analyzing cross-sectional data. While longitudinal data from follow-up surveys were collected, we excluded these from the current analysis. Our decision to concentrate on cross-sectional data was driven by the need for a sufficient sample size to conduct robust machine learning analyses. The lower follow-up rate of participants made the longitudinal data less suitable for this purpose. However, future research could explore the temporal development of depressive and anxiety symptoms using the follow-up data, potentially identifying preceding risk and resilience factors.

Even though the classification model for depressive symptoms performed significantly better than the no information rate, it had a rather low sensitivity. This indicates difficulties in reliably detecting participants with clinically relevant depressive symptoms above the cut-off. One potential reason for this might be the psychometric characteristics of the used outcome. There is, for instance, evidence that the PHQ-2 tends to have low specificity in detecting clinically relevant depressive symptoms [33]. This might have led to an over-detection of individuals with “truly” clinically relevant depressive symptoms, as is also evident in our rather high prevalence rate. Thus, the model might have had difficulties in identifying robust associations between several relevant variables, symptom constellations, and depressive symptoms, as the prevalence of the latter might have been overestimated (though this could also be attributed to self-selection bias, discussed below). The rationale for choosing this short screening tool was to find a balance between survey brevity and accuracy. Shorter surveys tend to have better response rates [49], and the aim of the original study was to assess a wide range of different constructs [9]. For future research, however, a combination of the PHQ-2 with the PHQ-9 might yield better accuracy in identifying the criterion of depressive symptoms [33]. Another alternative would be to address the imbalance between depressive symptoms (two items) and symptoms of anxiety (eight items) by assessing general distress using the Kessler Psychological Distress Scale (K10; Kessler et al. [28]). The K10 includes 10 questions about emotional states, serves as a brief screening tool to identify levels of distress and has been used effectively in adolescents (e.g., Boyes et al. [11]). Generally, many of the variables were assessed by only one item. While this again had clear benefits regarding survey length, it limits generalizability and the ability to capture nuances in the constructs. Although the outcome measure for the anxiety model had also only moderate accuracy [46], the performance seemed to be less affected by this than the depression model. This might indicate that the training fit was better due to slightly higher prevalence of cases with relevant anxiety symptoms than depression symptoms. Another explanation could be that the relations between some potential risk-factors and symptom patterns of depression were more complex and harder to detect for the elastic net regressions, especially compared to the anxiety model.

In conclusion, both models showed significant improvements in classifying outcomes over respective no-information rates. The performance of both models, which is satisfactory in some areas only, must be considered in the context of the partially imprecise outcome measures and can still be regarded as accurate in general.

Generally, we found a comparably high prevalence rate of mental health problems in our sample. This might have been due to a self-selection bias in our sample: As the Corona Health App was advertised in the German media and participation in the survey was voluntary, one could argue that individuals who were already feeling unwell or were heavily worried by the pandemic were more motivated to participate in the study. Moreover, one might wonder who downloaded the app completely voluntarily, and who was advised or encouraged by their parents. 14.71% of participants reported being in psychotherapy at the time of the survey, a notably higher rate than the German national average of adolescents (1.4% of all German children/adolescents, highest rate of 2.5% among 15- to 19-year-olds; Jaite et al. [25]). Thus, worried parents might have encouraged their children to seek help or partake in scientific studies regarding mental health, leading to a potentially un-representative sample. This potential bias, however, may have actually enhanced our analyses and models. By providing a more balanced sample than the general population, which typically includes far more healthy individuals than those with anxiety or depressive symptoms, we potentially improved accuracy. Highly imbalanced data can skew model classification, as the majority class tends to be overrepresented and over-classified [30]. Some researchers deliberately overrepresent the minority class—in our case, participants with clinically relevant symptoms—by under-sampling the majority or over-sampling the minority [18]. Consequently, we believe the classification analyses yielded robust results, despite the possibility of self-selection bias.

## Conclusion

In conclusion, our data gathered from individuals across Germany based on a self-downloaded App give novel insights into symptom constellations associated with mental health problems in adolescents. The overlapping and distinct symptoms of anxiety and depression can help informing prevention programs aiming at early detection of mental health issues. Employing advanced data-driven methodologies on extensive datasets not only advances scientific knowledge but also paves the way for more effective and personalized strategies to foster resilient mental health among today's youth.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13034-024-00793-1>.

