



Development and Validation of a Simple Tool for Predicting Pandemic-Related Psychological Distress Among Health Care Workers

Kristina Adorjan^{1,2,3} · Mark Sen Dong¹ · Paul R. Wrtil^{4,5} · Niklas A. Schmacke⁶ · Tobias Weinberger^{7,8} · Julius Steffen^{7,8} · Andreas Osterman^{4,5} · Alexander Choukér⁹ · Tonina T. Mueller^{7,8} · Tarek Jebrini¹ · Hauke Felix Wiegand¹⁰ · Oliver Tüscher^{10,13,14} · Klaus Lieb^{10,13} · Veit Hornung⁶ · Peter Falkai¹ · Matthias Klein^{11,12} · Oliver T. Keppler^{4,5} · Nikolaos Koutsouleris¹

Received: 27 July 2023 / Revised: 18 November 2023 / Accepted: 20 December 2023
© The Author(s) 2024

Abstract

We aimed to develop a simple predictive model that enables health care workers (HCWs) to self-assess pandemic-related psychological distress in order to assist them to find psychological support to avert adverse distress-related outcomes. In a pilot study, we recruited and followed longitudinally 220 HCWs at the Hospital of the Ludwig Maximilian University Munich (H-LMU) during the first wave of the COVID-19 pandemic (March–July 2020). In this sample, we evaluated whether a machine-learning model with sociodemographic, epidemiological, and psychological data could predict levels of pandemic-related psychological distress. To maximise clinical utility, we derived a brief, 10-variable model to monitor distress risk and inform about the use of individualised preventive interventions. The validity of the model was assessed in a subsequent cross-sectional study cohort (May–August 2020) consisting of 7554 HCWs at the H-LMU who were assessed for depressiveness after the first wave of the pandemic. The model predicted psychological distress at 12 weeks with a balanced accuracy (BAC) of 75.0% (sensitivity, 73.2%; specificity, 76.8%) and an increase in prognostic certainty of 41%. In the derivation cohort, the brief model maintained a BAC of 75.6% and predicted depressiveness ($P < .001$), resilience ($p < .001$), and coping ($p < .001$). Furthermore, it accurately stratified HCWs' psychological trajectories of global and affective burden as well as behavioural adaptation over the 12-week follow-up period. Our clinically scalable, 10-variable model predicts individual COVID-19 pandemic-related psychological distress outcomes. HCWs may use our associated predictive tool to monitor personal and team-based risk and learn about risk preventive interventions based on an intuitive risk stratification.

Keywords SARS-CoV-2 · Health care workers · Pandemic · Distress · Mental health · Risk prediction · Intervention · Translation

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused more than 641 million confirmed infections and 6,632,193 deaths from coronavirus disease 2019 (COVID-19) worldwide (as of November 29, 2022) (COVID-19 Map, n.d.). A pandemic of such magnitude poses considerable challenges to health care systems. Not only does it stretch hospitals to their limits in their capacity to care for patients (Butler et al., 2020), but it places health care workers (HCWs) at severe risk. In particular, in the early phase of the pandemic, HCWs faced

considerable psychological pressure from inadequate personal protective equipment and a high risk of exposure and infection, overwork, stigmatisation, isolation, frustration, lack of contact with families, and exhaustion (Kang et al., 2020). Thus, in this pandemic context, HCWs often experienced mental health problems, such as distress, anxiety, depressive symptoms, insomnia, and fear (Gilan et al., 2020; Lai et al., 2020). These problems may not only affect HCWs' concentration, understanding, and decision-making abilities but also have long-term impacts on their overall well-being (Kang et al., 2020).

Many studies on HCWs in the early phase of the COVID-19 pandemic have provided data on the prevalence of anxiety (13 to 70%) (Cai et al., 2020; Lu et al., 2020), depressive symptoms (12 to 50%) (Firew et al., 2020; Huang et al., 2020; Lai et al., 2020), and post-traumatic stress symptoms

Kristina Adorjan and Mark Sen Dong contributed equally.

Extended author information available on the last page of the article

(27 to 72% (Lai et al., 2020; Huang et al., 2020). Sleep disturbances were reported in 24 to 38% of employees (Huang et al., 2020; Kang et al., 2020; Lai et al., 2020) and one study found stress symptoms in 22% of respondents (Mo et al., 2020). Most studies identified similar risk factors for psychological distress, i.e., contact with patients with SARS-CoV-2, female sex, reduced (perceived) health status, worries about family members, and poor sleep quality (Rossi et al., 2020; Xiao et al., 2020). To date, several COVID-19 pandemic-related psychological stress scales have already emerged or validated. The COVID-19 Phobia Scale assesses DSM-V specific phobia criteria relating to the pandemic (Arpaci et al., 2020). The Coronavirus Anxiety Scale (Lee, 2020) was highly reliable and showed relationships between COVID-19 diagnosis, history of anxiety, COVID-19 feat and function all impairments. The COVID-19 Anxiety Syndrome Scale (Nikčević & Spada, 2020) identifies maladaptive coping, avoidance, checking, worrying and threat monitoring associated with COVID-19. However, none of the existing scales were developed specifically for the HCWs and provided predictive capability for the prevention of psychological stress. Furthermore, predictive tools for psychological distress are scarce because studies were limited by heterogeneous designs, different definitions of outcomes and, furthermore, a lack of internal tests of generalizability and external validity (Kang et al., 2020; Mo et al., 2020; Xiao et al., 2020).

Robust machine-learning approaches have shown promising results in outcome prediction (Chekroud et al., 2016; Koutsouleris et al., 2021) across various risk assessment applications (Burkhardt et al., 2020; Chand et al., 2020). They offer significant advantages over traditional statistical methods by examining multiple predictive variables and identify multi-dimensional interactions between them (Walter et al., 2019). Thus, an accurate machine-learning tool for self-assessment of pandemic-related psychological distress could aid affected HCWs to select risk-adaptive preventive support by following a stepwise intervention model.

In this work, we employed machine learning to investigate the predictive value of sociodemographic, epidemiological, and psychological variables in two different studies (Care Corona Immune Study (CCI): longitudinal pilot study, N=220; All Corona Care Study (ACC): cross-sectional validation study, N=7554) measuring pandemic-related psychological distress in HCWs at the Hospital of the Ludwig Maximilian University Munich (H-LMU), Munich, Germany. We aimed to develop a simple and scalable tool that uses baseline variables to estimate individualised risk for adverse affective outcomes, empowering HCWs to seek individualised psychological support to mitigate their personal risk. Behavioral health care should be achieved that addresses prevention, detection, and early intervention of mental health problems among HCWs.

Methods

To report on the derivation and validation of our predictive tool, we followed the internationally established Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis recommendations (Collins et al., 2015; Rector et al., 2012; Bertram & Hambleton, 2016).

Data Source and Study Participants

We obtained data from the populations of two studies performed at H-LMU, the longitudinal pilot study (CCI; N=220) [Supplementary information C1] and the cross-sectional validation study (ACC; N=7554) [Supplementary information C2]. Both had investigated the effects of the COVID-19 pandemic on HCWs (Weinberger et al., 2021; Wratil et al., 2022). We trained and developed the machine learning models on the CCI dataset and applied to the ACC dataset to evaluate the models' generalisability and construct validity.

Predictor Variables and Outcomes

We extracted all 39 features from the pilot study's dataset for the machine learning analyses [Supplementary information T1]. To define the target variable for prediction, we summed the participants' scores across all pandemic-related psychological-behavioural stress items from CCI questionnaire recorded at 12 weeks after the beginning of the study: 'worried about health' (CCI-25), 'feeling distressed' (CCI-27), 'developed obsessive behaviours' (CCI-23), 'suffered anxiety and somatic disturbances' (CCI-26), 'suffered sleep disturbances' (CCI-24) and 'haunted by intrusions and nightmares' (CCI-22). We defined our prediction targets as 'distressed' (equal to or greater than the 75% percentile of the summary distress score) or 'non-distressed' (lower than the 75% percentile). The cut-off percentile was based on an earlier study where COVID-19 pandemic-related distress was observed in 22% of HCWs (Mo et al., 2020).

Machine Learning Strategy

Details of our machine learning strategy can be found in the Supplementary information C3. We used the open-source machine learning software NeuroMiner to train and validate two different classifiers (<https://github.molgen.mpg.de/pages/LMU-Neurodiagnostic-Applications/NeuroMiner.io/>) in the CCI sample: a full feature classifier trained with all 39 baseline variables (classifier 1) [Supplementary information T1], and a 28 variables non-psychological classifier trained without the psychological-behavioural baseline features (classifier 2) [Supplementary information T2]. The input variables of both classifiers underwent the same pre-processing pipeline before entering the models. All models

are trained using the linear kernel Support Vector Machine (SVM). The optimisation metric is Balanced Accuracy defined as $BAC = (\text{sensitivity} + \text{specificity}) \div 2$. In order to achieve optimal prediction performances with the least number of variables, we incorporated greedy forward search wrapper in the model training. The wrapper is a feature selection technique which iteratively adds 5% of the variables and trains a new model at each step. Next, the wrapper selects the best performing model and uses the model's input feature subset as the optimal feature set. We employed the randomly pooled repeated nested cross-validation (P-CV) strategy to train and validate the two classifiers and avoid overfitting. Our P-CV strategy consists of 10 permutations and 10 folds at the inner cross-validation cycle (CV1) and 10 permutations and 10 folds at the outer cross-validation cycle (CV2). The model's final performance is calculated from the mean performances of all 100 CV2 models. Classifier performances include sensitivity, specificity, balanced accuracy, area-under-the-curve, false positive rate, positive and negative likelihood ratios, prognostic summary index PSI (Linn & Grunau, 2006), and number-needed-to-diagnose (Larner, 2018). Each participant's final Out-Of-Training (OOT) ensemble prediction was generated by computing the outcome predictions across all 100 CV2 models using majority voting.

