Predicting and Understanding Vaccine Hesitancy in Adults: Insights into Modifiable Risk Factors from a Mixed Methods Multiple Population Study Combining Machine Learning and Thematic Analysis

Omid V. Ebrahimi[1*], Ella Marie Sandbakken[2,3], Sigrun Marie Moss[2], Sverre Urnes Johnson[2,4], Asle Hoffart[4,2], Sarah Bauermeister[5], Ole André Solbakken[2], Lars T. Westlye[2,6,7], and Esten H. Leonardsen[2,6]

[1]Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom

[2]Department of Psychology, University of Oslo, Oslo, Norway

[3]Department of Psychology, Oslo New University College, Oslo, Norway

[4]Modum Bad Psychiatric Hospital and Research Center, Vikersund, Norway

[5]Department of Psychiatry, University of Oxford, Oxford, United Kingdom

[6]Centre for Precision Psychiatry, Division of Mental Health and Addiction, Oslo University Hospital & Institute of Clinical Medicine, University of Oslo

[7]KG Jebsen Centre for Neurodevelopmental Disorders, University of Oslo, Oslo, Norway

Author Note

* Corresponding author:

Omid V. Ebrahimi

University of Oxford, Department of Experimental Psychology

Anna Watts Building, Woodstock Rd

Oxford OX2 6GG

Citation: Ebrahimi, O. V., Sandbakken, E. M., Moss, S. M., Johnson, S. U., Hoffart, A., Bauermeister, S., Solbakken, O. A., Westlye, L. T., & Leonardsen, E. H. (2024, July 4). Predicting and understanding vaccine hesitancy in adults: insights into modifiable risk factors from a mixed methods multiple population study combining machine learning and thematic analysis. Retrieved from: https://osf.io/preprints/psyarxiv/zmfjv. doi: 10.31234/osf.io/zmfjv

Words: 6072, **Tables:** 2, **Figures:** 3

Supplementary Tables: 5, **Supplementary Figures:** 1, **Supplementary Appendices:** 2

Abstract

Background

Vaccine hesitancy, the delay in acceptance or reluctance to vaccinate, ranks among the top threats to global health and undermines efforts to control vaccine-preventable diseases. The identification of modifiable risk factors and psychological processes underlying vaccine hesitancy is imperative to inform targeted interventions aimed at increasing vaccination uptake.

Methods

Leveraging gradient boosting machines and thematic analysis, this mixed methods multiple population study aimed to identify modifiable predictors of vaccine hesitance in the general adult population. Predictors of vaccine hesitance were investigated in 2926 Norwegian adults (M_{age} = 37.91, 79.69% female) during the COVID-19 pandemic, before the predictive utility of these variables was investigated in an independent sample of 734 adults in the United Kingdom (M_{age} = 40.34, 57.08% female), enabling an examination of generalizability across populations. Two teams of authors independently conducted the machine learning and thematic analyses, blind to the analytic procedure and results of each other.

Results

The machine learning model performed well in discerning vaccine hesitant (*n* = 248, 8.48% and *n* = 109, 14.85%, Norway and UK, respectively) from vaccine uptaking individuals (*n* = 2678, 91.52% and *n* = 625, 85.15%), achieving an AUC of 0.94 (sensitivity = 0.81; specificity = 0.98) in the Norwegian sample, and an AUC of 0.98 (sensitivity = 0.83; specificity = 0.97) in the out-of-sample replication in the UK. The mixed methods investigation identified five categories of modifiable risk factors and psychological processes tied to vaccine hesitancy, including illusion of invulnerability, doubts about vaccine efficacy, distrust in official entities, minimization of the societal impact of COVID-19, and health-related fears tied to vaccination.

Moreover, the media's portrayal of rare incidents as fear amplifiers, and stigmatizing presentation of unvaccinated individuals were provided as additional underlying motives leading to vaccine reluctance and polarization. The thematic analysis further revealed information overload, fear of needles, previous negative vaccination experiences, fear of not getting healthcare follow-up after vaccination if needed, and vaccine aversion due to underlying (psychiatric) illness (e.g., eating disorders) as motives underlying vaccine hesitance.

Conclusions

The identified influential predictors were consistent across the two European samples, highlighting their generalizability across European populations. These predictors offer insights about modifiable factors that could be adapted by public health campaigns in mitigating misconceptions and fears related to vaccination toward increasing vaccine uptake. Moreover, the results highlight the media's responsibility, as mediators of the public perception of vaccines, to minimize polarization and provide accurate portrayals of rare vaccine-related incidents, reducing the risk aggravating fear and reactance to vaccination.

