Artificial Intelligence In Medicine 157 (2024) 102991

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed

Machine learning for anxiety and depression profiling and risk assessment in the aftermath of an emergency

Guillermo Villanueva Benito [a,c], Ximena Goldberg [a], Nicolai Brachowicz [a], Gemma Castaño-Vinyals [a,c,d], Natalia Blay [b], Ana Espinosa [a], Flavia Davidhi [a], Diego Torres [a], Manolis Kogevinas [a], Rafael de Cid [b], Paula Petrone [a,*]

[a] Barcelona Institute for Global Health (ISGlobal), C/ del Dr. Aiguader, 88, Barcelona 08003, Catalonia, Spain
[b] Genomes for Life-GCAT lab. CORE program. Germans Trias I Pujol Research Institute (IGTP), Camí de les Escoles, s/n, Badalona 08916, Catalonia, Spain
[c] Universitat Pompeu Fabra (UPF), Spain
[d] CIBER de Epidemiología y Salud Pública (CIBERESP), Spain

ARTICLE INFO

Keywords:
Machine learning
Mental health
COVID-19
Preparedness

ABSTRACT

Background & objectives: Mental health disorders pose an increasing public health challenge worsened by the COVID-19 pandemic. The pandemic highlighted gaps in preparedness, emphasizing the need for early identification of at-risk groups and targeted interventions. This study aims to develop a risk assessment tool for anxiety, depression, and self-perceived stress using machine learning (ML) and explainable AI to identify key risk factors and stratify the population into meaningful risk profiles.
Methods: We utilized a cohort of 9291 individuals from Northern Spain, with extensive post-COVID-19 mental health surveys. ML classification algorithms predicted depression, anxiety, and self-reported stress in three classes: healthy, mild, and severe outcomes. A novel combination of SHAP (SHapley Additive exPlanations) and UMAP (Uniform Manifold Approximation and Projection) was employed to interpret model predictions and facilitate the identification of high-risk phenotypic clusters.
Results: The mean macro-averaged one-vs-one AUROC was 0.77 ($\pm$ 0.01) for depression, 0.72 ($\pm$ 0.01) for anxiety, and 0.73 ($\pm$ 0.02) for self-perceived stress. Key risk factors included poor self-reported health, chronic mental health conditions, and poor social support. High-risk profiles, such as women with reduced sleep hours, were identified for self-perceived stress. Binary classification of healthy vs. at-risk classes yielded F1-Scores over 0.70.
Conclusions: Combining SHAP with UMAP for risk profile stratification offers valuable insights for developing effective interventions and shaping public health policies. This data-driven approach to mental health preparedness, when validated in real-world scenarios, can significantly address the mental health impact of public health crises like COVID-19.

1. Introduction

Mental health disorders are a leading cause of disease and disability worldwide. Anxiety alone has an estimated global prevalence of 4.05 % and the number of people affected increased >55 % from 1990 to 2019 [1]. According to the World Health Organization, at a global level, over 280 million people are estimated to suffer from depression, equivalent to 3.8 % of the world's population [2,3]. Notably, there is a significant treatment gap in mental health: the proportion of people with depression receiving at least minimally adequate depression treatment ranged from 22.4 % in high-income countries to 3.7 % in lower middle-income countries [3]. This treatment gap worsened in 2020, when the COVID-19 pandemic and subsequent lockdowns dramatically impacted mental health, exacerbating conditions for those already in treatment and affecting many without previous mental health issues [4]. Therefore, examining the COVID-related drivers of anxiety, depression, and stress is a key focus of this article.

Although the consequences of the pandemic on population mental health were anticipated early after the first COVID-19 outbreak, healthcare systems were poorly prepared to implement effective

* Corresponding author.
E-mail address: paula.petrone@isglobal.org (P. Petrone).

https://doi.org/10.1016/j.artmed.2024.102991
Received 2 January 2024; Received in revised form 23 September 2024; Accepted 26 September 2024
Available online 29 September 2024

mitigation measures [5]. Rather than adopting a proactive preventive approach to the emerging mental health crisis, the response was largely limited to passive case management in specialized healthcare centers. A fundamental lesson from the COVID-19 pandemic has been that health policies often lack preparedness and prevention strategies, with mental health being no exception. A treatment-focused strategy proved ineffective in addressing the sudden increase in mental health disorders. Instead of overloading specialized mental health settings, mitigation efforts should focus on integrating services into general healthcare, early identification of at-risk groups, and implementation of targeted measures [6]. Consequently, the development of a tool for identifying individuals and groups at risk of anxiety, depression, and stress—who may benefit from proactive intervention—is a primary objective of this study. Such an innovative tool would have a substantial impact not only on addressing current needs but also on preparing for future public health emergencies like COVID-19.

Developing a mental health risk assessment tool is challenging due to the multifactorial nature of mental health conditions. Over 176 non-genetic risk factors have been identified, each with a low to moderate effect (RR between 0.1 and 2.5) [7]. The range of identified exposures include lifestyle choices (e.g., smoking, physical activity), proximal factors (e.g., domestic violence, caregiving), socioeconomic conditions (e.g., gender, migrant status, unemployment), and environmental hazards (e.g., air pollution) [8]. Additionally, these factors interact in complex, bidirectional ways. A primary aim of this study is to create an automated tool that integrates these diverse exposures and socioeconomic drivers into a comprehensive mental health risk prediction model. This tool will assess the contribution of each factor to mental health disorders, guiding mental health clinicians in their assessment and thus, facilitating targeted and effective interventions.

Precision medicine is an innovative approach to disease prevention, diagnosis, and treatment that considers individual differences in genetics, environment, and lifestyle [8]. The emergence of machine learning (ML) and artificial intelligence (AI) has created new opportunities for this systemic approach to healthcare by integrating data from multiple sources to provide comprehensive risk assessments, which have proven successful for many chronic conditions [9,10]. In particular, the field of "digital psychiatry" focuses on developing ML and AI methods to assess, diagnose, and treat mental health issues [11,12]. These algorithms utilize various data sources beyond electronic health records (EHRs), including smartphones, wearables, genetics, demographics, and lifestyle factors [13–16]. However, while healthcare records offer high-quality data, their reliability decreases during emergencies such as the COVID-19 pandemic. During the pandemic, online surveys became essential for monitoring population mental health [17]. The provision of updated, individualized risk information through online surveys was crucial for effective response policies. For instance, two recent studies developed machine learning approaches to predict the impact of COVID-19 on healthcare workers' mood and identify key predictive variables related to stress [17,18]. In line with this, our study aims to integrate questionnaire data and clinical history with predictive modeling to enable personalized risk assessment for mental health outcomes, enhancing emergency preparedness across the population.

The present study capitalizes on the multifactorial aspect of the COVICAT dataset, a unique cohort of over 10,000 adult individuals from Northern Spain (Catalonia), with longitudinal data collected since 2014 and mental health surveys conducted following the COVID-19 pandemic outbreak [19,20]. By applying machine learning classification algorithms on this rich dataset, our goal is to develop a tool for individualized risk assessment of three adverse mental health outcomes: anxiety, depression, and self-perceived stress.

Many machine learning algorithms used for risk assessment are often viewed as "black boxes," making it difficult to understand the factors contributing to risk. Particularly important in biomedicine is the deployment of explainable AI (XAI) to enhance the interpretability and trustworthiness of algorithms. Clinicians and patients need to trust the AI tools to the extent that they can economize on human oversight, monitoring and verification of the system's recommendations [21–23]. In mental health, explainability is also becoming a common best practice in studies reporting machine learning algorithms for the identification of individuals at risk of mental health disorders [21]. New XAI methods provide several key benefits. First, they enhance transparency by showing which variables are important. Second, they improve trustworthiness by helping identify biases in the data. Finally, by understanding the patterns that drive predictions, these methods enable researchers to gain new insights and generate new hypotheses from the data [22,23].

In this study, we propose a novel approach for identifying the main risk phenotypes of mental health disorders, specifically anxiety, depression, and self-perceived stress. We leverage the AI explainability method SHAP (SHapley Additive exPlanations) [24], an algorithm that identifies the main determinants of mental health risk prediction by assessing feature importance. Additionally, we employ Uniform Manifold Approximation and Projection (UMAP) [25], a powerful dimensionality reduction technique, to visualize clustering patterns in high-dimensional data. The innovative combination of SHAP and UMAP allows us to stratify the population into meaningful phenotypic clusters, revealing key signatures of mental health symptoms within each risk group. This strategy enables the design of targeted interventions for participants with similar needs.

Our objective is to develop a machine learning-based risk prediction tool for anxiety, depression, and stress using multifactorial data from the COVICAT cohort, which includes lifestyle, socioeconomic, and demographic factors collected in the aftermath of the COVID-19 pandemic. This tool is designed to assess mental health risks during and after emergency situations, such as pandemics and other public health crises. We will employ explainable AI methods to identify key drivers of these disorders, using SHAP and UMAP to stratify the population into distinct risk profiles. Once validated and implemented, this tool will enable healthcare providers to deliver more effective, targeted interventions, mitigating the long-term societal and economic burden of mental health issues and improving preparedness for future emergencies.

2. Methods

2.1. Data collection

Data was sourced from COVICAT (COVID-19 cohort in Catalonia), a COVID-19 population-based cohort of the Catalan population. Its primary aim is to describe the health impact of the COVID-19 pandemic on the adult population in Catalonia. The cohort builds on pre-existing cohort studies, as previously detailed [20,26–30]. The largest proportion of the participants (88 %) were sourced from the GCAT|Genomes for Life Study. The GCAT cohort study includes middle-aged participants (40–65 years old at recruitment, 2014–2017) residing in Catalonia [31] and covers genetics and non-genetic exposure variables, including direct links to electronic healthcare registers (clinical diagnoses, treatments, and prescriptions). Participants from the pre-outbreak cohort were invited to participate in the COVICAT study, which included a survey conducted between June and November 2020, and a follow-up survey between June and August 2021. Behavior and quality of life, as well as personal data, were obtained through questionnaires, while health status data (from 2010 to 2020) was sourced by linking data collected in the public health administration databases of Catalonia. Environmental variables (e.g., exposures to air pollutants and green spaces) were estimated from the participants' residential addresses using models developed by the ELAPSE project and MODIS, respectively [28,32]. For demographic characteristics of the COVICAT cohort, refer to Table 1 and reference [20]. The mental health outcomes of interest are depression, anxiety, and self-perceived stress scores.

Table 1
Description of sociodemographic characteristics ($n = 9515$), COVICAT study.

Demographic characteristics	N(%)
Gender	
Women	5527(59.49)
Men	3764 (40.51)
Age groups	
≤49	2541(27.35)
50–59	4306(46.35)
≥60	2444(26.3)
Education	
Primary or lower	994(10.7)
Secondary	3825(42.25)
Graduate or above	4372(47.05)
COVID-19 disease	
All cases	329(4.17)
Severe COVID-19	57(0.61)
No diagnosis	8904(95.83)
Pre-pandemic mental health diagnosis	
Any diagnosis	522(5.62)
No diagnosis	8769(94.38)
Number of people in household	
Living alone	1260(13.56)
Two persons	2887(31.07)
Three persons	2424(26.09)
Four or more persons	2720(29.28)
Employment status	
Currently unemployed	2406(25.9)
Currently employed	5130(55.2)
Others	1755(18.9)

2.2. Data preprocessing

2.2.1. Data selection

Only participants who completed the self-reported web questionnaire were included. Participants who were interviewed over the phone and completed a reduced questionnaire were excluded. This selection process resulted in a dataset with a total of 9291 participants and 161 variables. Within the final dataset, only 0.34 % of the data was missing, and these missing values were distributed among a small subset of 31 variables. The majority of the variables in the dataset had complete information for all participants.