Additionally, variables selected by at least 50% of the models within each ensemble classifier were used to create brief versions of classifier 1 and 2. These classifiers were retrained with the data-driven condensed variable sets using the same settings described above excluding the wrapper feature selection. These two additional retrained classifiers are called the brief psychological pilot study model (brief psychological model, classifier 3) and the brief non-psychological pilot study model (brief non-psychological model, classifier 4).

Post hoc Predictive Pattern Extraction

Four different post hoc methods, as implemented in NeuroMiner, were used to identify and visualise the predictive patterns from the machine learning models ensemble [Supplementary information C4]. First, we computed mean feature weights across the ensemble classifier by averaging the normalised SVM models' weights directly extracted from the SVM models (Gaonkar et al., 2015). Second, we calculated Spearman correlation coefficients between each variable and the predicted scores to compare univariate and multivariate feature weights. Third, we calculated the pattern element stability termed as cross-validation ratio (CVR), by computing the mean and standard error of all SVM weight vectors concatenated across the entire nested cross-validation structure (Koutsouleris et al., 2021). Finally, we employed sign-based consistency to statistically test the predictive stability

of variables, including a correction for multiple testing using the False-Discovery-Rate (Gómez-Verdejo et al., 2019). The combination of the four predictive pattern visualisation methods can give us high certainty of the feature importance found by the machine-learning models.

Construct Validity Analysis Using ACC Sample

We tested the brief psychological model's construct validity in the 7554 participants drawn from the ACC dataset. To this end, we identified 7 matching variables which were assessed in both CCI and ACC datasets [Supplementary information T5]. The 3 remaining variables in the brief psychological model with no matching ACC questions were coded as missing values. The brief psychological model was applied to the ACC study participants to produce prediction labels and decision scores. We visualised the decision score distributions for each score level of the ordinal variables from the ACC study which were not available in the CCI dataset: 'depressiveness' (ACC-12), 'resilience' (ACC-13), 'stress recovery' (ACC-14) and 'return to normal' (ACC-15) [Supplementary information T5]. Following the construct validity theory (Strauss & Smith, 2009), these four variables were used as proxy measures to indirectly evaluate whether the model reflects the predictive pattern in the validation sample. We also measured the prevalence of distressed outcome prediction in each score level of the four variables. Finally, we used Kendall's tau-b (Tau-b) to evaluate the association between the variables of the ACC study and the decision scores.

Assessment of Psychological and Behavioural Trajectories

In the pilot study, linear mixed-effects models were used to evaluate the brief psychological model's stratification effect on the participants' self-reported psychological and behavioural measures spanning from the baseline assessment, over the 24-day time point to the 3-month final examination. To this end, for each participant and time point, we computed three domain measures reflecting general psychological burden, affective burden and behavioural adaptation by summing the single questionnaire items belonging to these domains of the CCI study [Supplementary information C1]. Then, we categorised study participants into low-risk, intermediate-risk, high-risk and ultra-high-risk outcome classes by defining the median, 80% and 90% percentiles of the OOT decision score distribution as cut-offs thresholds for the outcome categories. Domain-by-outcome category trajectories were visualised in Fig. 3. Then, domain measures were entered as dependents while examination time point, and outcome labels entered as within-subject and between-subject fixed factors in the mixed-effects analyses. Main effects of examination time point and outcome label as well

as their interaction were assessed for statistical significance at $\alpha=0.05$. Finally, if both main effects were significant, we conducted estimated marginal means analyses to determine significant differences between outcome categories. *P*-values were adjusted for false recovery rate (FDR-corrected) (Gómez-Verdejo et al., 2019).

Prediction Model Implementation

For providing individualised psychological support to help HCWs to mitigate their personal risk against COVID-related stress, we deployed our two brief classifiers (classifier 3, classifier 4) on our online NeuroMiner Model Library (<http://www.proniapredictors.eu>) (Supplementary information C5).

Results

The population and baseline characteristics of the pilot and validation study datasets are described in Table 1.

Model Prediction Performances and Predictive Pattern in the Pilot Dataset

In the CCI dataset, 56 of 220 (25.4%) HCWs had reported pandemic-related psychological distress at the time-point analysed. Classifier 1 predicted these poor outcomes with a BAC of 75% and increased the prognostic certainty by a PSI of 41%. The BAC for COVID-19 services was 73%; 70% for non-COVID-19 services; and 74% for hospital administration (Table 2). Classifier 2 achieved a BAC of 67% and a prognostic gain by a PSI of 25%. The BAC for COVID-19 services was 64%, for non-COVID-19 services 67%, and 49% for hospital administration.

The significant predictors for a poor course in classifier 1 (feature selection probability > 50%; significance threshold, $\alpha=0.05$, FDR-corrected) were development of obsessive behaviours; female sex; sleep disturbances; worrying about health; anxiety and somatic disturbances; feeling stressed; and worries about contact to SARS-CoV-2. The significant predictors for a good outcome are nightmare; contact with infected patients; and contact with infected persons (Fig. 1-A1). These

Table 1 Epidemiological information of 337 Healthcare Workers participating in the CCI pilot study, and epidemiological information and Anti-SARS-CoV-2 Antibody status of 7554 Healthcare Workers participating in the AAC validation study [cite PMID: 34379308]

	CCI pilot study	ACC validation study			
		Total	Anti-SARS-CoV-2 Antibody		
			Positive	Total	%
Number of participants	337	166	7554	2.2	(1.89–2.55)
Age group (Y)					
≤ 30	116	64	2170	2.95	(2.32–3.75)
31–40	106	39	1951	2	(1.47–2.72)
41–50	63	29	1430	2.03	(1.42–2.90)
51–60	38	23	1467	1.57	(1.05–2.34)
> 60	14	11	536	2.05	(1.15–3.64)
Gender					
Female	220	115	5431	2.12	(1.77–2.54)
Male	117	51	2118	2.41	(1.84–3.15)
3rd Gender	0	0	5	0	
Patient care occupations					
Nurse	167	68	2185	3.11	(2.46–3.93)
Physician	105	38	1345	2.83	(2.07–3.85)
Other	0	17	1199	1.42	(0.88–2.26)
Total	272	123	4729	2.6	(2.18–3.10)
Non-patient care occupations					
Administration/IT	27	15	822	1.82	(1.11–2.99)
Research	0	12	977	1.23	(0.70–2.14)
Transportation	17	1	28	3.57	(0.63–17.71)
Cleaning personnel	6	4	119	3.36	(1.32–8.33)
Other	15	11	879	1.25	(0.70–2.23)
Total	65	43	2825	1.52	(1.13–2.04)

% percent of participants with positive SARS-CoV-2 test result, 95% CI Binominal 95% confidence intervals (95% CI) were calculated using the Wilson score interval

Table 2 Overview of prediction performance measures of the full-feature, reduced-feature, and non-psychological models as measured in the validation sample of pilot study participants

Machine learning models	TP	TN	FP	FN	Sens [%]	Spec [%]	BAC [%]	AUC	FPR [%]	LR+	LR-	PSI [%]	NND
Full psychological model													
Whole cohort	41	126	38	15	73.2	76.8	75.0	0.80	23.2	3.16	0.35	41.3	2.00
COVID-19 services	10	61	11	6	62.5	84.7	73.6	0.82	15.3	4.09	0.44	38.7	2.12
Non-COVID-19 services/Central emergency	15	53	17	8	65.2	75.7	70.5	0.72	24.3	2.69	0.46	33.8	2.44
Hospital administration	16	12	10	1	94.1	54.6	74.3	0.81	45.5	2.07	0.11	53.9	2.06
Brief psychological model													
Whole cohort	41	128	36	15	73.2	78.0	75.6	0.80	22.0	3.34	0.34	42.8	1.95
COVID-19 services	10	61	11	6	62.5	84.7	73.6	0.82	15.3	4.09	0.44	38.7	2.12
Non-COVID-19 services/Central emergency	15	55	15	8	65.2	78.6	71.9	0.74	21.4	3.04	0.44	37.3	2.28
Hospital administration	16	12	10	1	94.1	54.6	74.3	0.81	45.5	2.07	0.11	53.9	2.06
Full non-psychological model													
Whole cohort	44	90	74	12	78.6	54.9	66.7	0.65	45.1	1.74	0.39	25.5	2.99
COVID-19 services	8	56	16	8	50.0	77.8	63.9	0.58	22.2	2.25	0.64	20.8	3.60
Non-COVID-19 services/Central emergency	20	33	37	3	87.0	47.1	67.1	0.61	52.9	1.65	0.28	26.8	2.93
Hospital administration	16	1	21	1	94.1	4.6	49.3	0.58	95.5	0.98	1.29	–	–
Brief non-psychological model													
Whole cohort	44	90	74	12	78.6	54.9	66.7	0.65	45.1	1.74	0.39	25.5	2.99
COVID-19 services	8	56	16	8	50.0	77.8	63.9	0.58	22.2	2.25	0.64	20.8	3.60
Non-COVID-19 services/Central emergency	20	33	37	3	87.0	47.1	67.1	0.61	52.9	1.65	0.28	26.8	2.93
Hospital administration	16	1	21	1	94.1	4.6	49.3	0.58	95.5	0.98	1.29	–	–