Supplementary Material 1.

## Author contributions

Conceptualization: M.W. and J.G.; Data curation: M.W. and J.G.; Formal analysis: M.W. and J.G.; Funding acquisition: M.R.; Methodology: M.W. and J.G.; Project administration: R.P. and M.R.; Resources: R.P. and M.R.; Software: R.P.;

Supervision: L.D. and G.H.; Validation: M.W. and J.G.; Visualization: M.W. and J.G.; Writing—original draft: M.W., J.G., L.D. and G.H.; Writing—review and editing: M.W., J.G., R.P., M.R., L.D. and G.H.

## Funding

J.G. receives funding from the Nationales Pandemie Kohorten Netz - Therapeutische Interventionsplattform (NAPKON-TIP) network, which is funded by the Federal Ministry of Education and Research. R.P. is supported by grants in the projects COMPASS and NAPKON. The COMPASS and NAPKON projects are part of the German COVID-19 Research Network of University Medicine ("Netzwerk Universitätsmedizin"), funded by the German Federal Ministry of Education and Research (funding reference 01KX2021). L.D. is supported by the German Research Foundation (DFG) as part of the Collaborative Research Centre 265 Losing and Regaining Control over drug intake (402170461, Project A02) and a DFG grant to L.D. on ADHD (533682086).

Open Access funding enabled and organized by Projekt DEAL.

## Availability of data and materials

The data and scripts that support the findings of this study are openly available at <https://osf.io/xfnuv/>.

## Declarations

### Ethics approval and consent to participate

The Corona Health app study was conducted in accordance with the German medical products law. The data protection officer and the ethics committee of the University of Würzburg, Germany, approved the study (No. 130/20-me). The procedures used in this study adhere to the tenets of the Declaration of Helsinki. Participation was voluntary, and no financial compensation was provided. Each participant provided informed consent.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

Received: 18 March 2024 / Accepted: 5 August 2024

Published online: 17 August 2024

## References

1. AlAzzam M, Abuhammad S, Abdalrahim A, Hamdan-Mansour AM. Predictors of depression and anxiety among senior high school students during COVID-19 pandemic: the context of home quarantine and online education. *J School Nurs*. 2021;37(4):241–8. <https://doi.org/10.1177/1059840520988548>.
2. Alizadeh S, Shahrousvand S, Sepandi M, Alimohamadi Y. Prevalence of anxiety, depression and post-traumatic stress disorder symptoms in children and adolescents during the COVID-19 pandemic: a systematic review and meta-analysis. *J Public Health*. 2023. <https://doi.org/10.1007/s10389-023-02168-w>.
3. Altmann A, Toloşi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. *Bioinformatics*. 2010;26(10):1340–7. <https://doi.org/10.1093/bioinformatics/btq134>.
4. APA. Diagnostic and statistical manual of mental disorders. The American Psychiatric Association; 2013.
5. Askenazy F, Caci H, Myquel M, Darcourt G, Lecrubier Y. Relationship between impulsivity and platelet serotonin content in adolescents. *Psychiatry Res*. 2000;94(1):19–28. [https://doi.org/10.1016/S0165-1781\(00\)00124-4](https://doi.org/10.1016/S0165-1781(00)00124-4).
6. Bantjes J. Digital solutions to promote adolescent mental health: opportunities and challenges for research and practice. *PLoS Med*. 2022;19(5):e1004008. <https://doi.org/10.1371/journal.pmed.1004008>.
7. Barendse MEA, Flannery J, Cavanagh C, Aristizabal M, Becker SP, Berger E, Breaux R, Campione-Barr N, Church JA, Crone EA, Dahl RE, Dennis-Tiwary TA, Dvorsky MR, Dziura SL, van de Groep S, Ho TC, Killoren SE, Langberg JM, Larginho TL, Pfeifer JH. Longitudinal change in adolescent depression and anxiety symptoms from before to during the COVID-19 pandemic. *J Res Adolesc*. 2023;33(1):74–91. <https://doi.org/10.1111/jora.12781>.