2.2.2. Categorization of mental health outcomes into severity classes

The primary objective of the risk assessment tool to be developed is to predict depression, anxiety, and self-perceived stress experienced by participants at the time of data collection using a supervised machine learning model. To achieve this, participant outcomes of interest were assessed for all participants to train and validate the model. These outcomes were measured using the Hospital Anxiety and Depression Scale (HADS) [33]. The HADS scale for depression and anxiety serves both diagnostic and severity assessment purposes through its depression and anxiety subscales. Each subscale consists of 7 items and yields scores ranging from 0 to 21. We used a threshold score of 11 for each subscale to identify severe anxiety and severe depression, while mild levels of anxiety were identified with a score of 8 and mild levels of depression with a score of 6, following the validated cut-off points for the Spanish population [34].

The global HADS score has been proposed as an effective measure of self-perceived stress or emotional distress [35]. Nevertheless, specific recommended threshold values for characterizing the severity of self-perceived stress have not yet been established. In this study, we used a threshold of 23 to identify severe self-perceived stress and a threshold of 18 to characterize mild levels of self-perceived stress. To ensure consistency, we based these threshold values on the quantiles associated with the established cut-off points for depression and anxiety.

Consequently, target outcomes—depression, anxiety, and self-perceived stress—were categorized into three levels of severity: *healthy*, *mild*, and *severe*. This categorization resulted in imbalanced datasets, with the "*severe*" category having the lowest representation. The resulting distribution of categories is shown in Table 2.

2.2.3. Imputation

Imputation was conducted using the k-Nearest Neighbors algorithm from the scikit-learn Python library. After benchmarking multiple imputation methods for classifier performance, the kNNImputer demonstrated the best performance for our relatively complete dataset.

The validity of the imputation approach was confirmed using the two-sample Kolmogorov- Smirnov statistical test [36]. Using the conventional significance threshold of p-values at 0.05, the analysis did not reveal any significant discrepancy between observed and imputed feature distributions. It is important to note that for machine learning analysis, imputation was exclusively performed on the training set. This was done without including the outcomes of interest to prevent the creation of artificial correlations between the outcomes and predictors. The imputation model developed on the training set was subsequently applied to fill in the missing data in the test set.

2.2.4. Categorical variable encoding

The dataset had 114 categorical and 47 numerical variables, respectively. Ordinal encoding was used for ordinal categorical variables, while nominal categorical variables were one-hot encoded. The resulting dataset after one-hot encoding consisted of 161 variables to be included as *features* in the machine learning models.

2.2.5. Feature scaling

We applied standard scaling to the already encoded and imputed data. Similarly, data scaling was performed on the training set and the scaling model derived based on the training set was later applied to scale the test set.

2.3. Machine learning-based predictive modeling

2.3.1. Classification tasks

For each mental health condition, we developed a three-class classification machine learning algorithm to predict the individual risk (*healthy, mild, or severe*) of adverse mental health conditions. In the multiclass classification, individuals are assigned to the severity category for which they have the highest predicted probability.

2.3.2. Machine learning classifiers

We evaluated and compared several state-of-the-art machine learning classification algorithms, including random forest (RF), extreme gradient boosting (XGBoost), support vector machine (SVM), naive Bayes (NB), multi-layer perceptron (MLP), and logistic regression (LR). All classifiers were trained using balanced class weights to address class imbalance.

2.3.3. Evaluation of the models

To evaluate algorithm performance, a nested (stratified) 5-fold cross-validation procedure was adopted. This strategy involves nesting two 5-fold cross-validation loops: the inner loop is used to optimize model hyperparameters, and the outer loop provides an unbiased estimate of

Table 2
Severity class distribution for each mental health condition.

			Number of patients (%)			
	Healthy		Mild		Severe	
Depression	6746	(72.6)	1988	(21.4)	547	(5.9)
Anxiety	6421	(69.1)	1775	(19.1)	1095	(11.8)
Self-perceived stress	6541	(70.4)	1803	(19.4)	947	(10.2)

the performance of the best model [37]. Importantly, each participant was left out and used as part of a test set fold once, while the participants in the remaining four folds were used as the training set for model learning.

Models were evaluated on the test set using the macro-averaged one-vs-one area under the receiver operating characteristic curve (AUROC_ovo), which measures the unweighted pairwise discriminability of classes and is insensitive to changes in class distribution, making it appropriate for imbalanced multiclass datasets [38]. Additional evaluation metrics include the one-vs-rest area under the receiver operating characteristic curve (AUROC_ovr), along with recall and precision values derived from multiclass confusion matrices [39]. Performance metrics were calculated for each outer cross-validation iteration and reported as mean ± standard deviation.

2.3.4. Feature selection

In machine learning, feature selection involves choosing variables or features that are informative for a given outcome based on a predictive model. Ideally, selected features should represent the variance but not be collinear or interdependent. By incorporating a feature selection step in our machine learning applications, we aim to leverage several advantages, including enhanced computational efficiency, streamlined model interpretation, and improved prediction accuracy. Moreover, this strategy helps mitigate the risk of overfitting, ensuring the robustness and reliability of our models [40].

Feature selection was performed in two steps. First, highly correlated features (Pearson's coefficient >0.85) were removed. Next, we applied mRMR (minimum Redundancy - Maximum Relevance), a filter-based feature selection method that ranks features based on their relevance for predicting each of the three categories (*healthy*, *mild*, or *severe*) of the outcomes—depression, anxiety, and self-perceived stress—and the redundancy among selected features [41]. Specifically, we used the FCQ (F-test Correlation Quotient) variant proposed in [42], where the F-statistic measures feature relevance, the Pearson correlation coefficient measures redundancy, and a quotient scheme balances relevance and redundancy.

Importantly, feature selection was performed only using the training set, resulting in a ranking of features based on the minimal-redundancy and maximal-relevance criterion. The number of selected top-ranked features was treated as a hyperparameter, and its optimal value was determined during hyperparameter tuning.

2.3.5. Hyperparameter tuning

To select the optimal hyperparameter values of each machine learning classifier, we maximized the macro averaged AUROC_ovo. We performed 50 hyperparameter optimization trials for the identification of the best hyperparameter values. The XGBoost classifier parameters were tuned using a combination of uniform and logarithmic distributions: max_depth and n_estimators were adjusted with uniform distributions, while learning_rate and reg_lambda were fine-tuned using logarithmic and uniform distributions, respectively. Additional details on parameter tuning for other models, including Random Forest, Support Vector Machine, Naïve Bayes, Multi-Layer Perceptron, and Logistic Regression, are provided in Table S1.

For comparing the performance of various machine learning classifiers, the number of selected top-ranked features was predetermined before conducting hyperparameter tuning. However, for the ultimately identified best-performing algorithm, hyperparameter tuning was executed for different numbers of top-ranked features. The final model was selected using the one-standard-error (OSE) rule [43], which identifies the model with the fewest features that is within one standard error of the best performance. This approach balances model accuracy and complexity, minimizing overfitting while ensuring high performance. The final model classification performance for each outcome is displayed in Fig. 2.

2.4. Model interpretation using SHAP

After deriving the mental health risk assessment tool for anxiety, depression, and self-perceived stress, and validating it with a separate test set, we employed SHAP (SHapley Additive exPlanations) for model interpretation. SHAP was utilized to identify the most significant drivers of these conditions by quantifying the contribution of each feature to model predictions [24]. This approach provides transparency by ranking features based on their relevance and explaining their impact on the predictions. SHAP represents a state-of-the-art method for interpreting machine learning models.

For each prediction, SHAP values were calculated to assess the importance of individual features, which can be positive or negative depending on whether they increase or decrease the probability of a particular severity category (*healthy*, *mild*, *severe*). To ensure the stability and reliability of interpretations, SHAP values were computed for the test set across all cross-validation iterations, ensuring that values were derived for each participant who was not included in the training folds. The SHAP values were computed using the Python package shap and the TreeExplainer algorithm [24]. Fig. 3 shows the most relevant features for the prediction of each outcome.

2.5. Stratification by personalized risk profile using SHAP and UMAP

After calculating individual and global risk drivers with SHAP, we next employed a recent unsupervised learning technique, Uniform Manifold Approximation and Projection (UMAP) [25], to stratify the population at risk into meaningful subgroups with similar risk profiles. UMAP, a dimensionality reduction method, helps visualize clustering patterns in high-dimensional data, providing insight into the structure of risk profiles (i.e. individuals similarly impacted by risk factors). This stratification aims to facilitate targeted interventions for identified risk groups. The process involved several key steps.

Initially, we performed post-processing of SHAP values by focusing on those computed for the *severe* mental health outcome category to identify risk factors positively associated with severe outcomes. We then selected participants with severe mental health conditions as the at-risk population and isolated SHAP values for features consistently identified across outer cross-validation folds to ensure stability. These SHAP values were visualized using UMAP to reveal clustering patterns, followed by the application of agglomerative hierarchical clustering [43] to define clusters and compute characteristic risk factors for each cluster.

A risk factor was considered characteristic of a cluster if its SHAP values were positive for at least 95 % of the cluster's members, with mean-magnitude SHAP values used as risk scores to assess significance. This approach provided a structured stratification of the population, enhancing our capability to implement targeted interventions.

As a result, we are able to: 1) identify risk factors and their associated risk scores; and 2) stratify the population in different risk profiles allowing for the characterization of the highest-risk and significantly prevalent profiles. The full machine learning workflow is illustrated in Fig. 1. Feature importance results utilizing SHAP are displayed in Fig. 3. Fig. 4 shows the population stratification by risk profile, and the characteristic features for each risk cluster.

3. Results

3.1. Classification performance

In this section, we present the results of developing a classification tool to assess the risk of mental health issues, including anxiety, depression, and self-perceived stress. We begin by comparing the performance of various classification algorithms using a 4-fold cross-validation approach. Additionally, we evaluate the impact of feature selection on classification performance through several metrics, including precision, recall, AUROC_ovo, AUROC_ovr, and F1-Score.