COVID-19 coronavirus disease 2019, TP number of true positives, TN number of true negatives, FP number of false positives, FN number of false negatives, Sens Sensitivity, Spec Specificity, BAC Balanced Accuracy, AUC Area-under-the Curve, LR+ Positive Likelihood Ratio, LR- Negative Likelihood Ratio, PSI Prognostic Summary Index, NND Number Needed to Diagnose

10 features were used to retrain the brief psychological model described above [Supplementary information T3]. The brief psychological model achieved a BAC of 75% for the whole cohort and increased the prognostic certainty by a PSI of 42%. The BAC was 73% for COVID-19 services, 71% for non-COVID-19 services, and 74% for hospital administration. Significant poor outcome predictors in the classifier 2 were reports catarrh; headache; age; sex and working in hospital administration. While contact with infected patients and contact with infected persons were predictive of good-outcome (Fig. 1-B1). These 7 variables were used to retrain the brief non-psychological model [Supplementary information T4]. The brief non-psychological model's prediction performances are identical to classifier 2 (Table 2).

Based on the results from the 4 post hoc predictive pattern analyses, the 5 most important poor-outcome predictors in the brief psychological model were worries about health, feeling stressed, nausea, while contact with infected personnel and contact with infected patients were the only predictors of a good course (Supplementary information S4). Supplementary information S5 shows that the only significant predictors from the brief non-psychological model were contact with infected personnel and contact with infected patients. All other predictors were not significant.

In order to provide stratified intervention recommendations to help HCWs, we have defined four different risk categories in the implemented prediction app. The risk categories are: No psychological stress (mean prediction score ≥ 0); Mild psychological stress (mean prediction score < 0 and ≥ -0.419 , 75–100 percentile); Moderate psychological stress (mean prediction score < -0.419 and ≥ -1 , 35–75 percentile) and Severe psychological stress (mean prediction score < -1 , 0–35 percentile). The percentiles of the risk categories are calculated based on the mean prediction score distribution from our training data (Supplementary information C5).

Construct Validity Analysis of the Brief Validation Study Model in the Validation Study Dataset

We examined the construct validity of the brief psychological model in the validation study dataset. The percentage of participants predicted as distressed as well as mean prediction scores were computed for each score group on the 'depressiveness', 'resilience', 'stress recovery' and 'return to normal' items from the validation study questionnaire (Fig. 2) (Supplementary information T5). Figure 2(A–D) showed that as the distress severities in the 4 items were increasing, the

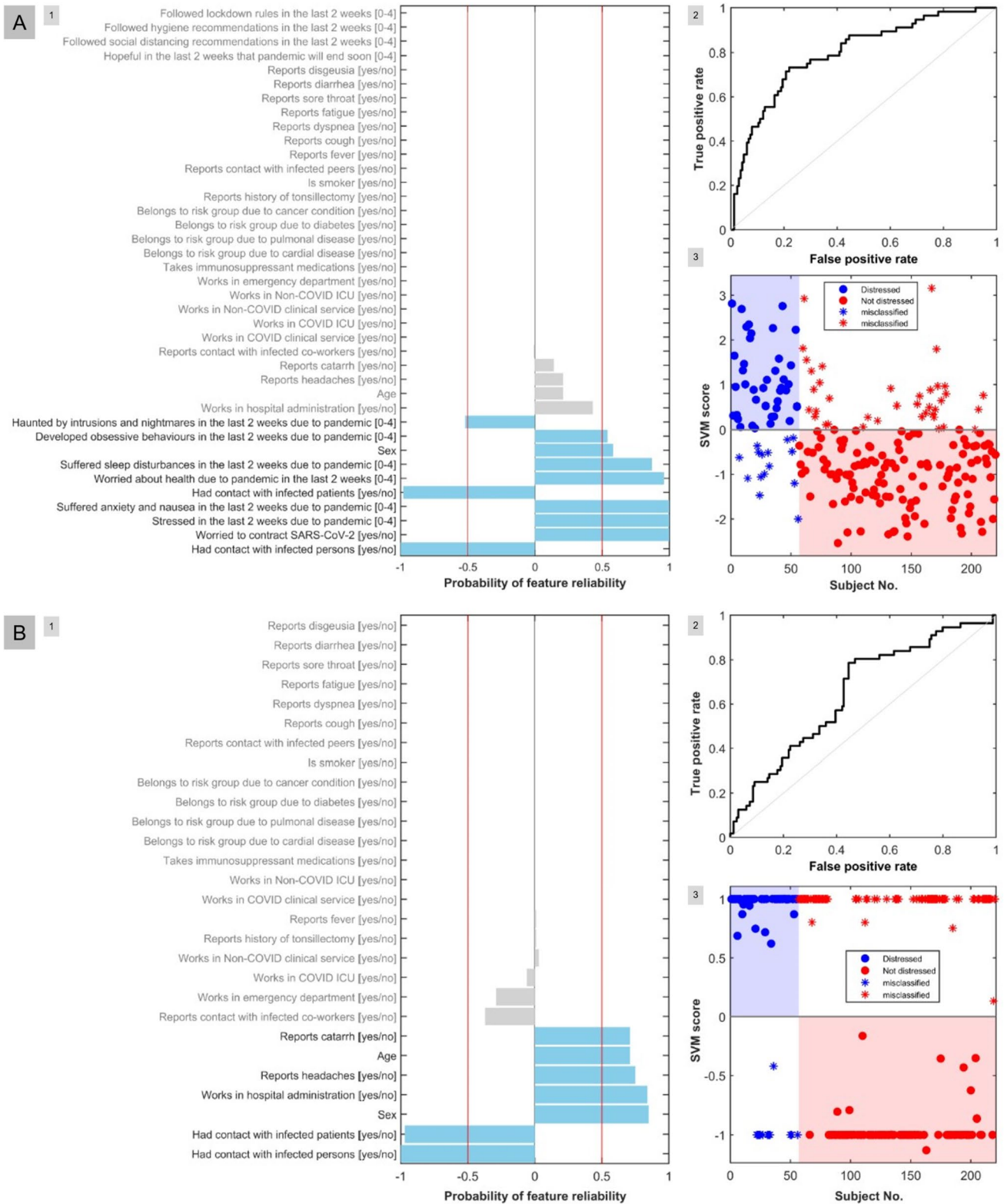


Fig. 1 Receiver Operating Characteristic (ROC) analysis and classification plots of the brief prognostic models developed using the entire variable pool (**A**) and the pool after removal of the psychological variables (**B**). The feature reliability profile of the prognostic model was trained on all 39 features (A1). Positive values indicated higher feature values in the distressed vs. the non-distressed outcome persons.

Subplot A2 shows the receiver-operator-curve analysis of the model and A3 shows classification plot with correctly and wrongly classified individual participants. Subplot B shows the respective analysis steps for the model trained on 28 features after removing the 11 psychological variables from the feature pool

percentage of predicted distressed participants were also rising significantly ($p < 0.001$). Figure 2(E–H) showed a clear trend that the decision scores were increasing significantly ($p < 0.001$) with higher distress severity in all items of the validation study. The Kendall's Tau-b correlation tests also showed that the prediction scores are significantly correlated with the items. Depressiveness has the highest tau-b correlation of 0.34 ($p < 0.001$), while resilience has the lowest correlation of 0.26 ($p < 0.001$).

Trajectory Analysis of General, Behavioural, and Affective Burden in Stratified Predicted Outcome Groups in the Pilot Study Dataset

Table 3 shows the results of longitudinal mixed-effects model analysis of 'general psychological burden', 'behavioural adaptation' and 'specific affective' burden in participants of the pilot study from baseline to 105 days. The general burden model achieved an Akaike Information Criterion (AIC) of 1,221. The main effects of visits ($F = 26.2$, $p < 0.001$) and stratified outcome groups ($F = 139.8$, $p < 0.001$) significantly affects the general burden. In the affective burden model (AIC = 1232), the main effects of visits ($F = 15.6$, $p < 0.001$) and stratified outcome groups ($F = 133.1$, $p < 0.001$) as well as the interaction between visits and outcome groups ($F = 3.4$, $p = 0.003$) have significant effects. The interaction between the two effects is also significant ($p = 0.003$). In the behavioural adaptation model (AIC = 1282), the main effects of visits ($F = 36.7$, $p < 0.001$) and stratified outcome group ($F = 4$, $p < 0.008$) have significant effects. Figure 3 shows the trajectory analysis of pilot study participants in the four stratified outcome groups. The trajectory of all psychological variables showed decreased compliant values between baseline and 105 days (Fig. 3A–C). The outcome groups remained significantly separable in both general burden and specific affective burden throughout the entire follow-up, but not significantly separable in terms of behavioural adaptation (Fig. 3D–F).