8. Barratt ES. Factor analysis of some psychometric measures of impulsiveness and anxiety. *Psychol Rep.* 1965;16(2):547–54. <https://doi.org/10.2466/pr0.1965.16.2.547>.
9. Beierle F, Schobel J, Vogel C, Allgaier J, Mulansky L, Haug F, Haug J, Schlee W, Holfelder M, Stach M, Schickler M, Baumeister H, Cohrdes C, Deckert J, Deserno L, Edler J-S, Eichner FA, Greger H, Hein G, Pryss R. Corona health—a study- and sensor-based mobile app platform exploring aspects of the COVID-19 pandemic. *Int J Environ Res Public Health.* 2021;18(14):14. <https://doi.org/10.3390/ijerph18147395>.
10. Blake MJ, Trinder JA, Allen NB. Mechanisms underlying the association between insomnia, anxiety, and depression in adolescence: implications for behavioral sleep interventions. *Clin Psychol Rev.* 2018;63:25–40. <https://doi.org/10.1016/j.cpr.2018.05.006>.
11. Boyes A, Levenstein JM, McLoughlin LT, Driver C, Mills L, Lagopoulos J, Hermens DF. A short-interval longitudinal study of associations between psychological distress and hippocampal grey matter in early adolescence. *Brain Imaging Behav.* 2024. <https://doi.org/10.1007/s11682-023-00847-6>.
12. Bührmann L, Van Daele T, Rinn A, De Witte NAJ, Lehr D, Aardoom JJ, Loheide-Niesmann L, Smit J, Ripper H. The feasibility of using Apple's ResearchKit for recruitment and data collection: considerations for mental health research. *Front Digit Health.* 2022. <https://doi.org/10.3389/fdgth.2022.978749>.
13. Carroll MK, Cecchi GA, Rish I, Garg R, Rao AR. Prediction and interpretation of distributed neural activity with sparse models. *NeuroImage.* 2009;44(1):112–22. <https://doi.org/10.1016/j.neuroimage.2008.08.020>.
14. Crowe K, Spiro-Levitt C. Sleep-related problems and pediatric anxiety disorders. *Child Adolesc Psychiatr Clin N Am.* 2021;30(1):209–24. <https://doi.org/10.1016/j.chc.2020.09.004>.
15. Cummings, C. M., Caporino, N. E., & Kendall, P. C. (2014). Comorbidity of anxiety and depression in children and adolescents: 20 years after. *Psychological bulletin, 140*(3), 816.
16. Fava M, Alpert JE, Carmin CN, Wisniewski SR, Trivedi MH, Biggs MM, Shores-Wilson K, Morgan D, Schwartz T, Balasubramani GK, Rush AJ. Clinical correlates and symptom patterns of anxious depression among patients with major depressive disorder in STAR\*D. *Psychol Med.* 2004;34(7):1299–308. <https://doi.org/10.1017/s0033291704002612>.
17. Franzen PL, Buysse DJ. Sleep disturbances and depression: risk relationships for subsequent depression and therapeutic implications. *Dialog Clin Neurosci.* 2008;10(4):473–81. <https://doi.org/10.31887/DCNS.2008.10.4/plfranzen>.
18. Ganganwar V. An overview of classification algorithms for imbalanced datasets. *Int J Emerg Technol Adv Eng.* 2012;4(2):42–7.
19. Garber J, Weersing VR. Comorbidity of anxiety and depression in youth: implications for treatment and prevention. *Clin Psychol Sci Pract.* 2010;17(4):293–306. <https://doi.org/10.1111/j.1468-2850.2010.01221.x>.
20. Gilman R, Huebner ES. Characteristics of adolescents who Report very high life satisfaction. *J Youth Adolesc.* 2006;35(3):293–301. <https://doi.org/10.1007/s10964-006-9036-7>.