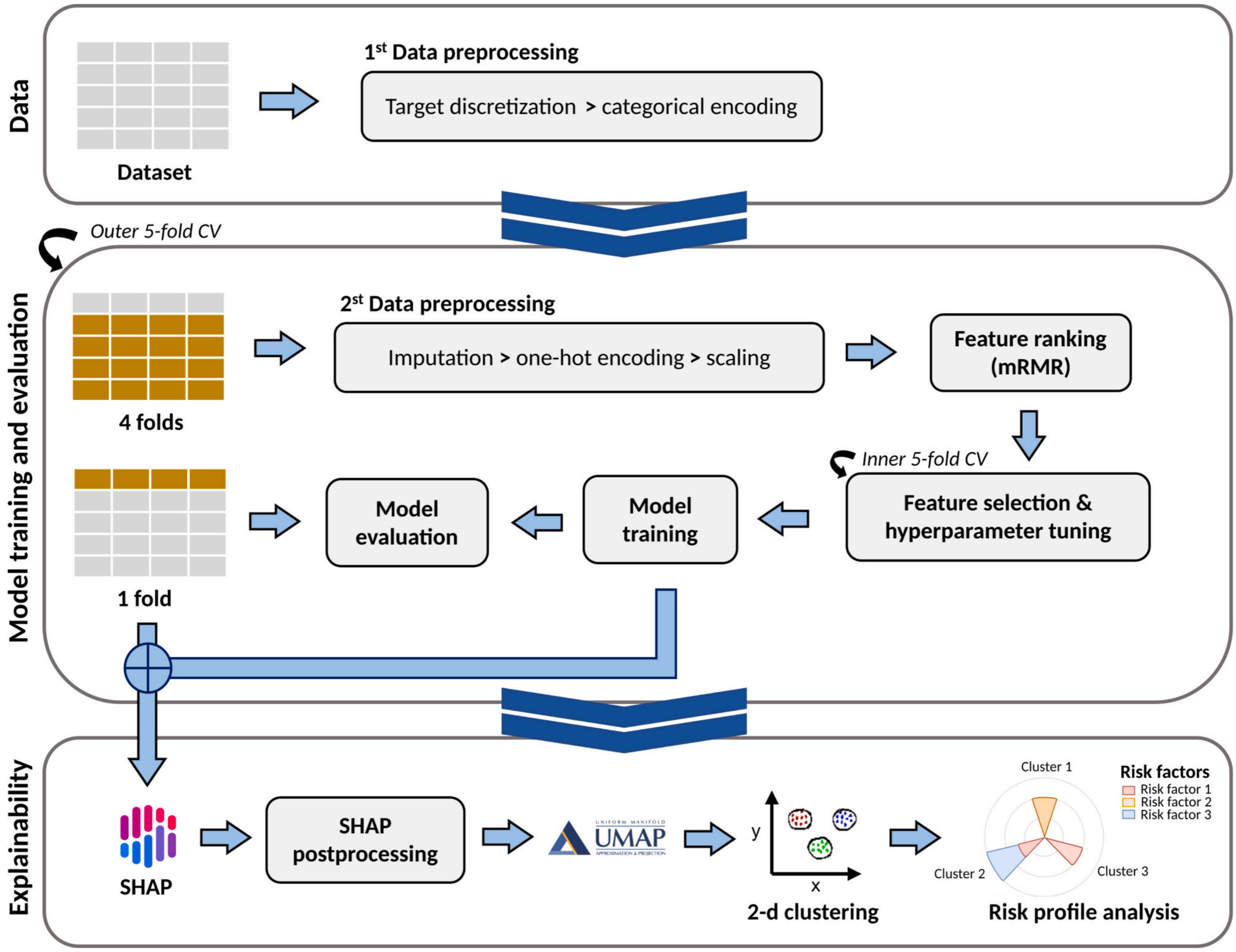

Fig. 1. Machine learning workflow. A stratified 5-fold cross-validation (CV) procedure was adopted. The procedure begins by splitting the dataset into 5 folds (outer CV): one fold is used as the validation set (20 %), while the remaining 4 folds (80 %) serve as the training set. An additional inner CV loop is employed to optimize the number of selected top-ranked features and the hyperparameters of the models. The best model is chosen based on performance on the training set folds and tested on the validation fold. This process is repeated for each of the 5 outer cross-validation iterations. SHAP values are used in combination with UMAP to explain the model's predictions and perform a cluster-based risk profile analysis.

The XGBoost algorithm proved to be the best-performing algorithm across the three outcomes as reported in Table 3 and Table 4, and was therefore chosen for the final models for feature selection optimization.

For depression, anxiety, and self-perceived stress, the predictive models demonstrated very good performance. The XGBoost models achieved macro-averaged one-vs-one (AUROC_ovo) scores of 0.780 (± 0.018) for depression, 0.723 (± 0.009) for anxiety, and 0.735 (± 0.010) for self-perceived stress. In Tables 3 and 4, we observe that there is no significant difference in classifier performance between using 100 features versus 25 features, suggesting that feature selection does not majorly impact performance in this case. However, it is noteworthy that the best results were achieved with a relatively small number of features: 13 (± 3) for depression, 19 (± 6) for anxiety, and 21 (± 4) for self-perceived stress, the variability of this feature selection resulting from an average across folds. This implies that many features may be non-informative, redundant, or potentially introducing noise. Practically, this finding suggests that the original questionnaire could be streamlined by reducing the number of questions to 20–25, simplifying the survey for respondents while still maintaining strong predictive accuracy.

Table 3
Comparison of multiple machine learning classifier performance. Values highlighted in bold indicate the highest performance. The displayed scoring metric is the macro-averaged one-vs-one area under the receiver operating characteristic curve (AUROC_ovo), presented as mean (± standard deviation) over outer cross-validation iterations. Metrics are shown for two different fixed numbers of top-ranked features: 25 and 100.

	#features 25			#features 100		
	Depression	Anxiety	Self-perceived stress	Depression	Anxiety	Self-perceived stress
Random Forest	0.772 (± 0.008)	0.717 (± 0.011)	0.729 (± 0.011)	0.774 (± 0.015)	0.721 (± 0.005)	0.731 (± 0.004)
XGBoost	**0.775 (± 0.007)**	**0.722 (± 0.014)**	**0.735 (± 0.009)**	**0.780 (± 0.018)**	**0.723 (± 0.009)**	**0.735 (± 0.010)**
Support Vector Machine	0.743 (± 0.005)	0.711 (± 0.010)	0.719 (± 0.003)	0.749 (± 0.009)	0.710 (± 0.005)	0.719 (± 0.007)
Naive Bayes	0.730 (± 0.008)	0.698 (± 0.010)	0.708 (± 0.009)	0.705 (± 0.017)	0.664 (± 0.011)	0.674 (± 0.009)
Multi-Layer Perceptron	0.762 (± 0.013)	0.722 (± 0.011)	0.724 (± 0.017)	0.775 (± 0.012)	0.718 (± 0.008)	0.730 (± 0.006)
Logistic Regression	0.764 (± 0.014)	0.720 (± 0.009)	0.731 (± 0.009)	0.771 (± 0.014)	0.721 (± 0.012)	0.734 (± 0.010)

Table 4
Comparison of machine learning classifier performance for depression. The displayed macro-averaged scoring metrics include the one-vs-one area under the receiver operating characteristic curve (AUROC_ovo), precision, recall, and F1-Score, presented as mean (± standard deviation) over outer cross-validation iterations. Metrics are shown for two different fixed numbers of top-ranked features: 25 and 100. The best-performing model (XGBoost) was determined based on the AUROC_ovo. We showed the corresponding results for anxiety and self-perceived stress in Supplementary Materials.

	#features 25				#features 100			
	AUROC_ovo	Precision	Recall	F1-Score	AUROC_ovo	Precision	Recall	F1-Score
Random Forest	0.772 (± 0.008)	0.474 (± 0.011)	0.578 (± 0.012)	0.466 (± 0.013)	0.774 (± 0.015)	0.481 (± 0.003)	0.581 (± 0.008)	0.48 (± 0.01)
XGBoost	**0.775** (± 0.007)	0.504 (± 0.006)	0.593 (± 0.008)	0.523 (± 0.008)	**0.78** **(± 0.018)**	0.505 (± 0.005)	0.591 (± 0.007)	0.525 (± 0.007)
Support Vector Machine	0.743 (± 0.005)	0.487 (± 0.008)	0.578 (± 0.008)	0.488 (± 0.022)	0.749 (± 0.009)	0.514 (± 0.011)	0.566 (± 0.025)	0.519 (± 0.01)
Naive Bayes	0.74 (± 0.008)	0.513 (± 0.063)	0.371 (± 0.015)	0.357 (± 0.027)	0.705 (± 0.017)	0.372 (± 0.118)	0.337 (± 0.003)	0.289 (± 0.006)
Multi-Layer Perceptron	0.762 (± 0.013)	0.473 (± 0.01)	0.57 (± 0.012)	0.475 (± 0.033)	0.775 (± 0.012)	0.491 (± 0.022)	0.568 (± 0.028)	0.497 (± 0.035)
Logistic Regression	0.764 (± 0.014)	0.486 (± 0.003)	0.58 (± 0.004)	0.499 (± 0.003)	0.771 (± 0.014)	0.491 (± 0.007)	0.587 (± 0.011)	0.508 (± 0.008)

In Fig. 2-A, AUROC_ovr values for each mental health condition and severity category are displayed using a boxplot. Predictive models performed best in predicting the *severe* category, with AUROC_ovr values of 0.861 (± 0.012) for depression, 0.821 (± 0.011) for anxiety, and 0.823 (± 0.013) for self-perceived stress. The *healthy* category also performed well, with AUROC_ovr values of 0.802 (± 0.011) for depression, 0.777 (± 0.014) for anxiety, and 0.783 (± 0.011) for self-perceived stress. However, the *mild* category had the lowest performance, with AUROC_ovr values of 0.685 (± 0.014) for depression, 0.628 (± 0.008) for anxiety, and 0.641 (± 0.019) for self-perceived stress. This outcome is expected, possibly due to the more distinct characteristics of severe and healthy individuals. Additionally, the ambiguity in determining mild cases using the HADS scale could contribute to these results.

Confusion matrices in Fig. 2-B1, -B2, and -B3 show moderately high recall values for the *healthy* and *severe* categories. Specifically, recall values for the *severe* category were 0.60 (± 0.02) for depression, 0.57 (± 0.01) for anxiety, and 0.55 (± 0.05) for self-perceived stress. However, recall values for the *mild* category were lower, with less than half of the participants with mild mental health outcomes being correctly identified. This finding is significant because recall indicates the model's

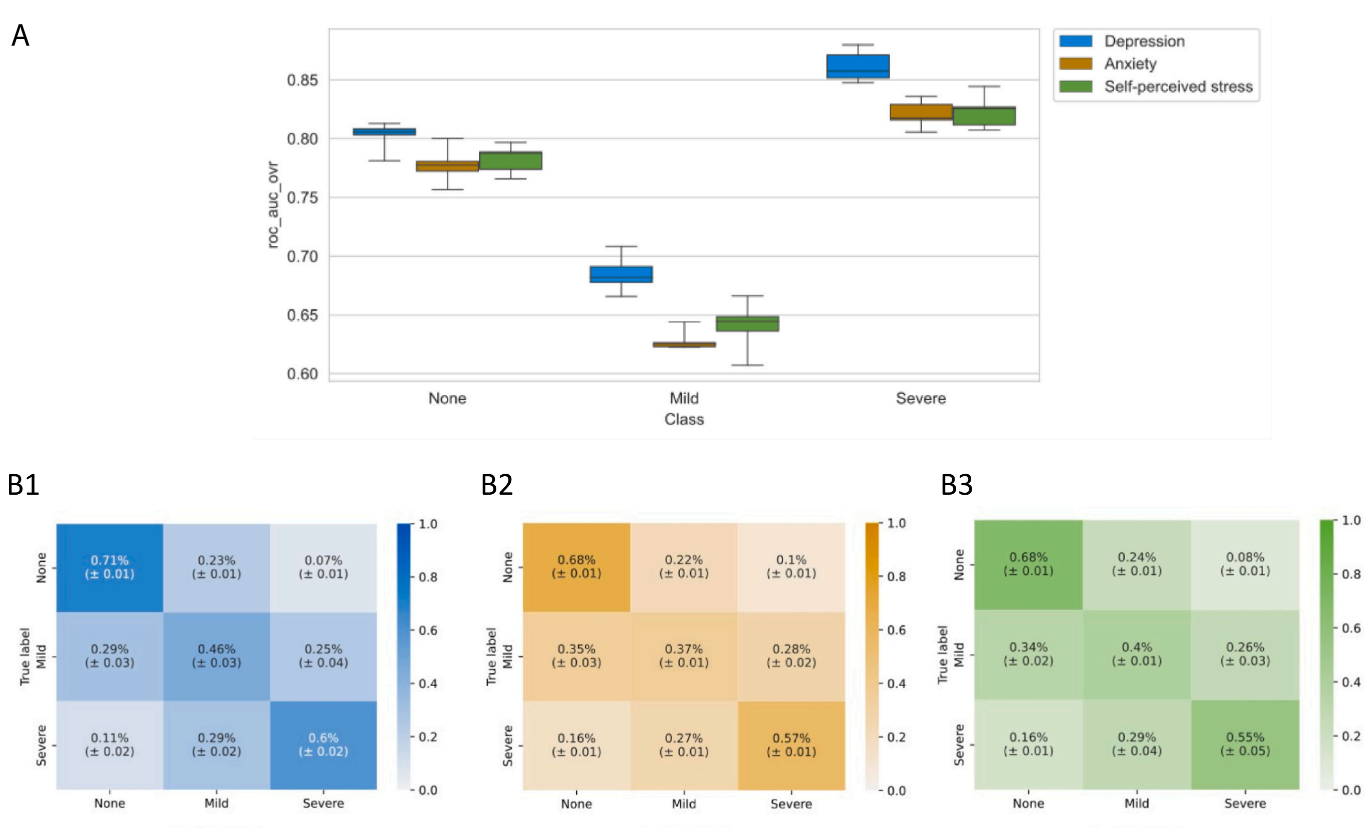

Fig. 2. Final model classification performance. Top: Boxplot of AUROC_ovr for each classification task. Performance scores are shown for each severity category (*healthy*, *mild*, and *severe*) and mental health outcome (depression, anxiety, and self-perceived stress). Bottom: Normalized (true label) confusion matrices for each classification task. Standard deviations are displayed in the form (± standard deviation) and below annotated numbers. Colors identify mental health conditions: Blue: depression; Orange: anxiety; Green: self-perceived stress. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

effectiveness in identifying all relevant instances of a specific health outcome, thus minimizing the risk of false negatives. While this may result in some false positives, high recall is crucial for a screening method like the one proposed here.