Discussion

In this study, we provided evidence that machine learning can be used to predict the individual risk for pandemic-related psychological distress in HCWs. Our aim was to develop a simple, clinically scalable decision-support tool to inform individual risk-adapted psychological support following a stepwise prevention model to avoid the development of absenteeism and mental disorders in vulnerable individuals (Bakkeli, 2022; Holmlund et al., 2022). We found that our tool identified these individuals, as defined by the upper quartile of the distressed outcome distribution (Mo et al., 2020), with a BAC of 75% in keeping with a recent study

focusing on the prediction of resilience to pandemic-related psychological distress (Lieslehto et al., 2022). We also observed that the tool increased prognostic certainty by 42% and performed equally well across different categories of HCWs, ranging from front-line healthcare professionals to hospital administration staff. Recent data showed that the seamless interaction of different types of HCWs is required to maintain the functionality of hospitals under the pressure of a pandemic situation (Bakkeli, 2022). In this regard, our findings support the potential utility and scalability of our tool to safeguard a high quality of care in a pandemic situation through the prevention of adverse mental health-related outcomes. Specifically, based on repeatedly quantifying the HCWs' risk, our tool could help fast-track HCWs with high levels of predicted distress outcomes to online face-to-face interventions (e.g., preventive CBT), while recommending app-supported protective measures such as progressive muscle relaxation and mindfulness exercises in cases with milder levels of predicted future distress. This stepwise approach allows simple measures to be applied in mild cases without the need for human intervention, which in turn help save scarce healthcare resources and focus them on infected patients. Further studies should examine whether recommending different measures for the different risk categories was clinically effective.

Our models used epidemiological and psychological parameters in the longitudinal pilot study assessed after 12 weeks to predict pandemic-related psychological distress and to identify HCWs at risk, and the models showed higher accuracy than pre-test outcome probabilities. The prediction performances from our full-feature, reduced-feature, and non-psychological models showed that including psychological parameters increased the prediction accuracy. Our brief models showed that reducing the number of variables increased the prediction accuracy, indicating that distress can be accurately predicted in HCW with only a few variables, minimizing the data acquisition cost. Through the post hoc predictive pattern extraction, we identified several variables previously described as being relevant for risk of pandemic-related psychological distress in HCWs, including subjective stress, concern for health, and anxiety (Rossi et al., 2020; Xiao et al., 2020). Our models also identified new predictors including 'developing obsessive behaviours due to the pandemic'. Contrary to our expectations, contacts with infected individuals and patients and working on COVID-19 wards were predictors of good outcomes. These findings may indicate that HCWs working on COVID-19 wards were well informed, able to cope well with the situation and, thus, felt less stressed.

Through the construct validity examinations conducted in the validation study dataset, we found an increase of the percentage of distressed predictions with higher scores on the validation study items 'depressiveness', 'resilient', 'stress recovery' and 'return to normal'. A clear trend of significant

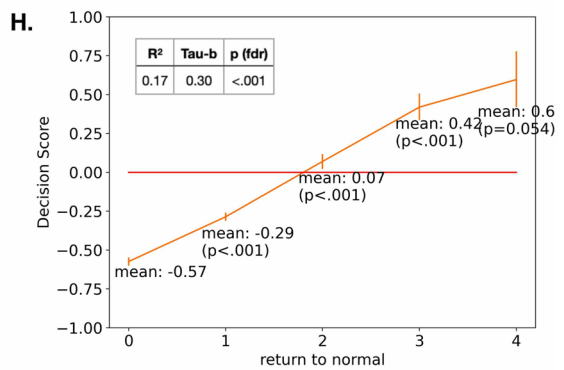
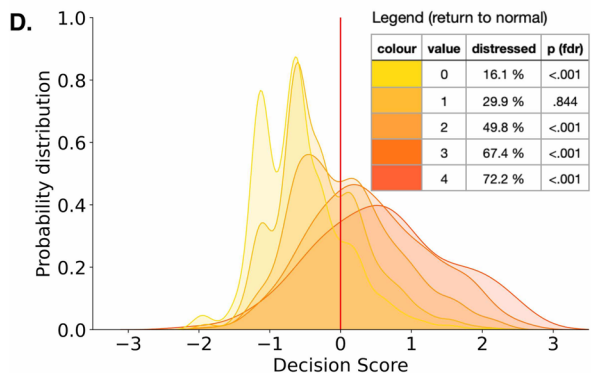
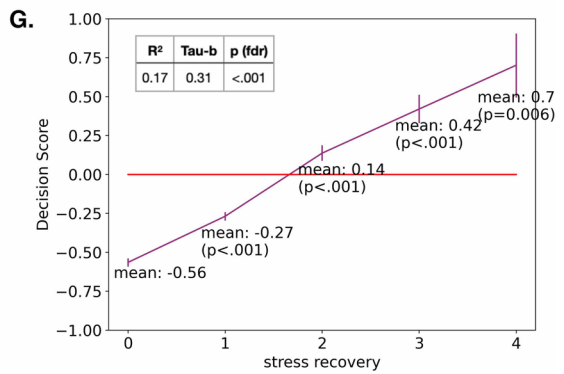
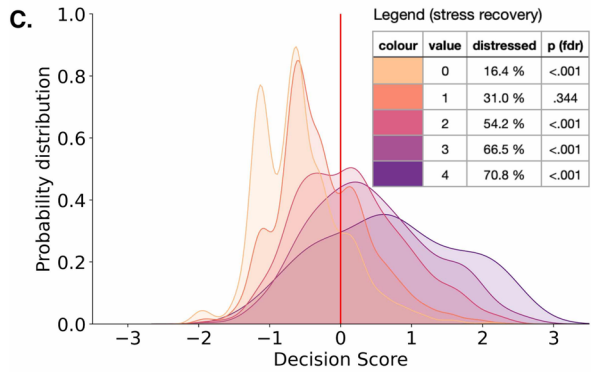
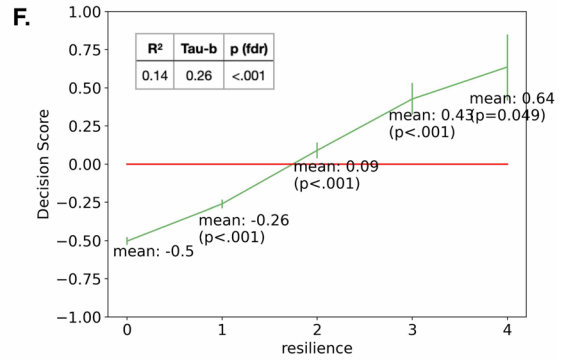
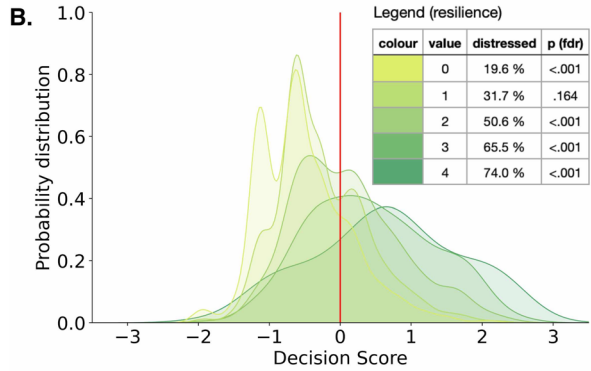
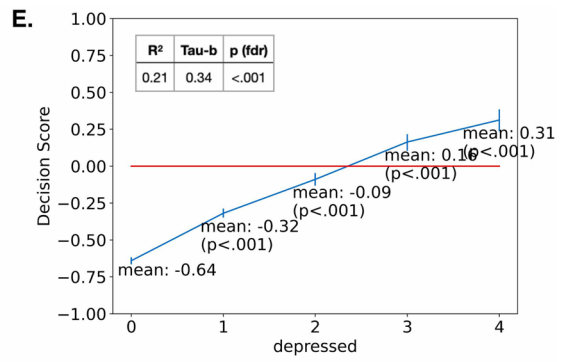
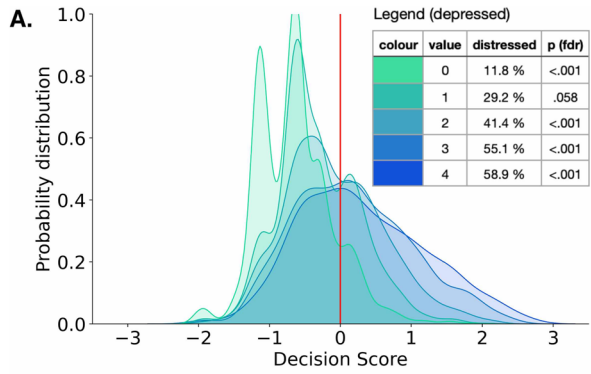