21. Gradaris M, Kahn M, Micic G, Short M, Reynolds C, Orchard F, Bauducco S, Bartel K, Richardson C. Sleep's role in the development and resolution of adolescent depression. *Nat Rev Psychol.* 2022;1(9):512. <https://doi.org/10.1038/s44159-022-00074-8>.
22. Guerrero-Muñoz D, Salazar D, Constain V, Perez A, Pineda-Cañar CA, García-Perdomo HA. Association between family functionality and depression: a systematic review and meta-analysis. *Korean J Family Med.* 2021;42(2):172–80. <https://doi.org/10.4082/kjfm.19.0166>.
23. Guessoum SB, Lachal J, Radjack R, Carretier E, Minassian S, Benoit L, Moro MR. Adolescent psychiatric disorders during the COVID-19 pandemic and lockdown. *Psychiatry Res.* 2020;291:113264. <https://doi.org/10.1016/j.psychres.2020.113264>.
24. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. *BMC Psychiatry.* 2010;10:113. <https://doi.org/10.1186/1471-244X-10-113>.
25. Jaite C, Seidel A, Hoffmann F, Matzejat F, Bachmann J. Guideline-based psychotherapy of children and adolescents in Germany: frequency, treatment modalities, and duration of treatment. *Deutsches Ärzteblatt Int.* 2022;119(8):132–3. <https://doi.org/10.3238/arztebl.m2022.0106>.
26. Jamieson D, Shan Z, Lagopoulos J, Hermens DF. The role of adolescent sleep quality in the development of anxiety disorders: a neurobiologically-informed model. *Sleep Med Rev.* 2021;59:101450. <https://doi.org/10.1016/j.smrv.2021.101450>.
27. Johnson D, Dupuis G, Piche J, Clayborne Z, Colman I. Adult mental health outcomes of adolescent depression: a systematic review. *Depress Anxiety.* 2018;35(8):700–16. <https://doi.org/10.1002/da.22777>.
28. Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, Howes MJ, Normand S-LT, Manderscheid RW, Walters EE, Zaslavsky AM. Screening for serious mental illness in the general population. *Arch Gen Psychiatry.* 2003;60(2):184–9. <https://doi.org/10.1001/archpsyc.60.2.184>.
29. Kuhn M. Building predictive models in R using the caret package. *J Stat Softw.* 2008;28:1–26. <https://doi.org/10.18637/jss.v028.i05>.
30. Kumar A, Sheshadri HS. On the classification of Imbalanced datasets. *Int J Comput Appl.* 2012;44(8):1–7. <https://doi.org/10.5120/6280-8449>.
31. Landis JR, Koch GG. The measurement of Observer Agreement for Categorical Data. *Biometrics.* 1977;33(1):159–74. <https://doi.org/10.2307/2529310>.
32. Lehtimäki S, Martic J, Wahl B, Foster KT, Schwalbe N. Evidence on Digital Mental Health Interventions for adolescents and Young people: systematic overview. *JMIR Mental Health.* 2021;8(4):e25847. <https://doi.org/10.2196/25847>.
33. Levis B, Sun Y, He C, Wu Y, Krishnan A, Bhandari PM, Neupane D, Imran M, Brehaut E, Negeri Z, Fischer FH, Benedetti A, Thombs BD, Levis L, Riehm A, Saadat K, Azar N, Zhang M, Y. Accuracy of the PHQ-2 alone and in Combination with the PHQ-9 for screening to detect Major Depression: systematic review and Meta-analysis. *JAMA.* 2020;323(22):2290–300. <https://doi.org/10.1001/jama.2020.6504>.
34. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *J Psychosom Res.* 2005;58(2):163–71. <https://doi.org/10.1016/j.jpsychores.2004.09.006>.