Despite achieving high precision values for the *healthy* category, the models showed low precision values for the *mild* and *severe* categories individually. Specifically, precision values were notably low for the *severe* category, achieving 0.26 ($\pm$ 0.02), 0.35 ($\pm$ 0.03), and 0.34 ($\pm$ 0.02) for depression, anxiety, and self-perceived stress respectively. Further details are given in Supplementary Materials (Supplementary Table S3). Low precision in the context of health screening suggests that a significant number of individuals identified as having a specific condition by the screening method may not actually have the condition upon further examination or confirmation. This result suggests that combining two classes (*mild* and *severe*) may yield better results.

To evaluate the classifier's performance, we calculated metrics for the combined *mild* + *severe* category, representing individuals at risk of experiencing a range of mental health outcomes. We computed the F1-Score, precision, recall, and accuracy for the binary risk assessment, as detailed in Table 5 and the Supporting Materials (see confusion matrices). The binary classifier demonstrated strong performance, with an F1-Score exceeding 0.70 and both recall and precision surpassing 0.75 and 0.70, respectively. These results suggest that the binary classification model is highly effective at identifying individuals at risk, offering enhanced efficiency and accuracy compared to models that differentiate outcomes by severity. This improved performance makes the binary classifier a potentially more valuable tool for clinicians in screening and risk assessment.

3.2. Mental health risk drivers

In this section, the objective is to understand the factors contributing to the prediction of mental health adverse outcomes. To this end, we utilize SHAP to assess the relative importance of features for predicting the *severe* category for anxiety, depression and self-perceived stress. Fig. 3-A1, -A2 and -A3 show the SHAP values computed for the *severe* category in the form of beeswarm plots for depression, anxiety and self-perceived stress, respectively. Features are ranked in descending order based on their global feature importance, and positive variable specific SHAP values indicate an increased risk of a negative mental health outcome. Fig. 3-B compares global feature importance across mental health conditions. The barplot groups informative features per type (Exposure to covid-19, Lifestyle, Health and General) and compares their informative value among the different outcomes.

The DUKE_IDX emerged as the most globally significant feature across mental health conditions. Following closely was the general health status scale (HEALTH_STATUS), identified as the second most crucial feature for depression and anxiety. For self-perceived stress, the second most globally important feature was the presence or absence of a chronic mental health condition (CHR_DEP). SHAP beeswarm plots showed that high values of DUKE_IDX, low values of HEALTH_STATUS, and high values of CHR_DEP were associated with an increased risk of a severe mental health outcome.

Changes in the physical activity (PHYSACT_CHANGES), the total number of COVID-19 symptoms by the time of the interview (N_SYMP_CUR) and whether or not sleep hours had increased, decreased or remained the same (SLEEP_CHANGES) were also found to be among the most important features for depression, anxiety and self-perceived stress, respectively. Notably, low values of PHYSACT_CHANGES, high values of N_SYMP_CUR and low values of SLEEP_CHANGES were associated with an increased risk of severe mental health outcomes. A more detailed description of the selected features is shown in Supplementary Materials (Supplementary Table S2).

Table 5
Performance metrics for binary risk assessment (*healthy* vs. *mild* + *severe*). Metrics for binary classification were derived from the confusion matrices of the three-class models shown in Fig. 2. For further details, please refer to the Supplementary Materials.

Metric	Depression	Anxiety	Self-perceived stress
Accuracy	0.75	0.71	0.72
Precision	0.72	0.70	0.70
Recall	0.80	0.75	0.75
F1-Score	0.76	0.72	0.72

3.3. Individual risk profiles analysis and stratification

In the previous sections, we developed tools for predicting depression, anxiety and self-perceived stress outcomes in the population and employed SHAP explainability to pinpoint key characteristics associated with a higher risk of adverse mental health outcomes.

In this section, we introduce a novel analytics and visualization tool designed to stratify the population into groups with similar risk factors for severe mental health outcomes, as determined by their SHAP values. This innovative tool combines SHAP with the dimensionality reduction method UMAP (Uniform Manifold Approximation and Projection). UMAP is a powerful dimensionality reduction technique that can help uncover intricate patterns and relationships within high-dimensional data. By incorporating UMAP, we gain the ability to transform complex data into a lower-dimensional representation while retaining the underlying structure of the data. This novel combination SHAP-UMAP enables us to visualize and analyze the population's risk factors for severe mental health outcomes in a more interpretable and informative manner.

This approach to population stratification is vital for the development of personalized and cost-effective solutions and treatments. People within the same group share common characteristics, making them suitable candidates for similar treatment strategies. UMAP, with its ability to reveal both local and global data structures, aids in the effective identification of such groups, enhancing the precision and affordability of tailored mental health interventions. We applied the SHAP-UMAP explainability workflow (Fig. 1) and provided a meaningful stratification of patients with a severe mental health condition.

Fig. 4-A1.1, -B1.1 and -C1.1 show UMAP two-dimensional visualizations of the population with severe outcomes for depression, anxiety and self-perceived stress, respectively. The visualizations showcase a varying number of clusters that reflect various risk profiles for each mental health condition. Specifically, we identified 9 clusters for depression, and 10 clusters for both anxiety and self-perceived stress. Clusters with a higher density of red-colored dots represent most vulnerable risk profiles. Individuals belonging to these clusters are associated with higher probabilities of experiencing severe mental health outcomes.

Fig. 4-A1.2, -B1.2 and -C1.2 show the set of identified risk factors and profiles for depression, anxiety and self-perceived stress, respectively. For each mental health condition, the identified clusters are regarded as risk profiles and each one is characterized by one to three risk factors. We quantitatively determined the most vulnerable risk profiles based on those exhibiting the highest risk scores. Interestingly, the same highest risk profiles were found for depression and anxiety: individuals with both a chronic mental health condition and poor social support (depression: cluster 6; anxiety: cluster 6) and people with both a poor or regular self-reported health status and poor social support (depression: cluster 1; anxiety: cluster 8). The latter was also found to be among the highest-risk profiles for self-perceived stress (cluster 6). Moreover, we identified two additional highest risk profiles for self-perceived stress: women whose sleep hours had decreased and, either had poor social support (cluster 3), or experienced anxiety about finding products to buy during the confinement (cluster 10). As expected, the identified high risk profiles align with the most red-colored clusters from two-dimensional UMAP embeddings. Additionally, there were clusters where no major

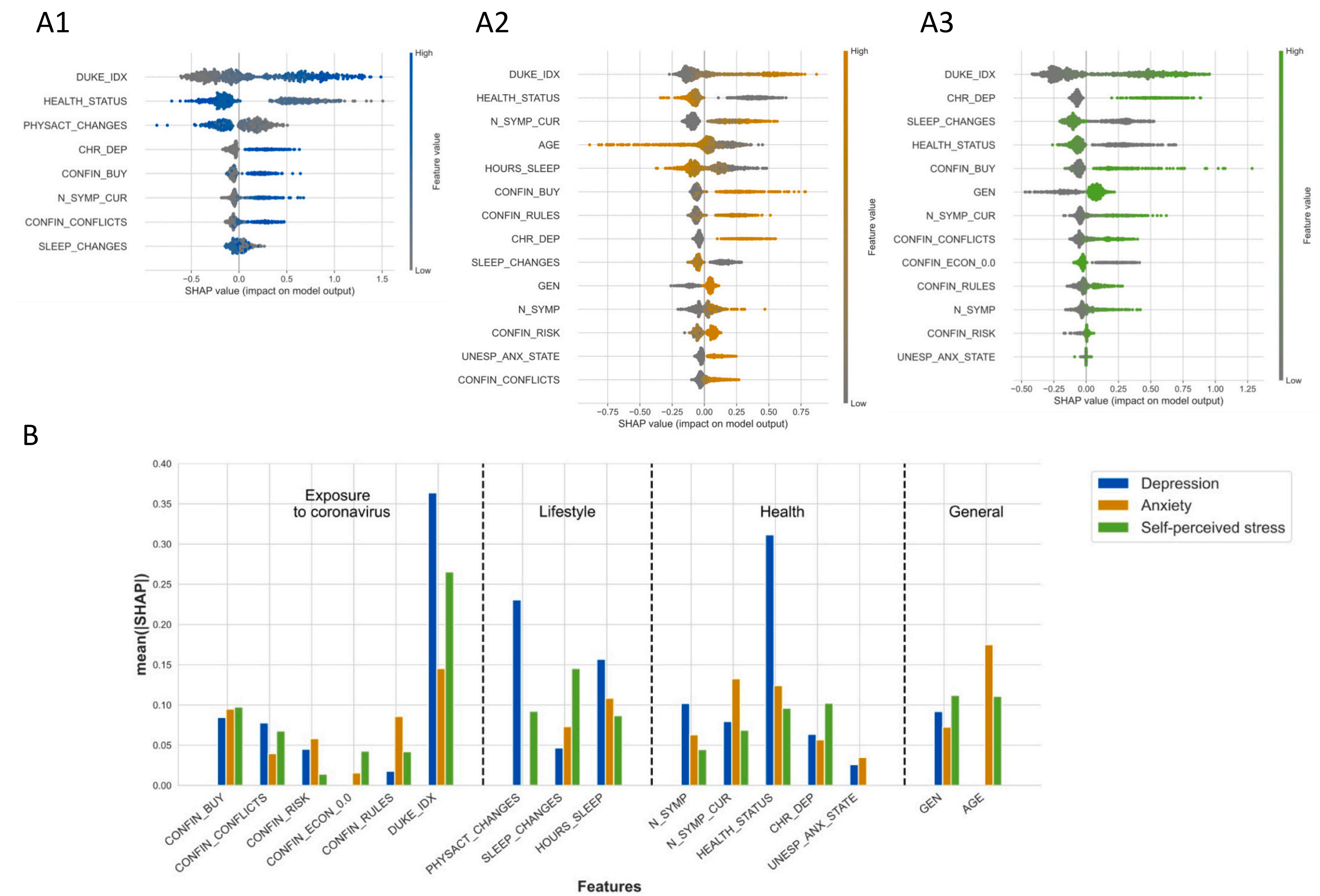

Fig. 3. SHAP values and global feature importance. Panel A: SHAP beeswarm plots for depression (A1), anxiety (A2) and self-perceived stress (A3). The SHAP beeswarm plot represents each patient by a single dot on each variable line with the color indicating the level of the respective feature value on its scale (gray indicates lower values, while colors blue, orange or green indicate higher values for depression, anxiety and self-perceived stress, respectively). The SHAP value determines the position of each data point on the x-axis. In cases where points overlap due to similar SHAP values, dots pile up along each feature row to show density. The plot also ranks the features based on their global feature importance (mean absolute SHAP value) for the model's output on the y-axis; Panel B: Bar plot comparing global feature importance (mean absolute SHAP value) between mental health outcomes. The x-axis displays the set of features that are more informative for all three mental health conditions. Features are arranged in four groups, with each group consisting of the same type of features. Blue: depression; Orange: anxiety; Green: self-perceived stress. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

characteristic risk factor was identified (depression: cluster 2; anxiety: cluster 2; and self-perceived stress: cluster 2). We observed that those clusters showed a high proportion of misclassifications (i.e. false negatives) Further details are given in Supplementary Materials (Supplementary Fig. S2).

