**Supplement 1. Items of the COVID-19 adapted version of the 7-item Vaccine Conspiracy Beliefs Scale:**

* Vaccine safety data for COVID-19 is faked;
* Immunizing children and other vulnerable groups is harmful and this fact is covered up;
* Pharmaceutical companies are covering up the dangers of COVID-19 vaccines;
* People are lied to about COVID-19 vaccine effectiveness;
* Vaccine effectiveness data is probably faked;
* People are lied to about vaccine safety;
* The government are trying to cover up links between the vaccines being tested and serious health problems

**Supplement 2. Hyperparameter tuning of the machine learning models**

**Strategy for hyperparameter tuning**

The hyperparameter tuning was conducted in a nested-cross-validation procedure with eight outer and four inner folds to identify the best hyperparameter configuration for each model. Due to the fact that classes were skewed we used random undersampling to prevent models potentially optimizing for the majority class. For the hyperparameter tuning we balanced classes on the whole dataset before starting the nested-cross-validation.

 In nested cross-validation outer folds are split in training and testing folds. We split the training data of each outer fold in four inner folds consisting of a training and validation data set each. Hyperparameter configurations were cross validated in the inner folds and the best configuration in the inner fold was applied to the outer fold testing data. We used a randomized search over our hyperparameter space and optimized for the metric accuracy. Random forest classifiers could vary in their number of decision trees (50, 100, 200, 500, 1000, 1500), maximum depth of decision trees (5, 10, 150 20, 25, 30, 35, 400, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, No restriction), maximum features allowed to use per split (20%, 40%, 60%, 80%), and the minimum samples required at each leaf node (10, 15, 20, 25, 30). Furthermore, we always used bootstrapping but varied the bootstrap sample size drawn from each training dataset (25%, 50%, 75%, 100%). Additionally, we pruned decision trees after calculation by varying the degree of model complexity penalization with the use of cost-complexity pruning (0.005, 0.01, 0.0125, 0.015, 0.0175). Our hyperparameter space resulted in 50400 possible configurations and we randomly searched 20% of this hyperparameter space (10080 configurations). Results of the hyperparameter tuning for all models are summarized in Table S2.

 Following hyperparameter tuning, the leave one site out and leave one person out cross-validation of the models with their respective best hyperparameter configuration was performed using random undersampling on each training fold before model calculation, whereas the test folds were based on the full data set.

**Table S2 (Supplement 2)**

*Hyperparameter tuning results and final hyperparameter configuration for each model*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Variables | Mean test accuracy  | Mean training accuracy | Number of trees | Max. depth | Max. features per split | Min. samples per leaf | bootstrap | Ccp alpha |
| Standard model (without middle category) |
| All variables included | 82.42% | 85.32% | 200 | 80 | 80% | 10 | 100% | 0.005 |
| Vaccination conspiracy belief excluded from model | 76.93% | 82.60% | 100 | 60 | 20% | 10 | 75% | 0.005 |
| Specific/General mistrust excluded from model | 67.07% | 74.82% | 1000 | 20 | 20% | 20 | 50% | 0.005 |
| Prediction using the most relevant variables |
| 10 best variables | 82.20% | 83.95% | 200 | 15 | 60% | 10 | 50% | 0.005 |
| 5 best variables | 81.97% | 83.11% | 1500 | 20 | 80% | 15 | 100% | 0.005 |

*Note.* Ccp alpha = alpha value for cost-complexity pruning

**Supplement 3. Strategy for permutation feature importance analysis**

We used permutation feature importance for the site cross-validation analysis. Each variable of the testing data was permuted ten times. The accuracy of models was then re-evaluated with the permutation of each variable and the overall accuracy of each resulting model with permutation was juxtaposed with the original model. For each variable, a score of the mean of the 10 differences in accuracy between the original model and the model with one of the 10 permutations of this variable was calculated to estimate its importance in a given model in terms of reduction in model-accuracy without meaningful information from this variable.