Fig. 2 The results of the brief psychological model's construct validity analysis in the validation study dataset using 4 items from the validation study questionnaire. The depressiveness item is 'Because of the COVID-19 pandemic I often feel sad and/or depressed' (0='not at all'; 4='very often'). The resilience item is 'In general, I have problems dealing with stressful situations' (0='totally disagree'; 4='totally agree'). The stress recovery item is 'In general, it takes a long time for me to recover from stressful situations' ('0'='totally disagree'; '4'='totally agree'). The return to normal item is 'I was not able to return back to normal from the stressful situations' ('0'='totally disagree'; '4'='totally agree'). In panels (A)–(D), the percentage of participants labelled as distressed was computed for each score group on the depressiveness (A), resilience (B) stress recovery (C), and return to normal (D) items of the validation study questionnaire. The legend in each plot shows the exact percentage of each score group as well as the FDR-corrected *p* values when comparing each score group to the whole dataset. The red line indicates the decision boundary. Panels (E–H) depict the mean (data point) and 95% confidence interval (error bars) of the brief psychological model's prediction decision scores from each score group on the depressiveness (E), resilience (F) stress recovery (G), and return to normal (H) items of the validation study questionnaire. The annotated *p*-values below each data point are FDR-corrected *p*-value when comparing the current score group with the previous score group (e.g. group 2 vs group 1, group 1 vs group 0), therefore *p*-value for group 0 is not available. The exact *p*-values can be found in the supplementary Fig. S2. The table at the top left corner in (E, F, G, H) displaces the R-squared explained variance when fitting a linear regression model predicting item scores using the prediction scores. The Tau-b value is the Kendall's tau sub-b correlation between the item scores and the prediction scores. The *p*-value reported are FDR-corrected hypothesis test scores of the Tau-b correlation values

positive association between all validation study items and the prediction scores is also observed. These results indicated that the classifier can reflect the constructs of psychological distress in an independent external dataset, thus improving the reliability of the model in generalising to future unseen sites and increasing the applicability of the model.

Our trajectory analysis in the pilot study dataset showed that the model captured the construct of depressiveness well

and predicted different dynamics of complaint trajectories. All psychological variables showed highly significant main and interaction effects between predicted outcomes and trajectories. The trajectory analyses showed that the model could predict poor trajectories, in particular people who had already shown distress or higher stress levels with varying trajectories and a tendency to worsen or remain worse, whereas HCWs who did not show distress remained stable over time. On closer inspection, we found that HCWs with a poor course showed a higher stress load at the beginning of the observation period. This finding could be explained by the fact that we only started collecting data in March 2020, but the pandemic started affecting the work of HCWs in end-February 2020. However, a similar relation of higher initial distress with a burdensome course of mental health during the first wave of the pandemic has been reported for a population-based sample (Ahrens, Neumann, Kollmann, Brokelmann et al., 2021; Ahrens, Neumann, Kollmann, Plichta et al., 2021), which had the opportunity to rely on pre-pandemic data as the initial reference point. This may suggest that HCWs with a poor trajectory may have already had a mental disorder before, independent from the pandemic. For gaining better knowledge in this highly endangered group, further prospective studies with standardised and validated instruments are required. In the future this may enable us to offer optimized therapeutic interventions.

Finally, we managed to bring the experimental machine-learning model into a practical tool for clinical translations and individual interventions. The model implementation is realised through our self-developed platform NeuroMiner model library described in the "Methods" section. Figure 4(A) shows an example of the type of information the HCWs can receive from the pilot study model implementation. Here, we used the brief psychological model to stratify the HCWs on the basis of their

Table 3 Longitudinal mixed-effects model analysis of general psychological burden, specific affective burden and behavioural adaptation in 260 pilot study participants followed for 105 days between March and July 2020

	AIC	df1	df2	F-value	P-value
General burden	1221.3				
Main effect of Visit		2	440	26.2	<.001*
Main effect of stratified outcome group		3	635.7	139.8	<.001*
Interaction: Visit x outcome group		6	440.2	1.4	0.219
Affective burden	1231.6				
Main effect of visit		2	463	15.6	<.001*
Main effect of stratified outcome group		3	640.8	133.1	<.001*
Interaction: Visit x outcome group		6	463.9	3.4	.003*
Behavioural adaptation	1281.7				
Main effect of visit		2	365.6	36.7	<.001*
Main effect of stratified outcome group		3	518.5	4	.008*
Interaction: Visit x outcome group		6	365.4	0.9	0.532

AIC Akaike Information Criterion, *df1* degrees of freedom numerator, *df2* degrees of freedom denominator, *P*-value Bonferroni-corrected *P*-values from hypothesis tests

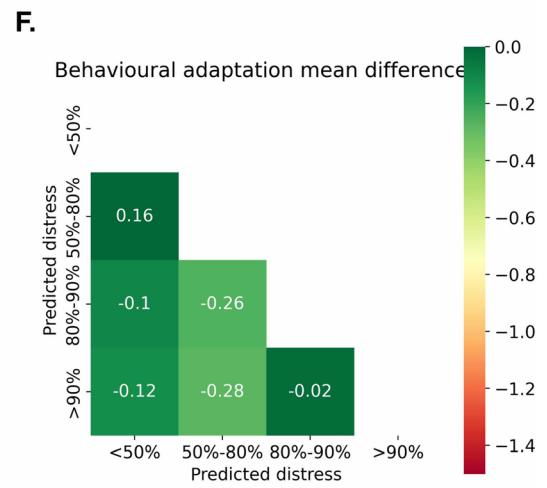
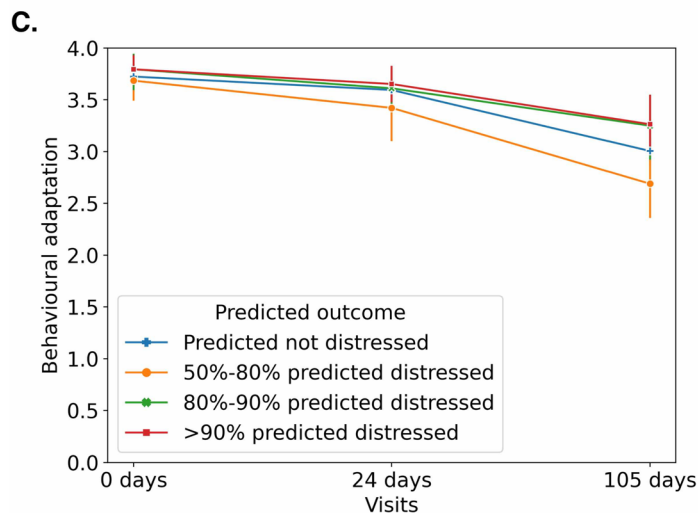
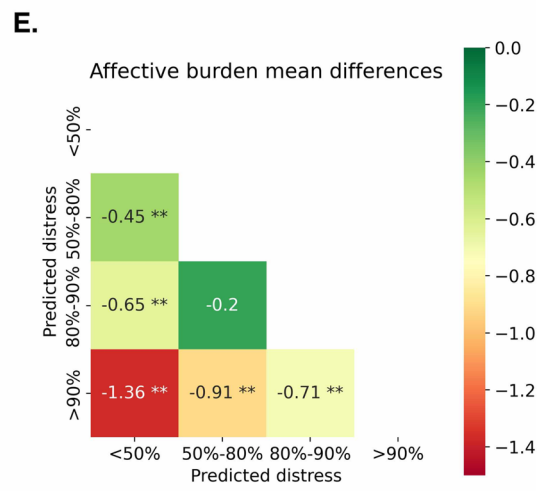
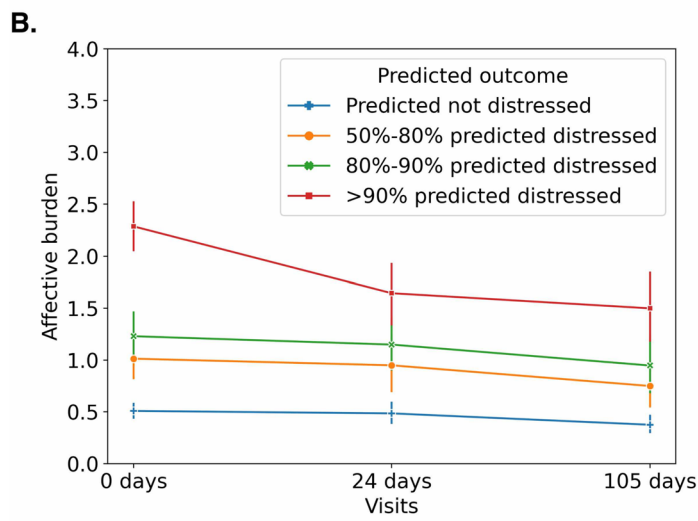
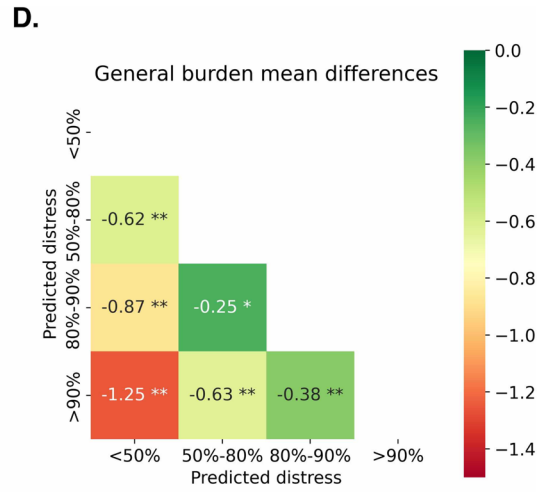
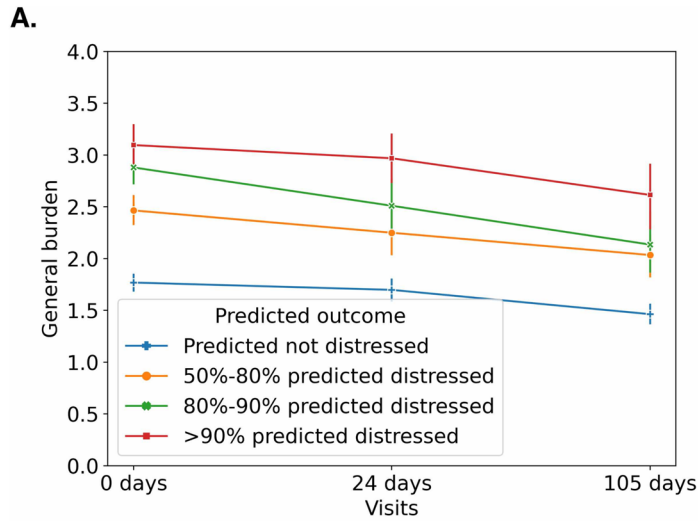