The list of identified risk factors for depression, anxiety and self-perceived stress is presented in Fig. 4-A1.2, -B1.2 and -C1.2, respectively. Certain risk factors are shared across multiple mental health conditions, such as experiencing anxiety about not being able to find products to buy during confinement, post-pandemic COVID-19 symptoms, and reduced sleep hours. However, the significance and prevalence of these common risk factors may vary among the different mental health outcomes. Notably, decreased sleep hours emerged as a more prevalent risk factor for self-perceived stress compared to anxiety. Additionally, specific risk factors were identified for each mental health outcome: decreased physical activity and conflicts with family members during confinement for depression; difficulty maintaining rules during confinement for anxiety; and being female for self-perceived stress.

SHAP also offers insights into crucial features at the individual level. The SHAP force plots in Fig. 4-A1.3, -B1.3, and -C1.3 provide a comprehensive breakdown of the model's decision process for individual predictions, illustrating the relative importance and directionality of each feature's influence. For instance, in the case of depression (Fig. 4-A1.3), the SHAP force plot illustrates the prediction breakdown for an individual from cluster 3, representing the highest-risk profile with a prevalence of 22.4 %. This cluster is characterized by individuals with poor social support as a primary risk factor. The plot reveals how various risk factors impact this individual, including family conflicts during confinement.

In the context of anxiety (Fig. 4-B1.3), the SHAP force plot represents an individual from cluster 8, the highest-risk profile characterized by poor social support, and a poor or regular self-reported health status. Lastly, Fig. 4-C1.3 displays a SHAP force plot for an individual from cluster 3 for self-perceived stress. This plot highlights the significant driving forces of poor social support, decreased sleep hours, and being female as the most important risk factors for this specific participant.

The novelty of the tool that we present here lies in its ability to stratify participants not only by overall risk but also by shared risk drivers. This stratification enables the identification of clusters that, theoretically, could be targeted by a common intervention. Furthermore, the SHAP force plot facilitates the ongoing monitoring of risk for an individual, providing insights into the specific risk drivers that influence that person's health outcome.

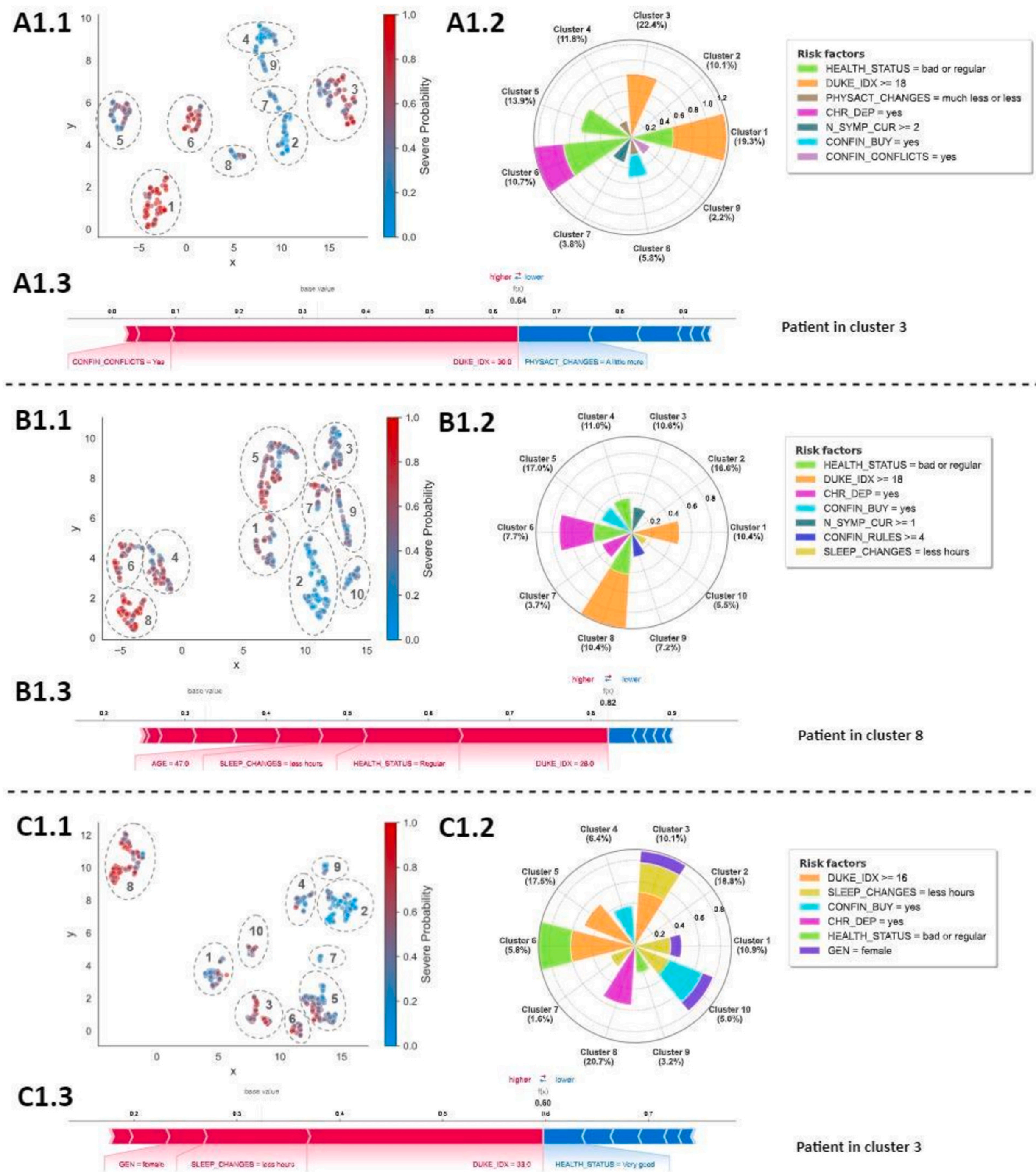

Fig. 4. Risk profiles and risk factors. Panel A1.1 shows 2-dimensional UMAP embeddings for depression. Each point represents a participant with severe depression and its color indicates the probability of a severe mental health condition. For each UMAP embedding, clusters of individuals are annotated with numbers. Panel A1.2 shows a stacked circular bar plot that displays the identified characteristic risk factors for each risk profile cluster. The radial axis represents the mean risk score (mean SHAP value). For each risk profile, risk factors are arranged in a stacked format, where the height of the stacked bar indicates the mean risk score. Stacked bar plots also show, in parentheses, the prevalence of each risk profile, i.e., the percentage of people within that cluster relative to the entire severe population. Panel A1.3 shows SHAP force plots for a representative individual belonging to an identified depression risk profile. Each feature value is represented by a red or blue arrow depending on whether it increases or decreases the predicted probability of severe depression, f(x). Similarly Panels B and C show the corresponding analyses for anxiety and self-perceived stress, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

The COVID-19 pandemic triggered a surge in mental health crises, underscoring the urgent need for tools to manage preparedness and optimize mental health resources [44]. In this work, we capitalize on a unique dataset of 9200 individuals to train and validate an AI algorithm for individualized risk assessment of adverse mental health outcomes, including anxiety, depression, and self-perceived stress. Additionally, our use of explainable AI (XAI) with SHAP allows us to identify the main informative factors for these outcomes. Finally, we utilize UMAP to visualize the SHAP values and stratify the population into highly vulnerable population subgroups, enabling the design of targeted interventions for those individuals with similar needs.

We developed three classification models to assess depression, anxiety, and self-perceived stress, categorizing individuals into three risk levels: *healthy*, *mild*, and *severe*. The most effective classifiers utilized an XGBoost engine with hyperparameter tuning, similar to other recent works in the field (e.g. [45]). All three models demonstrated strong performance, with AUROC_ovo scores of 0.78 for depression, 0.72 for anxiety, and 0.74 for self-perceived stress. These results align with published classifiers for mental health. For example, models for the open-access ADHD-200 dataset reported AUROCs ranging from 0.70 to 0.90 [46]. However, in our case, we use a three-class classifier and the metrics reported include precision, recall, F1-Score and the AUROC_ovr and AUROC_ovo which provide a comprehensive insight into the classifier's ability to distinguish between different risk levels.

Notably, our classifier excels at identifying the *severe* and *healthy* categories for all mental health outcomes, as indicated by the AUROC_ovr scores. The severe category performed exceptionally well, with AUROC_ovr values of 0.861 ($\pm$0.012) for depression, 0.821 ($\pm$0.011) for anxiety, and 0.823 ($\pm$0.013) for self-perceived stress. Similarly, the healthy category also showed strong performance, with AUROC_ovr values of 0.802 ($\pm$0.011) for depression, 0.777 ($\pm$0.014) for anxiety, and 0.783 ($\pm$0.011) for self-perceived stress. However, the models struggled most with predicting the mild category, which experienced the highest rate of misclassifications. This may be due to the *mild* class being less distinctly characterized by the cognitive and stress tests used as the gold standard. Additionally, the ambiguity in determining mild cases using the HADS scale could contribute to these results.

The most clinically relevant aspect is the prediction of the population at risk, which includes the combination of *mild* and *severe* classes. In binary classification (*healthy* vs at risk), our models demonstrated very high precision and recall for this combined group across all outcomes (precision >0.70, recall >0.75). High recall reflects the model's effectiveness in identifying all individuals at risk of a specific health outcome, thus minimizing the number of false negatives. High recall is crucial for a screening method like the one proposed here, indicating its potential value as a clinical tool. Given the imbalanced nature of our dataset, it is also noteworthy that the binary models achieved high F1-Scores for all outcomes (F1-Score >0.72), consistent with those reported in the literature for mental health applications(e.g. Ref [47]). While these results are very promising, further work is required to evaluate the technical feasibility and clinical utility of implementing these models in real-world scenarios.

Recent studies incorporate SHAP (SHapley Additive exPlanations) values into machine learning models to illustrate how feature values impact predictive risk scores and ascertain the directionality of specific features, consequently pinpointing SHAP-based risk factors [48–51]. In our work, utilizing SHAP has provided valuable insights into the significance of individual features in predicting mental health outcomes. The clarity offered by SHAP values has enhanced our understanding of the influence and direction of these features on predictions. SHAP values have enabled us to identify and understand the importance of different features, even those not traditionally considered risk factors, in predicting susceptibility to depression, anxiety, or stress. This includes identifying feature values that can be interpreted as potential risk drivers, which align with previously reported studies. Implementing SHAP values promotes transparency in model predictions and helps identify potential biases, making it a good practice in biomedical applications, especially in mental health. This approach ensures responsible AI usage by providing clear explanations for predictions.

In our work, SHAP has identified 16 important features as drivers of mental health disorders (Fig. 3), aligning well with what is reported in the literature. It is well established that the association between social support and mental health conditions [52] exist and some studies have proved that social support plays a positive role in predicting individuals' mental health during the COVID-19 pandemic [53–55]. Loneliness has also been associated with worse mental health during the pandemic [55–57], suggesting that those experiencing loneliness may feel detached from sources of support. Consistent with previous studies, we found that a high Duke index, i.e. poor social support, was one of the most predictive and highest-risk factors across mental health conditions. In addition, having a poor or regular self-reported health status and a chronic mental health condition were among the highest-risk factors for a severe mental outcome during the pandemic. Having a pre-existing physical and psychiatric condition have been associated with increased prevalence of mental health symptoms during epidemic outbreaks, including COVID-19 [57,58]. Moreover, people with chronic medical and mental illness or disability have been found to be at particular risk in disasters showing an increased vulnerability [59]. Interestingly, our findings suggest two population profiles to be highly vulnerable for the development of a severe post-pandemic depression and anxiety: 1) people having both a poor or regular self-reported health status and a chronic mental health condition; and 2) people having both a poor or regular self-reported health status and poor social support.