Keywords: Vaccine Hesitancy, Mixed Methods Study, Machine Learning, Extreme Gradient Boosting (XGBoost), Thematic Analysis, General Adult Population, COVID-19 Pandemic

Predicting and Understanding Vaccine Hesitancy in Adults: Insights into Modifiable Risk Factors from a Mixed Methods Multiple Population Study Combining Machine Learning and Thematic Analysis

Vaccine hesitancy is ranked among the world's leading health challenges, labeled as one of the top 10 threats to global health by the World Health Organization (WHO, 2019). This phenomenon is defined as a delay in acceptance or reluctance to vaccinate despite the availability of efficacious vaccines (Dubé et al., 2021; MacDonald et al., 2015). A key factor promoting vaccine hesitancy as a major global health problem is its ability to reverse progress made in tackling vaccine-preventable disease (Obregon et al., 2020; WHO, 2019). Moreover, vaccination is a key proactive measure in protecting against illness, with increases of vaccine coverage estimated to reduce an additional of 1.5 million disease-related deaths per year worldwide (Ryan & Malinga, 2021; WHO, 2019).

The onset of the recent SARS-CoV-2 pandemic highlighted the general public's dependency on the availability of efficacious vaccines across the globe, with the course of the pandemic being interconnected with the existence and routine uptake of vaccination (Anderson, Heesterbeek, et al., 2020; Sachs et al., 2022). The successful implementation of mass vaccination programs, however, depends on the public's willingness to vaccinate, rendering vaccine hesitancy as a critical problem during current and forthcoming health crises (Sachs et al., 2022; Wiysonge et al., 2022).

Several studies have outlined challenges with vaccine uptake across nations, calling for research to identify predictors of vaccine hesitance (Anderson, Vegvari, et al., 2020; Lazarus et al., 2021; Neumann-Böhme et al., 2020). Specifically, more knowledge is needed about the psychological mechanisms predicting vaccine hesitancy, with these processes likely to be informative across different contexts (Trogen & Pirofski, 2021). An optimal prediction of vaccine hesitancy would constitute the identification of a limited set of key variables with

high predictive power, maximizing applicability and practicality of developing public health measures designed to increase vaccine uptake (Chowdhury & Turin, 2020). Moreover, testing the predictive utility of identified predictors in independent samples is imperative to examine their generalizability across populations (Obermeyer & Emanuel, 2016). It would further be of added utility to identify modifiable factors that are subjectable to change (e.g., risk perception of vaccines) above more stable risk factors such as lower education levels and biological sex (Khairat et al., 2022).

One promising category of modifiable risk factors are cognitions and attitudes (MacDonald et al., 2015). This includes beliefs about the efficacy of vaccination in mitigating societal transmission rates, perceived danger of the virus for societal functioning, and degree to which individuals believe vaccination can effectively protect them against infection (e.g., Davis et al., 2022). In building models incorporating such modifiable factors, it is important to rely on the existing literature to investigate the comparative utility of novel variables against previously identified predictors of hesitancy, facilitating the identification of variables with the greatest importance in predicting hesitance. Previously, a range of demographic factors have been tied to vaccine hesitancy, including unemployment, lower education, younger age, rural residency, female sex, and migrant status (Lincoln et al., 2022; Salali & Uysal, 2020). Additionally, sociopolitical and contextual features previously related to hesitancy include governmental trust, COVID-19 anxiety, perceived risk of infection, and social media use (Lincoln et al., 2022; Shen, 2020).

Accordingly, controlling for the impact of these known risk factors, we leveraged the large-scale Norwegian MAP-19 study (Ebrahimi et al., 2022) to investigate the predictive utility of a range of beliefs and attitudes about vaccination, above and beyond sociodemographic variables previously identified to be related to vaccine hesitancy in the literature. To investigate the generalizability of the findings and the utility of identified key

predictors across nations, we aimed to perform an out-of-sample replication in a large-scale representative sample from the United Kingdom (McBride et al., 2020).

While machine learning models are powerful for identifying key factors that predict vaccine hesitancy, this predictive advantage can come at the cost of explanation (Bzdok et al., 2018; Stenwig et al., 2022). Moreover, a general limitation with quantitative analyses includes that the use of pre-defined and standardized measures reduces the opportunity to identify novel processes that may be related to hesitancy in an open and unrestricted manner (Mukumbang, 2023). Accordingly, to increase insight about cognitions and mechanisms related to vaccine hesitance, we supplemented these predictive models with a qualitative thematic analysis (Brooks et al., 2015) on all the subjects who provided written open-ended explanations for their hesitance. This mixed methods study design allows a more detailed and nuanced understanding of vaccination hesitance (Power et al., 2018). To avoid influence from one set of analytic results to the other and to further cross-validate the findings, two groups of researchers conducted each set of analyses (i.e. machine learning and thematic analysis) independently.