35. Lu W. Adolescent depression: national trends, risk factors, and healthcare disparities. *Am J Health Behav.* 2019;43(1):181–94. <https://doi.org/10.5993/AJHB.43.1.15>.
36. Magson NR, Freeman JYA, Rapee RM, Richardson CE, Oar EL, Fardouly J. Risk and protective factors for prospective changes in adolescent mental health during the COVID-19 pandemic. *J Youth Adolesc.* 2021;50(1):44–57. <https://doi.org/10.1007/s10964-020-01332-9>.
37. Moustafa AA, Tindle R, Frydecka D, Misiak B. Impulsivity and its relationship with anxiety, depression and stress. *Compr Psychiatr.* 2017;74:173–9. <https://doi.org/10.1016/j.comppsy.2017.01.013>.
38. Muñoz-Navarro R, Malonda E, Llorca-Mestre A, Cano-Vindel A, Fernández-Berrocal P. Worry about COVID-19 contagion and general anxiety: moderation and mediation effects of cognitive emotion regulation. *J Psychiatr Res.* 2021;137:311–8. <https://doi.org/10.1016/j.jpsychores.2021.03.004>.
39. Niermann HCM, Voss C, Pieper L, Venz J, Ollmann TM, Beesdo-Baum K. Anxiety disorders among adolescents and young adults: prevalence and mental health care service utilization in a regional epidemiological study in Germany. *J Anxiety Disord.* 2021;83:102453. <https://doi.org/10.1016/j.janxdis.2021.102453>.
40. Orgilés M, Serrano-Ortiz M, Espada JP, Morales A. Back to school after the pandemic: adjustment of Spanish children and adolescents. *Ann Psychol.* 2024;40(1):1. <https://doi.org/10.6018/analesps.530471>.
41. Otto C, Reiss F, Voss C, Wüstner A, Meyrose A-K, Hölling H, Ravens-Sieberer U. Mental health and well-being from childhood to adulthood: design, methods and results of the 11-year follow-up of the BELLA study. *Eur Child Adolesc Psychiatry.* 2021;30(10):1559–77. <https://doi.org/10.1007/s00787-020-01630-4>.
42. Perugi G, Del Carlo A, Benvenuti M, Fornaro M, Toni C, Akiskal K, Dell'Osso L, Akiskal H. Impulsivity in anxiety disorder patients: is it related to comorbid cyclothymia? *J Affect Disord.* 2011;133(3):600–6. <https://doi.org/10.1016/j.jad.2011.04.033>.
43. Pretorius C, Chambers D, Coyle D. Young people's online help-seeking and mental health difficulties: systematic narrative review. *J Med Internet Res.* 2019;21(11):e13873. <https://doi.org/10.2196/13873>.
44. Proctor C, Linley PA, Maltby J. Very happy youths: benefits of very high life satisfaction among adolescents. *Soc Indic Res.* 2010;98(3):519–32. <https://doi.org/10.1007/s11205-009-9562-2>.
45. Rabner J, Mian ND, Langer DA, Comer JS, Pincus D. The relationship between worry and dimensions of anxiety symptoms in children and adolescents. *Behav Cogn Psychother.* 2017;45(2):124–38. <https://doi.org/10.1017/S1352465816000448>.
46. Reardon T, Spence SH, Hesse J, Shakir A, Creswell C. Identifying children with anxiety disorders using brief versions of the Spence Children's anxiety scale for children, parents, and teachers. *Psychol Assess.* 2018;30(10):1342–55. <https://doi.org/10.1037/pas0000570>.
47. Richardson LP, Rockhill C, Russo JE, Grossman DC, Richards J, McCarty C, McCauley E, Katon W. Evaluation of the PHQ-2 as a brief screen for detecting Major Depression among adolescents. *Pediatrics.* 2010;125(5):e1097–103. <https://doi.org/10.1542/peds.2009-2712>.