The observation that depression and anxiety share many risk factors has been previously reported [60]. Despite being categorized as two separate classes of disorders, they often occur comorbidly and share many common symptoms [61]. Individuals having both a poor or regular self-reported health status and poor social support were also found to constitute a highly vulnerable subpopulation for self-perceived stress. Similarly, comorbidities between self-perceived stress and other mental health conditions have also been previously reported particularly in the context of the COVID-19 pandemic [62,63].

Post-pandemic decreased physical activity and reduced sleep hours were identified as highly prevalent risk factors for depression and self-perceived stress, respectively. Our findings are aligned with previous work showing that increased physical exercise is associated with better mental health [64], including during the COVID-19 pandemic [57,65]. Additionally, both decreased sleep duration and quality can adversely impact mental health [66,67]. Previous studies have shown that COVID-19 had a negative impact on sleep disorders [68], and that the sleep quality decreased during COVID-19 outbreak [69,70]. Notably, people with poorer sleep quality might perceive themselves as sleeping fewer hours, even if the actual time spent in bed remained the same. Similarly, experiencing anxiety about finding products to buy during the confinement was also found to be a highly prevalent risk factor for anxiety. During the COVID-19 pandemic, anxiety and fear related to panic buying behaviors were particularly seen across the world [71,72]. Conversely, it was also demonstrated that anxiety and stress served as panic buying predictors [73].

The identification of female gender as a highly prevalent risk factor for self-perceived stress aligns with prior research findings, indicating that women were more susceptible to mental health adversities during the COVID-19 pandemic [74,75]. This finding is also consistent with results from previous research which concluded that women exhibited an inherently increased psychological vulnerability during natural disasters and the COVID-19 pandemic [76,77]. Interestingly, we found that women showed sleep hours were reduced, and those that experienced anxiety about finding products to buy or had poor social support were among the highest-risk profiles for experiencing severe self-perceived stress following the COVID-19 pandemic.

Consistently with previous studies [78,79], we also found that having conflicts with other family members during confinement represented a risk factor for depression. Interestingly, we also found that experiencing difficulty maintaining rules during the confinement was a risk factor for anxiety, suggesting that prolonged restrictions and guidelines could have a negative impact on mental well-being, contributing to increased burnout across the population. Finally, having post-pandemic COVID-19 symptoms emerged as a risk factor for depression and anxiety. Specifically, for depression, having more than two symptoms was associated with increased risk, whereas for anxiety, the condition was relaxed to having at least one symptom. Our findings might be related with previous results showing that people with long COVID experienced a range of factors that negatively affect their mental health and well-being [80,81].

This multidisciplinary approach paves the way for more precise and effective mental health interventions. AI can assist both clinicians and patients in recognizing the need for clinical help through concise self-assessments. Our findings demonstrate that a shorter questionnaire, consisting of approximately 20 questions, is sufficient to maintain predictive power. This is consistent with previous studies that have highlighted the effectiveness of abbreviated questionnaires when combined with machine learning techniques [14,82]. However, our study is unique in that it not only identifies key risk factors but also stratifies the population into distinct risk phenotypes, offering a more detailed understanding of mental health needs.

Our study introduces a novel approach by combining SHAP with UMAP, an advanced technique for dimensionality reduction. By applying hierarchical clustering on SHAP local explanations within two-dimensional UMAP embeddings, we achieve a meaningful stratification of the at-risk population, revealing distinct subgroups with shared risk factors. These findings have significant potential for designing targeted interventions tailored to specific subgroups within the population. For instance, implementing personalized measures aimed at increasing sleep hours among women experiencing post-pandemic self-perceived stress could lead to notable improvements in mental health [90]. Similarly, targeted interventions focused on enhancing social activity and support, especially among vulnerable populations with prior health issues, could effectively alleviate the burden of mental health outcomes.

Our study has several limitations and our results need to be taken with caution at the time of implementing these tools in real-world scenarios. Firstly, data limitations involve those derived from the COVICAT cohort, which includes an overrepresentation of individuals aged 45 and above. Consequently, our data do not cover younger people, who are a particularly vulnerable group and showed a significant increase in the prevalence of depression and anxiety after the initial outbreak [83,89]. In addition, our data shows an overrepresentation of people with a college degree or higher. Many studies document the positive association between more schooling and higher income, better health, and longer life [91]. Better educated people may have had better resources to deal with the challenges posed by the pandemic. Moreover, the COVICAT cohort consists of residents from Catalonia, a particular region of Spain. The local aspects and cultural characteristics of this area may introduce additional biases into the model. Cultural factors, regional healthcare practices, and local economic conditions can significantly influence mental health outcomes and the effectiveness of interventions. As such, the insights and predictions generated by our model may not fully generalize to populations outside the specific region where the model was trained.

Secondly, while the technical approach is robust and can be easily scaled to include data from a larger population, this model has only been cross-validated within this specific cohort. To be effectively deployed in clinical practice, the model needs to be retrained with more diverse data and tested in other independent cohorts to further validate its generalizability across different populations. This includes incorporating data from various age groups, educational backgrounds, and geographic locations to mitigate potential biases. By expanding the dataset to encompass a broader range of demographic and cultural backgrounds, the model's accuracy and applicability across different settings can be significantly improved, reducing the impact of cohort-specific biases.

Thirdly, our study focused solely on identifying predictive risk factors and did not analyze protective risk factors, although a similar study could be conducted. Additionally, the list of identified risk factors should not be interpreted as an exhaustive compilation of all potential risk factors for adverse mental health conditions in the aftermath of the pandemic. Similarly, the identified risk profiles should not be regarded as comprehensive population stratifications but rather as meaningful subsets highlighting the most predictable and significant risk profiles. Moreover, it is important to note that while SHAP values can reveal associations learned from the data, they do not necessarily guarantee or reflect causal relationships. Finally, although we achieved high classification performance for both the *healthy* and *severe* categories, the *mild* category exhibited low performance, reflecting the difficulty of accurately classifying patients with mild mental health conditions, and raising questions about the feasibility of accurately diagnosing mild mental health outcomes.

Finally, in this study, although we do not have longitudinal data to directly measure the impact of the pandemic on all the lifestyle and demographic factors considered, we addressed this by including questions on perceived changes in areas such as buying habits, employment, physical exercise, health, sleep, nutrition,and several living conditions compared to the pre-COVID period considered as baseline. These perceptions provide indirect insights into how individuals feel their lives have been affected by the pandemic. While not a direct measure of change, this approach offers valuable context and helps us understand the broader impacts of the pandemic on mental health.

Despite the promising potential of AI in mental health, caution is warranted regarding its implementation. The technology presents challenges related to data safety and security, patient autonomy, data biases, ethical issues, language barriers, and the lack of accountability when errors occur, which may go unnoticed despite efforts in explainability and transparency [84–86]. Therefore, these tools should be used as complementary decision-support systems with a “human-in-the-loop” approach, that is, ensuring they are supervised and managed by trained practitioners to maintain responsible use and oversight [87,88]. With these regulatory measures in place, AI can significantly enhance mental health care by improving access and efficiency for a broader range of the population.

5. Conclusion

We developed a personalized risk assessment algorithm, utilizing a comprehensive questionnaire and clinical history data, aimed at reducing the misdiagnosis of mental health conditions in the aftermath of the COVID-19 pandemic. By identifying key risk drivers, we demonstrate that accurate prediction of mental health conditions can be achieved with a reduced questionnaire. Moreover, our novel combination of SHAP and UMAP visualization techniques provides a deeper understanding and interpretation of the model's predictions, enabling the identification of vulnerable population subgroups and their associated risk factors. We anticipate that our work and the personalized risk assessment and stratification tools developed will help set priorities for public health policies and facilitate the implementation of effective, targeted healthcare interventions tailored to specific subgroups within the population, particularly in preparation for future pandemic outbreaks.

Ethical approval

This study was performed in line with the principles of the Declaration of Helsinki. Ethical approval was obtained by the ethics committees at the Hospital Universitari Germans Trias i Pujol (CEI no. PI-20-182) and the Parc de Salut Mar (CEIM-PS MA no. 2020/9307/I). All participants provided informed consent and had consented to be recontacted during the first follow-up.

CRediT authorship contribution statement

Guillermo Villanueva Benito: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ximena Goldberg:** Writing – original draft, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Nicolai Brachowicz:** Data curation. **Gemma Castaño-Vinyals:** Writing – review & editing, Resources, Project administration, Data curation. **Natalia Blay:** Resources, Data curation.

Ana Espinosa: Resources, Data curation. **Flavia Davidhi:** Methodology, Data curation. **Diego Torres:** Software, Formal analysis. **Manolis Kogevinas:** Writing – review & editing, Resources, Funding acquisition. **Rafael de Cid:** Writing – review & editing, Resources, Funding acquisition, Data curation. **Paula Petrone:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used OpenAI's ChatGPT 3.5 (OpenAI, Inc.) in order to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare no competing interests.

Data availability

Due to ethical and legal constraints, the dataset used in this study is derived from a cohort requiring strict approval from the ethics committee, which limits direct public access. However, the data can be made available upon request. Interested researchers should contact the corresponding author to discuss access under specific conditions that comply with ethical guidelines.

The code to reproduce the present study's findings are available at https://github.com/guillermovb/Mental_health_COVID19. All data processing and modeling were conducted on Python 3.10.11 using standard libraries that are publicly available: pandas, numpy, scipy, matplotlib, seaborn, scikit-learn, xgboost, keras, mrmr, shap and umap.

Acknowledgements

ISGlobal acknowledge support from the Spanish Ministry of Science and Innovation and State Research Agency through the “Centro de Excelencia Severo Ochoa 2019-2023” Program (CEX2018-000806-S), and the grant CEX2023-0001290-S funded by MCIN/AEI/ 10.13039/501100011033, support from the Generalitat de Catalunya through the CERCA Program” and support from the Ministry of Research and Universities of the Government of Catalonia (2021 SGR 01563).

This study makes use of data generated by the GCAT-Genomes for Life, cohort study of the Genomes of Catalonia, Fundacio IGPT. IGPT is part of the CERCA Program/Generalitat de Catalunya. GCAT was funded by Acción de Dinamización del ISCIII-MINECO and the Ministry of Health of the Generalitat of Catalunya (ADE 10/00026); and have additional support by Spanish National Grant PI18/01512; and partially support of MENARINI. Additional data included in this study was obtained in part by the COVICAT Study Group (Cohort Covid de Catalunya) supported by ISGlobal and IGTP, EIT COVID-19 Rapid Response activity 20873A and SR20-01024 La Caixa Foundation. This study was carried out using anonymized data provided by the Catalan Agency for Quality and Health Assessment, within the framework of the PADRIS Program.

The authors of the study would like to acknowledge all GCAT project investigators who contributed to the generation of the GCAT data and the COVICAT study (Cohort Covid of Catalonia). A full list of the investigators is available from \url{www.genomesforlife.com}, specially former ones, Anna Carreras and Beatriz Cortés. We thank Dr. Joan Grifols on behalf of the Blood and Tissue Bank from Catalonia (BST) and all the GCAT volunteers that participated in the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.artmed.2024.102991.