In summary, this study sought to identify a) the most influential modifiable predictors of vaccine hesitancy while controlling for demographic characteristics related to the phenomena; b) investigate the generalizability of the findings by conducting an out-of-sample replication to examine the utility of identified influential predictors in an independent sample from another nation; in addition to c) provide a more nuanced and comprehensive understanding of vaccine hesitancy through a qualitative investigation of respondents' motvies underlying their hesitance.

Methods

The present study is part of The Norwegian COVID-19, Mental Health and Adherence Project (MAP-19), a large longitudinal study investigating mental health and preventive

health behaviors in the adult population during the COVID-19 pandemic (Ebrahimi et al., 2022). The study was approved by The Norwegian Regional Committee for Medical and Health Research Ethics (REK; reference: 125510) and the Norwegian Centre for Research Data (NSD; reference: 802810).

Study Design, Participants and Procedure

Eligible participants included all adults (age >= 18 years) residing in Norway who provided informed consent to participate in the study. Upon initial recruitment in March 2020 (T1), participants joined through participation in an online survey disseminated to a random selection of Norwegian adults through a Facebook Business algorithm, in addition to systematic dissemination of the survey via national, regional, and local information platforms (i.e. television, radio, and newspapers). This procedure is elaborated in detail elsewhere (Ebrahimi et al., 2021; Magnúsdóttir et al., 2022). The sampling procedure involved the recruitment of a proportional number of subjects from each region of the country with respect to the region's population size to yield a geographically representative sample of the adult population. Data from the eighth wave of the MAP-19 study (T8; collection period: January 2 to January 14, 2022) was used for the present study, where vaccine hesitancy along with beliefs and attitudes about vaccination were measured. 2926 adults provided data at this assessment wave, rendering them eligible for our analyses. 248 participants (8.48%) were vaccine hesitant, corresponding to the known rate of adults unvaccinated against COVID-19 (approximately 9%) in the Norwegian population around the measurement period (Norwegian Institute of Public Health, 2021). Respondents had the opportunity to receive a pair of noise cancelling headphones for participation. Data quality was further examined using attention checks (Braitman et al., 2022). This was done by asking participants to "Please provide the response "a little" if you are paying attention to this survey" (response options: not at all; a

little, moderately; a lot; extremely). 97.80% of the subjects passed these checks, with subjects failing the attention checks being excluded to maintain high data quality.

Measurement

A list of all variables used by the predictive algorithm in the present study, along with their description and coding, is provided in Supplementary Document 1. Examples of the categories of variables used are presented below.

Vaccine Hesitancy

Following assessment practices in the literature, vaccine hesitancy was measured by querying participants about whether they planned to vaccinate against COVID-19 upon availability of a vaccine, with adults responding no coded as hesitant (binary scale: hesitant: 1; willing: 0), yielding the label (i.e. outcome variable) of the study (Neumann-Böhme et al., 2020).

Sociodemographic Predictors

The subjects provided sociodemographic information such as their biological sex (females: 0; males: 1), age (continuous in years; age groups: 18-30 years; 31-44 years; 45-64 years; and 65 years and above), relationship status (in a relationship: 1; single: 0) and education level (compulsory school: 0; upper secondary high school: 1; student: 2; any university degree: 3). Participants further provided information about their current living situation, urban versus rural residency, cultural background, and employment status (see Supplementary Document 1 for full details).

Medical and Psychiatric Comorbidities

Information regarding the presence of a preexisting psychiatric comorbidity, medical comorbidity, and presence of a chronic illness was obtained. Mental health covariates were included using validated instruments measuring general anxiety symptoms (GAD-7; Spitzer et

al., 2006), sleep difficulties (BIS; Pallesen et al., 2008), post-traumatic stress symptoms (PCL-5; Bovin et al., 2016), and depression (PHQ-9; Kroenke et al., 2001).

Contextual Variables

Participants reported their general level of adherence to pandemic mitigation protocols including social distancing and hygienic behavior recommendations, governmental trust, frequency of information acquisition about the pandemic from different information sources, financial and occupational concerns, worries related to self-infection, and fear about significant others being infected by the SARS-CoV-2 virus (Ebrahimi et al., 2023; Supplementary Document 1).