**Supplement 4**

**Table S4**

*Mean values and standard deviations for all psychosocial predictors by site*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variables | UK (n=512) | USA (n=535) | AU (n=502) | GE (n=516) | HK (n=445) |
|  | M/% | SD | M/% | SD | M/% | SD | M/% | SD | M/% | SD |
| *Risk perception* |  |  |  |  |  |  |  |  |  |  |
| COVID anxiety | 57.00 | 28.08 | 53.63 | 32.80 | 51.93 | 29.46 | 58.66 | 29.56 | 55.98 | 24.29 |
| Close people have been infected  | 40.63% | - | 41.50% | - | 13.94% | - | 28.29% | - | 6.74% | - |
| Perceived infection risk  | 4.82 | 2.25 | 4.02 | 2.55 | 4.03 | 2.79 | 4.76 | 2.50 | 4.06 | 2.25 |
| Perceived consequences of infection | 42.75 | 26.55 | 37.48 | 30.70 | 48.75 | 32.14 | 48.20 | 30.24 | 46.54 | 27.44 |
| *Political views* |  |  |  |  |  |  |  |  |  |  |
| Political views | 3.66 | 1.43 | 4.17 | 1.72 | 3.57 | 1.36 | 3.65 | 1.17 | 3.74 | 1.18 |
| Primary information source | 2.41 | 1.24 | 2.46 | 1.31 | 2.58 | 1.28 | 2.60 | 1.19 | 2.37 | 1.29 |
| *Specific Mistrust* |  |  |  |  |  |  |  |  |  |  |
| Pandemic persecutory threat (PPS) | 2.48 | 7.03 | 4.70 | 10.71 | 7.20 | 12.20 | 3.11 | 7.11 | 5.27 | 9.05 |
| Pandemic paranoid conspiracy (PPS) | 3.85 | 3.22 | 3.48 | 3.44 | 4.10 | 3.25 | 2.93 | 3.01 | 3.52 | 2.73 |
| Pandemic interpersonal mistrust (PPS) | 3.66 | 4.83 | 4.71 | 5.92 | 5.50 | 5.94 | 4.05 | 5.20 | 6.01 | 5.36 |
| General Pandemic Paranoia (PPS) | 6.68 | 9.05 | 9.09 | 12.97 | 12.18 | 14.60 | 6.94 | 9.07 | 10.42 | 11.15 |
| Vaccine conspiracy beliefs | 16.32 | 11.39 | 20.55 | 12.77 | 22.41 | 12.31 | 18.62 | 11.83 | 19.91 | 11.30 |
| *General mistrust* |  |  |  |  |  |  |  |  |  |  |
| Ideas of reference (RGPTS) | 6.46 | 8.59 | 6.77 | 8.69 | 11.73 | 8.52 | 7.13 | 8.60 | 9.17 | 7.07 |
| Paranoid ideation (RGPTS) | 6.03 | 9.66 | 6.85 | 10.42 | 12.01 | 11.02 | 5.72 | 9.11 | 8.52 | 9.67 |
| General conspiracy mentality (CMQ) | 29.95 | 10.43 | 31.30 | 12.43 | 34.09 | 10.35 | 24.45 | 12.44 | 34.61 | 9.73 |
| *Social adversity* |  |  |  |  |  |  |  |  |  |  |
| Traumatic emotional neglect | 30.66% | - | 33.64% | - | 61.75% | - | 44.19% | - | 45.39% | - |
| Traumatic psychological abuse | 25.59% | - | 34.95% | - | 50.80% | - | 36.82% | - | 30.34% | - |
| Traumatic physical abuse | 20.90% | - | 30.52% | - | 39.64% | - | 30.62% | - | 12.41% | - |
| Traumatic sexual abuse | 18.16% | - | 28.41% | - | 38.84% | - | 19.96% | - | 17.08% | - |
| *Generalized beliefs*  |  |  |  |  |  |  |  |  |  |  |
| Negative beliefs about self (BCSS) | 2.92 | 4.70 | 2.60 | 4.73 | 5.35 | 5.81 | 3.29 | 4.85 | 3.48 | 5.31 |
| Negative beliefs about others (BCSS) | 4.52 | 5.83 | 5.76 | 6.79 | 6.63 | 6.97 | 5.55 | 5.96 | 2.88 | 5.17 |
| Positive beliefs about self (BCSS) | 9.83 | 6.15 | 14.10 | 6.94 | 8.72 | 6.39 | 12.80 | 6.46 | 11.38 | 6.21 |
| Positive beliefs about others (BCSS) | 9.83 | 6.09 | 11.14 | 6.76 | 8.81 | 6.39 | 10.81 | 5.87 | 8.54 | 6.02 |
| Low perceived social rank (SCS) | 5.54 | 1.71 | 6.49 | 1.99 | 5.17 | 2.00 | 6.34 | 1.80 | 6.08 | 1.50 |

 Note: PPS = Pandemic Paranoia Scale; CMQ = Conspiracy Mentality Questionnaire; RGTPS = Revised Green Paranoid Thoughts Scale; Scale; BCSS = Brief Core Schema Scales; SCS = Social Comparison Scale.