References

[1] Javaid SF, Hashim IJ, Hashim MJ, et al. Epidemiology of anxiety disorders: global burden and sociodemographic associations. Middle East Curr Psychiatry 2023;30:44.
[2] The World Health Organization. Depressive disorder (depression). https://www.who.int/news-room/fact-sheets/detail/depression; 31 March, 2023.
[3] Thornicroft G, Chatterji S, Evans-Lacko S, et al. Undertreatment of people with major depressive disorder in 21 countries. Br J Psychiatry 2017;210(2):119–24. https://doi.org/10.1192/bjp.bp.116.188078.
[4] Santomauro Damian F, Herrera Mantilla, Ana M, Shadid Jamileh, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet 2021;398(10312):1700–12. https://doi.org/10.1016/s0140-6736(21)02143-7.
[5] Holmes Emily A, O'Connor Rory C, Perry V Hugh, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. Lancet Psychiatry 2020;7(6):547–60. https://doi.org/10.1016/s2215-0366(20)30168-1.
[6] Jorm Anthony F, Patten Scott B, Brugha Traolach S, Mojtabai Ramin. Has increased provision of treatment reduced the prevalence of common mental disorders? Review of the evidence from four countries. World Psychiatry 2017;16(1):90–9. https://doi.org/10.1002/wps.20388.
[7] Arango Celso, Dragioti Elena, Solmi Marco, et al. Risk and protective factors for mental disorders beyond genetics: an evidence-based atlas. World Psychiatry 2021;20(3):417–36. https://doi.org/10.1002/wps.20894.
[8] Kline Adrienne, Wang Hanyin, Li Yikuan, et al. Multimodal machine learning in precision health: a scoping review. npj Digit Med 2022;5(1). https://doi.org/10.1038/s41746-022-00712-8. N/A.
[9] Ravaut Mathieu, Sadeghi Hamed, Leung Kin Kwan, et al. Predicting adverse outcomes due to diabetes complications with machine learning using administrative health data. npj Digi Med 2021;4(1). https://doi.org/10.1038/s41746-021-00394-8. N/A.
[10] Singh M, Kumar A, Khanna NN, Laird JR, Nicolaides A, Faa G, et al. Artificial intelligence for cardiovascular disease risk assessment in personalised framework: a scoping review. EClinicalMedicine 2024;73:102660. https://doi.org/10.1016/j.eclinm.2024.102660.
[11] Burr Christopher, Morley Jessica, Taddeo Mariarosaria, Floridi Luciano. Digital psychiatry: risks and opportunities for public health and wellbeing. IEEE Trans Technol Soc 2020;1(1):21–33. https://doi.org/10.1109/tts.2020.2977059.
[12] Chen Zhe, Kulkarni Prathamesh (Param), Galatzer-Levy Isaac R, et al. Modern views of machine learning for precision psychiatry. Journal 2022. https://doi.org/10.36227/techrxiv.19502131. N/A(N/A), N/A.
[13] Opoku Asare Kennedy, Terhorst Yannik, Vega Julio, et al. Predicting depression from smartphone behavioral markers using machine learning methods, hyperparameter optimization, and feature importance analysis: exploratory study. JMIR Mhealth Uhealth 2021;9(7):e26540. https://doi.org/10.2196/26540.
[14] Alhuwaydi Ahmed. Exploring the role of artificial intelligence in mental healthcare: current trends and future directions – a narrative review for a comprehensive insight. Risk Manag Healthc Policy 2024;17:1339–48. https://doi.org/10.2147/rmhp.s461562 (N/A).
[15] Iyortsuun Ngumimi Karen, Kim Soo-Hyung, Jhon Min, Yang Hyung-Jeong, Pant Sudarshan. A review of machine learning and deep learning approaches on mental health diagnosis. Healthcare 2023;11(3):285. https://doi.org/10.3390/healthcare11030285.
[16] Chikersal Prerna, Doryab Afsaneh, Tumminia Michael, et al. Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing. ACM Trans Comput-Human Interact 2021;28(1):1–41. https://doi.org/10.1145/3422821.
[17] Fluharty Meg, Bu Feifei, Steptoe Andrew, Fancourt Daisy. Coping strategies and mental health trajectories during the first 21 weeks of COVID-19 lockdown in the United Kingdom. Soc Sci Med 2021;279(N/A):113958. https://doi.org/10.1016/j.socscimed.2021.113958.
[18] Ćosić Krešimir, Popović Siniša, Šarlija Marko, Kesedžić Ivan, Jovanovic Tanja. Artificial intelligence in prediction of mental health disorders induced by the COVID-19 pandemic among health care workers. Croat Med J 2020;61(3):279–88. https://doi.org/10.3325/cmj.2020.61.279.
[19] Rezapour Mostafa, Hansen Lucas. A machine learning analysis of COVID-19 mental health data. Journal 2022. https://doi.org/10.21203/rs.3.rs-1129807/v1. N/A(N/A), N/A.
[20] Goldberg X, Castaño-Vinyals G, Espinosa A, et al. Mental health and COVID-19 in a general population cohort in Spain (COVICAT study). Soc Psychiatry Psychiatr Epidemiol 2022;57(12):2457–68. https://doi.org/10.1007/s00127-022-02303-0.
[21] Joyce Dan W, Kormilitzin Andrey, Smith Katharine A, Cipriani Andrea. Explainable artificial intelligence for mental health through transparency and interpretability for understandability. npj Digit Med 2023;6(1). https://doi.org/10.1038/s41746-023-00751-9. N/A.
[22] Lundberg Scott M, Erion Gabriel, Chen Hugh, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2020;2(1):56–67. https://doi.org/10.1038/s42256-019-0138-9.

[23] Minh Dang, Wang H Xiang, Li Y Fen, Nguyen Tan N. Explainable artificial intelligence: a comprehensive review. Artif Intell Rev 2022;55(5):3503–68. https://doi.org/10.1007/s10462-021-10088-y.

[24] Lu Xiaolei, Ma Jianghong, Zhang Haode. Asymmetric feature interaction for interpreting model predictions. Findings of the Assoc Comput Linguist ACL 2023 2023. https://doi.org/10.18653/v1/2023.findings-acl.286. N/A(N/A), N/A.

[25] McInnes Leland, Healy John, Saul Nathaniel, Großberger Lukas. UMAP: uniform manifold approximation and projection. J Open Source Softw 2018;3(29):861. https://doi.org/10.21105/joss.00861.

[26] Karachaliou Marianna, Moncunill Gemma, Espinosa Ana, et al. Infection induced SARS-CoV-2 seroprevalence and heterogeneity of antibody responses in a general population cohort study in Catalonia Spain. Sci Rep 2021;11(1). https://doi.org/10.1038/s41598-021-00807-4. N/A.

[27] Karachaliou Marianna, Moncunill Gemma, Espinosa Ana, et al. SARS-CoV-2 infection, vaccination and antibody response trajectories in adults: a cohort study in Catalonia. Journal 2022. https://doi.org/10.21203/rs.3.rs-1536936/v1. N/A (N/A), N/A.

[28] Kogevinas Manolis, Castaño-Vinyals Gemma, Karachaliou Marianna, et al. Ambient air pollution in relation to SARS-CoV-2 infection, antibody response, and COVID-19 disease: a cohort study in Catalonia, Spain (COVICAT study). Environ Health Perspect 2021;129(11). https://doi.org/10.1289/ehp9726. N/A.

[29] Kogevinas Manolis, Karachaliou Marianna, Espinosa Ana, et al. Long-term exposure to air pollution and COVID-19 vaccine antibody response in a general population cohort (COVICAT study, Catalonia). Environ Health Perspect 2023;131 (4). https://doi.org/10.1289/ehp11989. N/A.

[30] Delgado-Ortiz Laura, Carsin Anne-Elie, Merino Jordi, et al. Changes in population health-related behaviors during a COVID-19 surge: a natural experiment. Ann Behav Med 2023;57(3):216–26. https://doi.org/10.1093/abm/kaac054.

[31] Obón-Santacana Mireia, Vilardell Mireia, Carreras Anna, et al. GCAT|genomes for life: a prospective cohort study of the genomes of Catalonia. BMJ Open 2018;8(3): e018324. https://doi.org/10.1136/bmjopen-2017-018324.

[32] Ambler Tim, Cloud Nicholas. PM2. JavaScript frameworks for modern web dev. N/A(N/A). 2015. p. 53–72. https://doi.org/10.1007/978-1-4842-0662-1_4.

[33] Orbell Sheina, Schneider Havah, Esbitt Sabrina, et al. Hospital anxiety depression scale. Encycl Behav Med 2013:988–9. https://doi.org/10.1007/978-1-4419-1005-9_962. N/A(N/A).

[34] Terol-Cantero Maria C, Cabrera-Perona Victor, Martín-Aragón Maite. Revisión de estudios de la Escala de Ansiedad y Depresión Hospitalaria (HAD) en muestras españolas. Anales de Psicología 2015;31(2):494. https://doi.org/10.6018/analesps.31.2.172701.

[35] Cosco Theodore D, Doyle Frank, Ward Mark, McGee Hannah. Latent structure of the hospital anxiety and depression scale: a 10-year systematic review. J Psychosom Res 2012;72(3):180–4. https://doi.org/10.1016/j.jpsychores.2011.06.008.

[36] Abayomi Kobi, Gelman Andrew, Levy Marc. Diagnostics for multivariate imputations. J R Stat Soc Ser C Appl Stat 2008;57(3):273–91. https://doi.org/10.1111/j.1467-9876.2007.00613.x.

[37] Belyadi Hoss, Haghighat Alireza. Introduction to machine learning and Python. Mach Learn Guide Oil Gas Using Python 2021:1–55. https://doi.org/10.1016/b978-0-12-821929-4.00006-8. N/A(N/A).

[38] Cullmann Andreas Dominik. HandTill2001: multiple class area under ROC curve. CRAN: contributed packages. N/A(N/A), N/A.. 2012. https://doi.org/10.32614/cran.package.handtill2001.

[39] Fawcett Tom. An introduction to ROC analysis. Pattern Recogn Lett 2006;27(8): 861–74. https://doi.org/10.1016/j.patrec.2005.10.010.

[40] Cai Jie, Luo Jiawei, Wang Shulin, Yang Sheng. Feature selection in machine learning: a new perspective. Neurocomputing 2018;300(N/A):70–9. https://doi.org/10.1016/j.neucom.2017.11.077.

[41] Peng Hanchuan, Long Fuhui, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell Aug. 2005;27(8):1226–38. https://doi.org/10.1109/TPAMI.2005.159.

[42] Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. In: Computational systems bioinformatics. CSB2003. Proceedings of the 2003 IEEE bioinformatics conference. CSB2003; 2003. https://doi.org/10.1109/csb.2003.1227396. N/A(N/A), N/A.

[43] Ruppert David. The elements of statistical learning: data mining, inference, and prediction. J Am Stat Assoc 2004;99(466):567. https://doi.org/10.1198/jasa.2004.s339.

[44] Graham Andrea K, Lattie Emily G, Powell Byron J, et al. Implementation strategies for digital mental health interventions in health care settings. Am Psychol 2020;75 (8):1080–92. https://doi.org/10.1037/amp0000686.

[45] de Lacy Nina, Ramshaw Michael J, McCauley Elizabeth, et al. Predicting individual cases of major adolescent psychiatric conditions with artificial intelligence. Transl Psychiatry 2023;13(1). https://doi.org/10.1038/s41398-023-02599-9. N/A.

[46] Cao Meng, Martin Elizabeth, Li Xiaobo. Machine learning in attention-deficit/hyperactivity disorder: new approaches toward understanding the neural mechanisms. Transl Psychiatry 2023;13(1). https://doi.org/10.1038/s41398-023-02536-w. N/A.

[47] Iyortsuun Ngumimi Karen, Kim Soo-Hyung, Jhon Min, Yang Hyung-Jeong, Pant Sudarshan. A review of machine learning and deep learning approaches on mental health diagnosis. Healthcare 2023;11(3):285. https://doi.org/10.3390/healthcare11030285.

[48] Garriga Roger, Mas Javier, Abraha Semhar, et al. Machine learning model to predict mental health crises from electronic health records. Nat Med 2022;28(6): 1240–8. https://doi.org/10.1038/s41591-022-01811-5.

[49] Garcia-Gutiérrez Susana, Esteban-Aizpiri Cristobal, Lafuente Iratxe, et al. Machine learning-based model for prediction of clinical deterioration in hospitalized patients by COVID 19. Sci Rep 2022;12(1). https://doi.org/10.1038/s41598-022-09771-z. N/A.