Cognitions and Beliefs Related to Vaccination and the Pandemic

Attitudes and beliefs related to vaccination, the pandemic, and the SARS-CoV-2 virus were measured by asking subjects to rate the extent to which they agreed with different statements, measured on a five-point Likert scale (1–5; Completely Disagree: 1 to Completely Agree: 5). For instance, overconfidence in ability to avoid the SARS-CoV-2 virus was measured through the statement “I do not need to vaccinate as I have managed to avoid the coronavirus so far”. Other measured cognitions and beliefs included fear of vaccination due to side-effects or related to an underlying illness, extent of belief in whether vaccines were developed too fast to be safe, perceived dangerousness of COVID-19 as an overall societal problem, extent of trust in information disseminated about vaccines from health-care officials, extent of belief in the ability of vaccines to effectively mitigate societal transmission rates, protect against COVID-19 infection, and the perceived health risk associated with vaccinating (Supplementary Document 1). The extent to which the subjects believed in the superiority of natural immunity as compared to vaccination was further measured.

Machine Learning Analyses

The machine learning analysis was performed by the first and last author, blind to the analytic procedures and results of the qualitative team (second and third author).

Data Splitting

To evaluate the performance of our models in an unbiased manner, and obtain a singular list of variable importances, we reserved ⅓ of our dataset for testing (test set, n_{test} = 975) before doing any model fitting. The splitting was done while stratifying on vaccine hesitancy and a list of demographic and psychological variables (Supplementary Table S1) to get a similar distribution between the folds (i.e. subsets of data). We used the same strategy to split the remaining ⅔ of the dataset, constituting the training set (n_{train} = 1951), into five similar folds to facilitate cross-validation for tuning hyperparameters.

Variable Selection and Encoding

When predicting vaccine hesitancy, we used a set of 111 variables, comprising of mentioned demographic factors and psychological variables such as attitudes and cognitions all measured prior to or at the same time as the outcome (detailed in Supplementary Document 1). The outcome variable (vaccine hesitancy) was coded as a binary variable (hesitant: 1; willing: 0). For the predictors, missing data was retained and handled as a distinct level by the trained models (Shen et al., 2023).

Model Fitting

All analyses were run using Python 3.8.6 on a machine running Windows Server 2012. Gradient boosting machines (Friedman, 2001) were fitted using the XGBoost library version 1.6.1 (Chen & Guestrin, 2016). All models were trained to optimize the log loss:

$$\frac{1}{N}\sum_{i} -(y_i * \log(\hat{y}_i) + (1 - y_i) * \log(1 - \hat{y}_i))$$

where y_i denotes the response by person *i*, $\hat{y}_i$ the predicted response for person *i*, and *N* the number of participants. To find optimal hyperparameters, we performed an inner cross-validation in the training set, searching for parameters in a complete grid (Varma & Simon, 2006). Due to the sample size restrictions, we contrained our search to five parameters with either two or three possible values (Supplementary Table S2), resulting in 108 models per fold. For model selection in the cross-validation loop, we used the area under the receiver operating characteristic curve (ROC AUC, AUC for short; Dinga et al., 2019), selecting the hyperparameter settings that yielded the highest average AUC across folds.

Model Evaluation

In the final model evaluation, we mainly relied on the AUC. Given the ability of different performance metrics to capture unique properties of the model (Varoquaux & Colliot, 2023), we computed a set of complementary evaluation metrics, defined below. All metrics rely on the number of true positives (*tp*), true negatives (*tn*), false positives (*fp*) and false negatives (*fn*):

$$Accuracy = \frac{(tp + tn)}{(tp + tn + fp + fn)}$$

$$Balanced\ accuracy = \frac{1}{2}\left(\frac{tp}{tp + fn} + \frac{tn}{tn + fp}\right)$$

$$Precision = \frac{tp}{(tp + fp)}$$

$$Recall\ (Sensitivity) = \frac{tp}{(tp + fn)}$$

$$Specificity = \frac{tn}{(tn + fp)}$$

The aforementioned metrics are all contingent on a threshold to dichotomize what is classified as hesitant versus willing (the classification threshold), often set to 0.5. In the

present study, models were selected using the AUC, which does not rely on accurately setting this classification threshold. Thus, we observed that the threshold used by the best performing model was poorly tuned (Provost, 2008). To alleviate this, we used the ROC curve from the training data to set a moving classification threshold for dichotomization (Calvert & Khoshgoftaar, 2019). This was based on the True Positive Rate, $tpr = \frac{tp}{(tp + fn)}$, and the False Positive Rate, $fpr = \frac{fp}{(fp + tn)}$, at each threshold *t* in the ROC curve. From these, a geometric mean was computed:

$$g_t = \sqrt{tpr_t * (1 - fpr_t)}$$

Next, we selected the *t* yielding the highest geometric mean, $t^* = \max_t g_t$, which in the training set was empirically determined as $t^* = 0.13$. This value represents the optimal classification threshold and was used in the calculations of the aforementioned metrics in the test set.