Fig. 3 Trajectory analysis of general, behavioural, and affective burden in stratified, predicted outcome groups in the pilot study dataset. Group assignments were based on the full psychological model's decision scores in 260 CCI participants and differentiated between participants scoring below the median, between the median and the 80% percentile, between the 80 and 90% percentile, and above the 90% percentile of the decision score distribution. The psychological variables were averaged into 3 summary scores, measuring general burden ('worried about health', 'worried about health within the last two weeks', 'hopeful pandemic ends soon' [coding reversed], 'stressed due to the pandemic'), behavioural adaptation ('following social distancing rules', 'following hygiene recommendations', 'following lockdown rules'), and affective burden ('sleep disturbances', 'anxiety and somatic disturbances', 'intrusions and nightmares', 'obsessive disturbances'). Panels (A)–(C) display the results of the trajectory analysis of outcome groups defined by the brief psychological model for the 3-summary scores. Error bars indicate the 95% confidence intervals around the subgroups' mean scores at each visit. Panels (D)–(F) are showing the pairwise mean difference when comparing all outcome groups' psychological variables against each other. The value in each cell is calculated by subtracting the mean psychological variable value from the more distressed group with the less distressed group. The * symbol denotes FDR-corrected $p < .05$ and ** denotes $p < .001$ when comparing the 2 outcome groups using independent *t*-test. The exact *p*-values can be found in the supplementary Fig. S3

risk for developing pandemic-related distress. In the next step, we defined tailored therapeutic procedures suggested for each risk group. By assessing the severity of symptoms in HCWs, we were able to classify them into different risk categories, which enabled the suggestion of individually adapted interventions, such as smartphone-based mindfulness exercises and progressive muscle relaxation (PML) for risk category 1, tele-health COVID-19 psychiatric consultation for risk category 2, and outpatient preventive CBT interventions for risk category 3. From our results and literature findings (Mo et al., 2020), we defined the decision threshold value as below 2, i.e., employees with a score of 2 or more should be offered stress-relieving interventions. Additionally, the implementation also tracks the trend of the HCWs stress level as well as having an understanding of their peers who are facing similar situations. The model implementation achieved the research to application cycle and completed our study aim. Our tool

is constantly being developed within the framework of the University Medicine Network. In the future, it should be particularly applicable in crisis situations.

Limitations

Our study has the following limitations: Firstly, at the time of our study design and launch (April 2020), no standardised and validated instruments were available to assess COVID-related distress in HCWs [Supplementary information T9]. As a result, we developed self-administered questionnaires in both CCI and ACC studies. Although our questionnaire tool was based on many years of clinical and research experience, it could not be sufficiently validated due to the dynamics at the beginning of the pandemic. However, the questionnaire is currently under validation in different settings. Furthermore, because we are conducting the study among actively working employees, we are limited to assess only symptoms and could not provide diagnoses. We also aimed to explore different COVID specific risk factors which have not been investigated in the existing literature. Finally, few features were included in the CCI study. The CCI questionnaire did not include features (e.g., female, family concerns, poor sleep quality, decreased perceived health status); therefore, these variables were not included in the machine learning algorithm later in the manuscript.

Secondly, due to staff change and the limited staff availability during our study period, 40% of participants did not manage to complete the entire CCI study.

Thirdly, the variables that overlapped in the two studies were only available from HCWs at the H-LMU, and we were unable to validate the full feature classifiers in a multicentric approach (Steyerberg & Harrell, 2016). Lastly, our study only focused on creating a predictive tool for stratifying and identifying HCWs who might be at risk of suffering mental stress due to the pandemic, therefore interventions were not validated, and not carried out in a structural manner in the duration of our study.

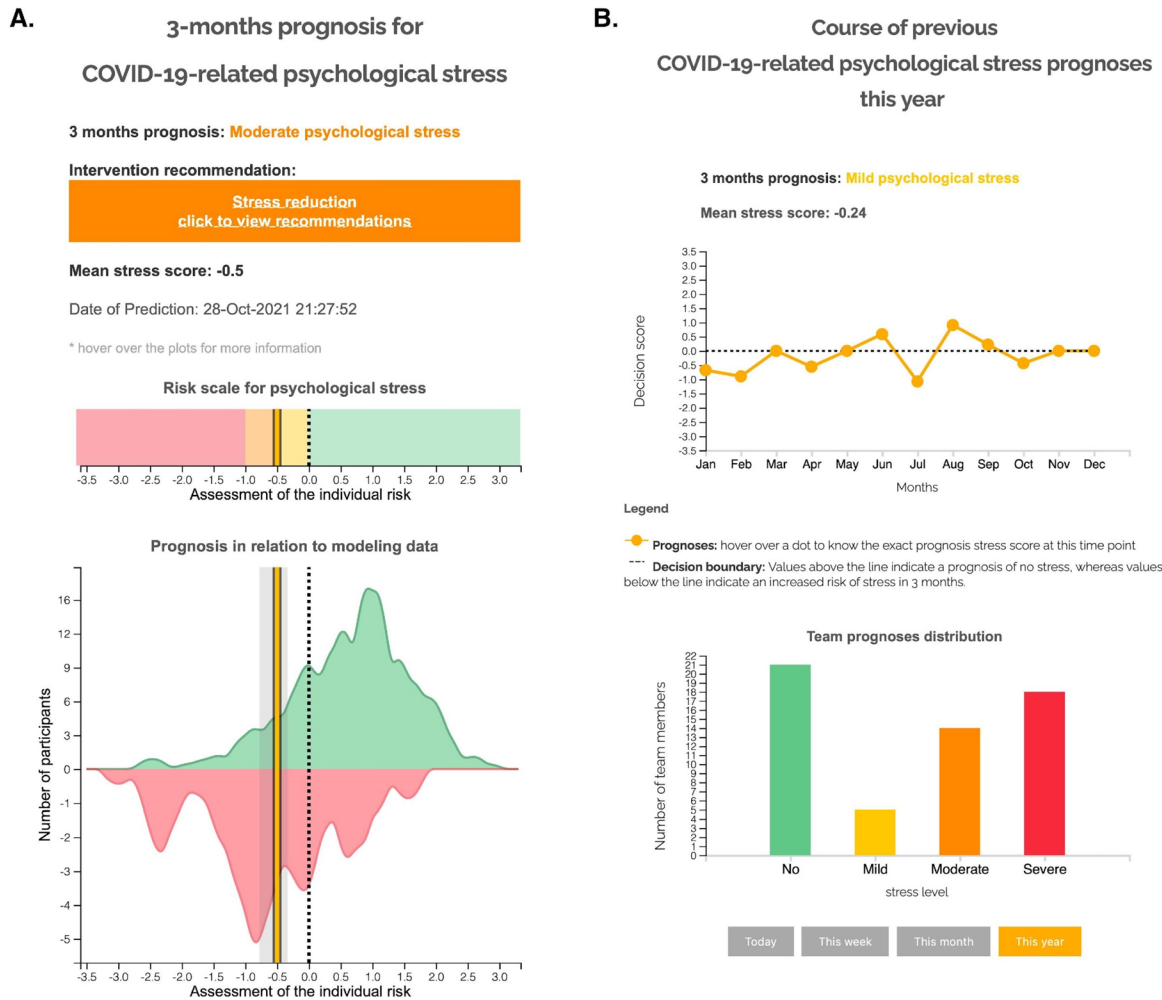


Fig. 4 Exemplary results display of the implemented pilot study model in the NeuroMiner model library. **A** Prognosis results display. Top: 3-month prognosis and actionable suggestions to reduce stress level, middle: the exact prognosis score and risk categorisation (Mo et al., 2020), bottom: prognosis score in relation to the training data population. Apart from providing immediate individual prognosis results, the model library also tracks the prognoses of the user over time, so the user can be informed about the trend of their COVID-