[50] Kumar Vishnu, Sznajder Kristin K, Kumara Soundar. Machine learning based suicide prediction and development of suicide vulnerability index for US counties. npj Mental Health Res 2022;1(1). https://doi.org/10.1038/s44184-022-00002-x. N/A.

[51] Oh Taeseob, Kim Dongkyun, Lee Siryeol, et al. Machine learning-based diagnosis and risk factor analysis of cardiocerebrovascular disease based on KNHANES. Sci Rep 2022;12(1). https://doi.org/10.1038/s41598-022-06333-1. N/A.

[52] Grey I, Arora T, Thomas J, Saneh A, Tohme P, Abi-Habib R. The role of perceived social support on depression and sleep during the COVID-19 pandemic. Psychiatry Res 2020;293:113452.

[53] Liu Yang, Xie Ya-Nan, Li Wen-Gang, et al. A machine learning-based risk prediction model for post-traumatic stress disorder during the COVID-19 pandemic. Medicina 2022;58(12):1704. https://doi.org/10.3390/medicina58121704.

[54] Kandeğer Ali, Aydın Memduha, Altınbaş Kürşat, et al. Evaluation of the relationship between perceived social support, coping strategies, anxiety, and depression symptoms among hospitalized COVID-19 patients. Int J Psychiatry Med 2021;56(4):240–54. https://doi.org/10.1177/0091217420982085.

[55] Hou Tianya, Xie Yawei, Mao Xiaofei, et al. The mediating role of loneliness between social support and depressive symptoms among Chinese rural adolescents during COVID-19 outbreak: a comparative study between left-behind and non-left-behind students. Front Psychiatry 2021;12. https://doi.org/10.3389/fpsyt.2021.740094 (N/A), N/A.

[56] Hervalejo Diego, Carcedo Rodrigo J, Fernández-Rouco Noelia. Family and mental health during the confinement due to the COVID-19 pandemic in Spain: the perspective of the counselors participating in psychological helpline services. J Comp Family Stud 2020;51(3–4):399–416. https://doi.org/10.3138/jcfs.51.3-4.014.

[57] Yuan Kai, Zheng Yong-Bo, Wang Yi-Jie, et al. A systematic review and meta-analysis on prevalence of and risk factors associated with depression, anxiety and insomnia in infectious diseases, including COVID-19: a call to action. SSRN Electron J 2022. https://doi.org/10.2139/ssrn.4001811. N/A(N/A), N/A.

[58] Horesh Danny, Kapel Lev-Ari Rony, Hasson-Ohayon Ilanit. Risk factors for psychological distress during the COVID-19 pandemic in Israel: loneliness, age, gender, and health status play an important role. Br J Health Psychol 2020;25(4): 925–33. https://doi.org/10.1111/bjhp.12455.

[59] Eisenman David P, Zhou Qiong, Ong Michael, et al. Variations in disaster preparedness by mental health, perceived general health, and disability status. Disaster Med Public Health Prep 2009;3(1):33–41. https://doi.org/10.1097/dmp.0b013e318193be89.

[60] Yuan Lulu, Lu Lu, Wang Xuehang, et al. Comorbid anxiety and depressive symptoms and the related factors among international medical students in China during COVID-19 pandemic: a cross-sectional study. BMC Psychiatry 2023;23(1). https://doi.org/10.1186/s12888-023-04638-7. N/A.

[61] Kalin NH. The critical relationship between anxiety and depression. Am J Psychiatry 2020;177(5):365–7. https://doi.org/10.1176/appi.ajp.2020.20030305.

[62] Yen Cheng-Fang, Hsu Chia-Chuang. Post-traumatic stress disorder in adolescents after a natural disaster. Comprehensive guide to post-traumatic stress disorders. N/A(N/A). 2016. p. 1401–19. https://doi.org/10.1007/978-3-319-08359-9_33.

[63] Quittkat HL, Düsing R, Holtmann FJ, Buhlmann U, Svaldi J, Vocks S. Perceived impact of Covid-19 across different mental disorders: a study on disorder-specific symptoms, psychosocial stress and behavior. Front Psychol 2020;11:586246. https://doi.org/10.3389/fpsyg.2020.586246.

[64] Ashdown-Franks G, Firth J, Carney R, et al. Exercise as medicine for mental and substance use disorders: a meta-review of the benefits for neuropsychiatric and cognitive outcomes. Sports Med 2020;50:151–70. https://doi.org/10.1007/s40279-019-01187-6.

[65] Creese Byron, Khan Zunera, Henley William, et al. Loneliness, physical activity, and mental health during COVID-19: a longitudinal analysis of depression and anxiety in adults over the age of 50 between 2015 and 2020. Int Psychogeriatr 2021;33(5):505–14. https://doi.org/10.1017/s1041610220004135.

[66] Scott AJ, Webb TL, Martyn-St James M, Rowse G, Weich S. Improving sleep quality leads to better mental health: a meta-analysis of randomised controlled trials. Sleep Med Rev 2021;60:101556. https://doi.org/10.1016/j.smrv.2021.101556.

[67] Blackwelder Amanda, Hoskins Mikhail, Huber Larissa. Effect of inadequate sleep on frequent mental distress. Prev Chronic Dis 2021;18(N/A). https://doi.org/10.5888/pcd18.200573. N/A.

[68] Krishnamoorthy Yuvaraj, Nagarajan Ramya, Saya Ganesh Kumar, Menon Vikas. Prevalence of psychological morbidities among general population, healthcare workers and COVID-19 patients amidst the COVID-19 pandemic: a systematic review and meta-analysis. Psychiatry Res 2020;293(N/A):113382. https://doi.org/10.1016/j.psychres.2020.113382.

[69] Targa Adriano DS, Benítez Iván D, Moncusí-Moix Anna, et al. Decrease in sleep quality during COVID-19 outbreak. Sleep Breathing 2021;25(2):1055–61. https://doi.org/10.1007/s11325-020-02202-1.

[70] Guo Yang-feng, Liao Min-qi, Cai Wei-li, et al. Physical activity, screen exposure and sleep among students during the pandemic of COVID-19. Sci Rep 2021;11(1). https://doi.org/10.1038/s41598-021-88071-4. N/A.

[71] Sim Kang, Chua Hong Choon, Vieta Eduard, Fernandez George. The anatomy of panic buying related to the current COVID-19 pandemic. Psychiatry Res 2020;288 (N/A):113015. https://doi.org/10.1016/j.psychres.2020.113015.

[72] Arafat SM, Yasir Kar, Kumar Sujita, Kabir Russell. Possible controlling measures of panic buying during COVID-19. Int J Mental Health Addict 2021;19(6):2289–91. https://doi.org/10.1007/s11469-020-00320-1.
[73] Lins Samuel, Koch Rita, Aquino Sibele, de Freitas Melo Cynthia, Costa Icaro Moreira. Anxiety, depression, and stress: can mental health variables predict panic buying? J Psychiatr Res 2021;144(N/A):434–40. https://doi.org/10.1016/j.jpsychires.2021.11.008.
[74] Almeida Marcela, Shrestha Angela D, Stojanac Danijela, Miller Laura J. The impact of the COVID-19 pandemic on women's mental health. Arch Womens Ment Health 2020;23(6):741–8. https://doi.org/10.1007/s00737-020-01092-2.
[75] Morin Charles M, Bjorvatn Bjørn, Chung Frances, et al. Insomnia, anxiety, and depression during the COVID-19 pandemic: an international collaborative study. Sleep Med 2021;87(N/A):38–45. https://doi.org/10.1016/j.sleep.2021.07.035.
[76] Leighton Caroline, Martínez Claudio. Gender and depression: women, transgender, and gender nonconforming depression. Depression Personal 2021;N/A(N/A): 281–311. https://doi.org/10.1007/978-3-030-77329-8_15.
[77] Tang Bihan, Liu Xu, Liu Yuan, Xue Chen, Zhang Lulu. A meta-analysis of risk factors for depression in adults and children after natural disasters. BMC Public Health 2014;14(1). https://doi.org/10.1186/1471-2458-14-623. N/A.
[78] Guo Yan, Cheng Chao, Zeng Yu, et al. Mental health disorders and associated risk factors in quarantined adults during the COVID-19 outbreak in China: cross-sectional study. J Med Internet Res 2020;22(8):e20328. https://doi.org/10.2196/20328.
[79] Mohanty Jayashree, Chokkanathan Srinivasan, Alberton Amy M. <scp>COVID</scp>-19–related stressors, family functioning and mental health in Canada: test of indirect effects. Family Relat 2022;71(2):445–62. https://doi.org/10.1111/fare.12635.
[80] Burton Alexandra, Aughterson Henry, Fancourt Daisy, Philip Keir EJ. Factors shaping the mental health and well-being of people experiencing persistent COVID-19 symptoms or 'long COVID': qualitative study. BJPsych Open 2022;8(2). https://doi.org/10.1192/bjo.2022.38. N/A.
[81] Goss Charles W, Duncan Jennifer G, Lou Sunny S, et al. Effects of persistent exposure to COVID-19 on mental health outcomes among trainees: a longitudinal survey study. J Gen Intern Med 2022;37(5):1204–10. https://doi.org/10.1007/s11606-021-07350-y.
[82] Tutun Salih, Johnson Marina E, Ahmed Abdulaziz, et al. An AI-based decision support system for predicting mental health disorders. Inf Syst Frontiers 2023;25(3):1261–76. https://doi.org/10.1007/s10796-022-10282-5.
[83] Shvetcov Artur, Whitton Alexis, Kasturi Suranga, et al. Machine learning identifies a COVID-19-specific phenotype in university students using a mental health app. Journal 2022. https://doi.org/10.1101/2022.12.07.22283234. N/A(N/A), N/A.
[84] Kretzschmar Kira, Tyroll Holly, Pavarini Gabriela, et al. Can your phone be your therapist? Young People's ethical perspectives on the use of fully automated conversational agents (Chatbots) in mental health support. Biomed Inf Insights 2019;11(N/A):117822261982908. https://doi.org/10.1177/1178222619829083.
[85] Doraiswamy Sathyanarayanan, Abraham Amit, Mamtani Ravinder, Cheema Sohaila. Use of telehealth during the COVID-19 pandemic: scoping review. J Med Internet Res 2020;22(12):e24087. https://doi.org/10.2196/24087.
[86] Chiruvella Varsha, Guddati Achuta Kumar. Ethical issues in patient data ownership. Interact J Med Res 2021;10(2):e22269. https://doi.org/10.2196/22269.
[87] Terra Mohamed, Baklola Mohamed, Ali Shaimaa, El-Bastawisy Karim. Opportunities, applications, challenges and ethical implications of artificial intelligence in psychiatry: a narrative review. Egypt J Neurol Psychiatry Neurosurg 2023;59(1). https://doi.org/10.1186/s41983-023-00681-z. N/A.
[88] Char Danton S, Abràmoff Michael D, Feudtner Chris. Identifying ethical considerations for machine learning healthcare applications. Am J Bioeth 2020;20(11):7–17. https://doi.org/10.1080/15265161.2020.1819469.
[89] Tan Xiangmin, He Yuqing, Ning Ni, et al. Shared decision-making in the treatment of adolescents diagnosed with depression: a cross-sectional survey of mental health professionals in China. J Psychiatr Ment Health Nurs 2024;31(3):340–51. https://doi.org/10.1111/jpm.12990.
[90] Liu Hui, Kong Luya, Sun Qian, Ma Xiaofeng. The effects of mindfulness-based interventions on nurses' anxiety and depression: a meta-analysis. Nurs Open 2023;10(6):3622–34. https://doi.org/10.1002/nop2.1610.
[91] Zajacova Anna, Lawrence Elizabeth M. The relationship between education and health: reducing disparities through a contextual approach. Annu Rev Public Health 2024;39(1):273–89. https://doi.org/10.1146/annurev-publhealth-031816-044628.