48. Rosen ML, Rodman AM, Kasperek SW, Mayes M, Freeman MM, Lengua LJ, Meltzoff AN, McLaughlin KA. Promoting youth mental health during the COVID-19 pandemic: a longitudinal study. *PLoS ONE*. 2021;16(8):e0255294. <https://doi.org/10.1371/journal.pone.0255294>.
49. Sahlqvist S, Song Y, Bull F, Adams E, Preston J, Ogilvie D, the iConnect consortium. Effect of questionnaire length, personalisation and reminder type on response rate to a complex postal survey: randomised controlled trial. *BMC Med Res Methodol*. 2011;11(1):62. <https://doi.org/10.1186/1471-2288-11-62>.
50. Scheiner C, Grashoff J, Kleindienst N, Buerger A. Mental disorders at the beginning of adolescence: prevalence estimates in a sample aged 11–14 years. *Public Health Pract*. 2022;4:100348. <https://doi.org/10.1016/j.puhip.2022.100348>.
51. Simon NM, Zalta AK, Otto MW, Ostacher MJ, Fischmann D, Chow CW, Thompson EH, Stevens JC, Demopulos CM, Nierenberg AA, Pollack MH. The association of comorbid anxiety disorders with suicide attempts and suicidal ideation in outpatients with bipolar disorder. *J Psychiatr Res*. 2007;41(3–4):255–64. <https://doi.org/10.1016/j.jpsychires.2006.08.004>.
52. Speyer LG, Eisner M, Ribeaud D, Luciano M, Auyeung B, Murray AL. Developmental relations between internalising problems and ADHD in childhood: a symptom level perspective. *Res Child Adolesc Psychopathol*. 2021;49(12):1567–79. <https://doi.org/10.1007/s10802-021-00856-3>.
53. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics (Oxford England)*. 2012;28(1):112–8. <https://doi.org/10.1093/bioinformatics/btr597>.
54. Stewart TM, Fry D, McAra L, Hamilton S, King A, Laurie M, McCluskey G. Rates, perceptions and predictors of depression, anxiety and Post traumatic stress disorder (PTSD)-like symptoms about Covid-19 in adolescents. *PLoS ONE*. 2022;17(4):e0266818. <https://doi.org/10.1371/journal.pone.0266818>.
55. Tang L, Yin R, Hu Q, Fan Z, Zhang F. The effect of childhood socioeconomic status on depressive symptoms in middle-old age: the mediating role of life satisfaction. *BMC Psychiatry*. 2022;22(1):1. <https://doi.org/10.1186/s12888-022-04046-3>.
56. Thielsch C, Ehring T, Nestler S, Wolters J, Kopei I, Rist F, Gerlach AL, Andor T. Metacognitions, worry and sleep in everyday life: studying bidirectional pathways using Ecological Momentary Assessment in GAD patients. *J Anxiety Disord*. 2015;33:53–61. <https://doi.org/10.1016/j.janxdis.2015.04.007>.
57. Tiirikainen K, Haravuori H, Ranta K, Kaltiala-Heino R, Marttunen M. Psychometric properties of the 7-item generalized anxiety disorder scale (GAD-7) in a large representative sample of Finnish adolescents. *Psychiatry Res*. 2019;272:30–5. <https://doi.org/10.1016/j.psychres.2018.12.004>.
58. Willis TA, Gregory AM. Anxiety disorders and sleep in children and adolescents. *Sleep Med Clin*. 2015;10(2):125–31. <https://doi.org/10.1016/j.jsmc.2015.02.002>.
59. Wu Q, Xu Y. Parenting stress and risk of child maltreatment during the COVID-19 pandemic: a family stress theory-informed perspective. *Dev Child Welf*. 2020;2(3):180–96. <https://doi.org/10.1177/2516103220967937>.
60. Zhang J, Paksarian D, Lamers F, Hickie IB, He J, Merikangas KR. Sleep patterns and mental health correlates in US adolescents. *J Pediatr*. 2017;182:137–43. <https://doi.org/10.1016/j.jpeds.2016.11.007>.
61. Zhou S-J, Zhang L-G, Wang L-L, Guo Z-C, Wang J-Q, Chen J-C, Liu M, Chen X, Chen J-X. Prevalence and socio-demographic correlates of psychological health problems in Chinese adolescents during the outbreak of COVID-19. *Eur Child Adolesc Psychiatry*. 2020;29(6):749–58. <https://doi.org/10.1007/s00787-020-01541-4>.
62. Zisner A, Beauchaine TP. Neural substrates of trait impulsivity, anhedonia, and irritability: mechanisms of heterotypic comorbidity between externalizing disorders and unipolar depression. *Dev Psychopathol*. 2016;28(4pt1):1177–208. <https://doi.org/10.1017/S0954579416000754>.
63. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Ser B (Stat Methodol)*. 2005;67(2):301–20. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>.

## Publisher's Note

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