19-related mental stress. **B** Longitudinal and team results display. Top: trend of the individual's previous prognoses, bottom: overall distributions of stress risk categories of the team where the user belongs. This adds an additional layer of transparency to the users so that they can have an anonymous understanding of the overall stress level of the team, which can help the colleagues and team members stay informed of the mental conditions of each other

Conclusion

In this study, we developed risk assessment tools for predicting pandemic-related distress in HCWs using machine learning. To our knowledge, this is the first study to demonstrate that augmenting human prognostic capabilities with machine learning pattern recognition improves prognostic accuracy to a degree that likely justifies clinical implementation of cybernetic decision-support tools. They can improve the stratification of risk in HCWs to ensure that they receive adequate support. We are planning a large-scale clinical and external validation study by developing and combining a health monitoring app-solution

with stratified employee assistance program interventions. In addition, a context adaptivity study (e.g., in ICUs) is also planned. These tools can help optimise resource allocation, prevent the development of mental disorders and predict human resource capacities in German hospitals.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s41347-023-00380-9>.

Funding Open Access funding enabled and organized by Projekt DEAL. This research project was supported by LMUexcellent, funded by the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government as well as the "Stiftungen zugunsten der Medizinischen Fakultät—Cluster: 2". The study was funded in part by the BMBF initiative "NaFoUniMedCovid19" (01KX2021), subproject B-FAST.

Declarations

Ethics Approval The Care Corona Immune (CCI) study was approved by the ethics committee of the Medical Faculty of the LMU (approval number: 20–247). The All Corona Care (ACC) study was approved by the ethics committee of the Medical Faculty of the LMU (approval number: 20–245),

Informed consent All healthcare workers (HCWs) aged over 18 years at the hospital-LMU were invited to participate in the Care Corona Immune (CCI) study as well as the All Corona Care (ACC) study, participation was voluntary, and all participants gave written informed consent to participate.

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.


References

- Arpaci, I., Karataş, K., & Baloğlu, M. (2020). The development and initial tests for the psychometric properties of the COVID-19 Phobia Scale (C19P-S). *Personality and Individual Differences, 164*, 110108. <https://doi.org/10.1016/j.paid.2020.110108>
- Ahrens, K. F., Neumann, R. J., Kollmann, B., Brokelmann, J., von Werthern, N. M., Malyshau, A., Weichert, D., Lutz, B., Fiebach, C. J., Wessa, M., Kalisch, R., Plichta, M. M., Lieb, K., Tüscher, O., & Reif, A. (2021). Impact of COVID-19 lockdown on mental health in Germany: Longitudinal observation of different mental health trajectories and protective factors. *Translational Psychiatry, 11*(1), 392. <https://doi.org/10.1038/s41398-021-01508-2>
- Ahrens, K. F., Neumann, R. J., Kollmann, B., Plichta, M. M., Lieb, K., Tüscher, O., & Reif, A. (2021). Differential impact of COVID-related lockdown on mental health in Germany. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA), 20*(1), 140–141. <https://doi.org/10.1002/wps.20830>
- Bakkeli, N. Z. (2022). Predicting psychological distress during the COVID-19 pandemic: Do socioeconomic factors matter? *Social Science Computer Review, 41*(4). <https://doi.org/10.1177/08944393211069622>
- Bartram, D., & Hambleton, R. K. (2016). The ITC guidelines: International standards and guidelines relating to tests and testing. In F. T. L. Leong, D. Bartram, F. M. Cheung, K. F. Geisinger, & D. Iliescu (Eds.), *The ITC international handbook of testing and assessment* (pp. 35–46). Oxford University Press. <https://doi.org/10.1093/med/psych/9780199356942.003.0004>
- Burkhardt, G., Adorjan, K., Kambeitz, J., Kambeitz-Ilankovic, L., Falkai, P., Eyer, F., Koller, G., Pogarell, O., Koutsouleris, N., & Dwyer, D. B. (2020). A machine learning approach to risk assessment for alcohol withdrawal syndrome. *European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology, 35*, 61–70. <https://doi.org/10.1016/j.euroneuro.2020.03.016>
- Butler, C. R., Wong, S. P. Y., Wightman, A. G., & O'Hare, A. M. (2020). US clinicians' experiences and perspectives on resource limitation and patient care during the COVID-19 pandemic. *JAMA Network Open, 3*(11), e2027315. <https://doi.org/10.1001/jamanetworkopen.2020.27315>
- Cai, H., Tu, B., Ma, J., Chen, L., Fu, L., Jiang, Y., & Zhuang, Q. (2020). Psychological impact and coping strategies of frontline medical staff in Hunan between January and March 2020 during the outbreak of coronavirus disease 2019 (COVID-19) in Hubei, China. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research, 26*, e924171. <https://doi.org/10.12659/MSM.924171>
- Chand, G. B., Dwyer, D. B., Erus, G., Sotiras, A., Varol, E., Srinivasan, D., Doshi, J., Pomponio, R., Pignoni, A., Dazzan, P., Kahn, R. S., Schnack, H. G., Zanetti, M. V., Meisenzahl, E., Busatto, G. F., Crespo-Facorro, B., Pantelis, C., Wood, S. J., Zhuo, C., ... & Davatzikos, C. (2020). Two distinct neuroanatomical subtypes of schizophrenia revealed using machine learning. *Brain: A Journal of Neurology, 143*(3), 1027–1038. <https://doi.org/10.1093/brain/awaa025>
- Chekroud, A. M., Zotti, R. J., Shehzad, Z., Gueorguieva, R., Johnson, M. K., Trivedi, M. H., Cannon, T. D., Krystal, J. H., & Corlett, P. R. (2016). Cross-trial prediction of treatment outcome in depression: A machine learning approach. *The Lancet. Psychiatry, 3*(3), 243–250. [https://doi.org/10.1016/S2215-0366\(15\)00471-X](https://doi.org/10.1016/S2215-0366(15)00471-X)
- Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *BMJ (clinical Research Ed.), 350*, g7594. <https://doi.org/10.1136/bmj.g7594>
- COVID-19 Map. (n.d.). Johns Hopkins Coronavirus Resource Center. Retrieved November 30, 2022, from <https://coronavirus.jhu.edu/map.html>
- Firew, T., Sano, E. D., Lee, J. W., Flores, S., Lang, K., Salman, K., Greene, M. C., & Chang, B. P. (2020). Protecting the front line: A cross-sectional survey analysis of the occupational factors contributing to healthcare workers' infection and psychological distress during the COVID-19 pandemic in the USA. *British Medical Journal Open, 10*(10), e042752. <https://doi.org/10.1136/bmjopen-2020-042752>
- Gaonkar, B., Shinohara, T., & R., Davatzikos, C., & Alzheimers disease neuroimaging initiative. (2015). Interpreting support vector machine models for multivariate group wise analysis in neuroimaging. *Medical Image Analysis, 24*(1), 190–204. <https://doi.org/10.1016/j.media.2015.06.008>
- Gilan, D., Röhke, N., Blessin, M., Kunzler, A., Stoffers-Winterling, J., Müssig, M., Yuen, K. S. L., Tüscher, O., Thurl, J., Kreuter, F., Sprengholz, P., Betsch, C., Stieglitz, R. D., & Lieb, K. (2020). Psychomorbidity, resilience, and exacerbating and protective factors during the SARS-CoV-2 pandemic. *Deutsches Arzteblatt International, 117*(38), 625–630. <https://doi.org/10.3238/arztebl.2020.0625>
- Gómez-Verdejo, V., Parrado-Hernández, E., Tohka, J., & Alzheimer's disease neuroimaging initiative. (2019). Sign-consistency based variable importance for machine learning in brain imaging. *Neuroinformatics, 17*(4), 593–609. <https://doi.org/10.1007/s12021-019-9415-3>
- Holmlund, L., Hellman, T., Engblom, M., Kwak, L., Sandman, L., Törnkvist, L., & Björk Brämberg, E. (2022). Coordination of return-to-work for employees on sick leave due to common mental disorders: Facilitators and barriers. *Disability and Rehabilitation, 44*(13), 3113–3121. <https://doi.org/10.1080/09638288.2020.1855263>
- Huang, J. Z., Han, M. F., Luo, T. D., Ren, A. K., & Zhou, X. P. (2020). [Mental health survey of medical staff in a tertiary infectious disease hospital for COVID-19]. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi = Zhonghua Laodong Weisheng*

- Zhiyebing Zazhi = *Chinese Journal of Industrial Hygiene and Occupational Diseases*, 38(3), 192–195. <https://doi.org/10.3760/cma.j.cn121094-20200219-00063>
- Kang, L., Li, Y., Hu, S., Chen, M., Yang, C., Yang, B. X., Wang, Y., Hu, J., Lai, J., Ma, X., Chen, J., Guan, L., Wang, G., Ma, H., & Liu, Z. (2020). The mental health of medical workers in Wuhan, China dealing with the 2019 novel coronavirus. *The Lancet. Psychiatry*, 7(3), e14. [https://doi.org/10.1016/S2215-0366\(20\)30047-X](https://doi.org/10.1016/S2215-0366(20)30047-X)
- Koutsouleris, N., Dwyer, D. B., Degenhardt, F., Maj, C., Urquijo-Castro, M. F., Sanfelici, R., Popovic, D., Oetzuerk, O., Haas, S. S., Weiske, J., Ruef, A., Kambeitz-Ilanovic, L., Antonucci, L. A., Neufang, S., Schmidt-Kraepelin, C., Ruhrmann, S., Penzel, N., Kambeitz, J., Haidl, T. K., ... & PRONIA Consortium. (2021). Multimodal machine learning workflows for prediction of psychosis in patients with clinical high-risk syndromes and recent-onset depression. *JAMA Psychiatry*, 78(2), 195–209. <https://doi.org/10.1001/jamapsychiatry.2020.3604>
- Lai, J., Ma, S., Wang, Y., Cai, Z., Hu, J., Wei, N., Wu, J., Du, H., Chen, T., Li, R., Tan, H., Kang, L., Yao, L., Huang, M., Wang, H., Wang, G., Liu, Z., & Hu, S. (2020). Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. *JAMA Network Open*, 3(3), e203976. <https://doi.org/10.1001/jamanetworkopen.2020.3976>
- Larner, A. J. (2018). Number needed to diagnose, predict, or misdiagnose: Useful metrics for non-canonical signs of cognitive status? *Dementia and Geriatric Cognitive Disorders Extra*, 8(3), 321–327. <https://doi.org/10.1159/000492783>
- Lee, S. A. (2020). Coronavirus Anxiety Scale: A brief mental health screener for COVID-19 related anxiety. *Death Studies*, 44(7), 393–401. <https://doi.org/10.1080/07481187.2020.1748481>
- Lieslehto, J., Rantanen, N., Oksanen, L.-M. A. H., Oksanen, S. A., Kivimäki, A., Paju, S., Pietiäinen, M., Lahdentausta, L., Pussinen, P., Anttila, V.-J., Lehtonen, L., Lallukka, T., Geneid, A., & Sanmark, E. (2022). A machine learning approach to predict resilience and sickness absence in the healthcare workforce during the COVID-19 pandemic. *Scientific Reports*, 12(1), 8055. <https://doi.org/10.1038/s41598-022-12107-6>
- Linn, S., & Grunau, P. D. (2006). New patient-oriented summary measure of net total gain in certainty for dichotomous diagnostic tests. *Epidemiologic Perspectives & Innovations: EP+I*, 3, 11. <https://doi.org/10.1186/1742-5573-3-11>
- Lu, W., Wang, H., Lin, Y., & Li, L. (2020). Psychological status of medical workforce during the COVID-19 pandemic: A cross-sectional study. *Psychiatry Research*, 288, 112936. <https://doi.org/10.1016/j.psychres.2020.112936>
- Mo, Y., Deng, L., Zhang, L., Lang, Q., Liao, C., Wang, N., Qin, M., & Huang, H. (2020). Work stress among Chinese nurses to support Wuhan in fighting against COVID-19 epidemic. *Journal of Nursing Management*, 28(5), 1002–1009. <https://doi.org/10.1111/jonm.13014>
- Nikčević, A. V., & Spada, M. M. (2020). The COVID-19 anxiety syndrome scale: Development and psychometric properties. *Psychiatry Research*, 292, 113322. <https://doi.org/10.1016/j.psychres.2020.113322>
- Rector, T. S., Taylor, B. C., & Wilt, T. J. (2012). Chapter 12: Systematic review of prognostic tests. *Journal of general internal medicine*, 27 Suppl 1(Suppl 1), S94–S101. <https://doi.org/10.1007/s11606-011-1899-y>
- Rossi, R., Socci, V., Pacitti, F., Di Lorenzo, G., Di Marco, A., Siracusano, A., & Rossi, A. (2020). Mental health outcomes among frontline and second-line health care workers during the coronavirus disease 2019 (COVID-19) pandemic in Italy. *JAMA Network Open*, 3(5), e2010185. <https://doi.org/10.1001/jamanetworkopen.2020.10185>
- Steyerberg, E. W., & Harrell, F. E. (2016). Prediction models need appropriate internal, internal-external, and external validation. *Journal of Clinical Epidemiology*, 69, 245–247. <https://doi.org/10.1016/j.jclinepi.2015.04.005>
- Strauss, M. E., & Smith, G. T. (2009). Construct validity: Advances in theory and methodology. *Annual Review of Clinical Psychology*, 5, 1–25. <https://doi.org/10.1146/annurev.clinpsy.032408.153639>
- Walter, M., Alizadeh, S., Jamalabadi, H., Lueken, U., Dannlowski, U., Walter, H., Olbrich, S., Colic, L., Kambeitz, J., Koutsouleris, N., Hahn, T., & Dwyer, D. B. (2019). Translational machine learning for psychiatric neuroimaging. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 91, 113–121. <https://doi.org/10.1016/j.pnpbp.2018.09.014>
- Weinberger, T., Steffen, J., Osterman, A., Mueller, T. T., Muenchhoff, M., Wratil, P. R., Graf, A., Krebs, S., Quartucci, C., Spaeth, P. M., Grabein, B., Adorjan, K., Blum, H., Keppler, O. T., & Klein, M. (2021). Prospective longitudinal serosurvey of healthcare workers in the first wave of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic in a quaternary care hospital in Munich, Germany. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 73(9), e3055–e3065. <https://doi.org/10.1093/cid/ciaa1935>
- Wratil, P. R., Schmacke, N. A., Osterman, A., Weinberger, T., Rech, J., Karakoc, B., Zeilberger, M., Steffen, J., Mueller, T. T., Spaeth, P. M., Stern, M., Albanese, M., Thun, H., Reinbold, J., Sandmeyer, B., Kressirer, P., Grabein, B., Falkai, P., Adorjan, K., ... & Keppler, O. T. (2022). In-depth profiling of COVID-19 risk factors and preventive measures in healthcare workers. *Infection*, 50(2), 381–394. <https://doi.org/10.1007/s15010-021-01672-z>
- Xiao, H., Zhang, Y., Kong, D., Li, S., & Yang, N. (2020). The effects of social support on sleep quality of medical staff treating patients with coronavirus disease 2019 (COVID-19) in January and February 2020 in China. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, 26, e923549. <https://doi.org/10.12659/MSM.923549>

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

Authors and Affiliations

Kristina Adorjan^{1,2,3}  · Mark Sen Dong¹ · Paul R. Wratil^{4,5} · Niklas A. Schmacke⁶ · Tobias Weinberger^{7,8} · Julius Steffen^{7,8} · Andreas Osterman^{4,5} · Alexander Choukér⁹ · Tonina T. Mueller^{7,8} · Tarek Jebrini¹ · Hauke Felix Wiegand¹⁰ · Oliver Tüscher^{10,13,14} · Klaus Lieb^{10,13} · Veit Hornung⁶ · Peter Falkai¹ · Matthias Klein^{11,12} · Oliver T. Keppler^{4,5} · Nikolaos Koutsouleris¹

✉ Kristina Adorjan
kristina.adorjan@med.uni-muenchen.de

✉ Mark Sen Dong
sen.dong@med.uni-muenchen.de

¹ Department of Psychiatry and Psychotherapy, University Hospital, LMU Munich, Nussbaum Str. 7, 80336 Munich, Germany

² Institute of Psychiatric Phenomics and Genomics, University Hospital, LMU Munich, Munich, Germany

³ Center for International Health, Ludwig Maximilians University Munich, Munich, Germany

⁴ Faculty of Medicine, Max Von Pettenkofer Institute National Reference Center for Retroviruses LMU Munich, Virology Munich, Germany

⁵ German Center for Infection Research (DZIF), Partner Site Munich, Munich, Germany

⁶ Gene Center and Department of Biochemistry, Ludwig-Maximilians-University Munich, Munich, Germany

⁷ Department of Medicine I, University Hospital, LMU Munich, Munich, Germany

⁸ DZHK (German Centre for Cardiovascular Research), Partner Site Munich Heart Alliance, Munich, Germany

⁹ Laboratory of Translational Research Stress and Immunity, Department of University Hospital, LMU Munich, Munich, Germany

¹⁰ Department of Psychiatry and Psychotherapy, University Medical Center, Johannes Gutenberg University Mainz, Mainz, Germany

¹¹ Emergency Department, University Hospital, LMU Munich, Munich, Germany

¹² Department of Neurology, University Hospital, LMU Munich, Munich, Germany

¹³ Leibniz Institute for Resilience Research (LIR) Mainz, Mainz, Germany

¹⁴ Institute for Molecular Biology (IMB) Mainz, Mainz, Germany