Variable Importance

To determine variable importance, we used the default XGBboost functionality based on calculating the average gain for each split using each predictor. For an individual branching in each decision tree, the gain denotes how much information is obtained by performing the split, measured by how much it decreases the prediction error:

$$Gain = \frac{1}{2} * \left(\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right) - \gamma$$

where $G_{\{L,R\}}$ and $H_{\{L,R\}}$ describe the first and second derivatives of the loss for all samples that end up in the given node of the tree, and λ (L2-loss) and γ (minimum loss reduction required for further splitting) are regularization terms. To investigate the differences between the willing versus the vaccine hesitant individuals in the most important variables,

we performed post-hoc inspections of the differences in the distributions of these variables across the two groups.

Independent Sample Replication

For replication, we used a large open-source study from the United Kingdom using identified variables resembling those in our original dataset (McBride et al., 2020). To maximize the overlap of variables in the UK dataset and the main (Norwegian) dataset while retaining a large enough sample, we used the data collected at wave 5 in the UK data. We trained a new model to predict vaccine hesitancy, encoded in the binary target variable *W5_C19_Vax_Self*, asking whether participants would take a vaccine for COVID-19 when it becomes available to them (McBride et al., 2020). We selected all variables in this data set which resembled the 10 most important variables in our model trained on the Norwegian dataset, resulting in the selection of 18 variables from the UK data (Figure 2 and Supplementary Table S4). Also here, we split the data into a test split and a training split (n_{train} = 490, n_{test} = 244) and performed an inner 5-fold cross-validation to find optimal hyperparameters (Varma & Simon, 2006). The hyperparameter settings tested were the same as for the original model (Supplementary Table S2).

Thematic Analysis

Without insight into the quantitative design or analysis of the material, the second and third authors independently analyzed the qualitative survey material. Out of the 248 vaccine hesitant subjects in the Norwegian sample, 53 filled out the open-ended question: "If you did not get vaccinated against COVID-19, could you please explain what contributed to you not wanting to get vaccinated?". Of the 53 responses, one was removed as it was too limited to analyze (the respondent merely answered "no"). We were left with 52 extracts (range: 4 to 206 words, mean response word length: 51 words), male: 9, $\overline{age}$: 51.8, range: 20-70; female: 43, $\overline{age}$: 40.3, range: 20-84.

We adopted a template analysis approach to thematic analysis (Brooks et al., 2015; King, 2012) with a critical realist epistemological stance (Wiltshire & Ronkainen, 2021). Aiming for transparency, we have incorporated an audit trail (as per for example Gale et al., 2013; Steltenpohl et al., 2023) to allow readers access to and overview of the analytic process (see Supplementary Document 2). The second and third author separately coded all the qualitative data material for manifest meaning and latent meaning. Based on this, we developed an initial template in the form of an elaborate, hierarchical coding template (8 superordinate codes, 2-6 subordinate under each, and up to 3 sub-subordinate codes, with a total of 50 codes across those three levels; Supplementary Document 2). This initial coding template was developed inductively to allow for open analysis and combating the risk of "foreclosure of analysis" (Braun & Clarke, 2022). Having developed the template, the second and third author separately coded the whole dataset using the template. Of 317 coding instances, 282 coding instances overlapped at main theme level (88.96%) between the authors, and 260 overlapped at sub-theme level (82.02%). Where the authors had coded differently, this was resolved through discussion and refinement of the coding template. Thereafter, and based on the separate coding and the discussions of the coding, we generated the final coding template, presented as a hierarchical thematic map, consisting of six main themes, 2-5 subthemes under each, and 0-6 sub-subthemes (see Figure 3 for overview). All extracts used in the article are translated from Norwegian and anonymized to maintain confidentiality.

Open Science, Transparency, and Reproducibility

The machine learning analyses were performed using Python 3.8.6, with our code openly available. The Jupyter Notebook documentation of the full procedure and findings can be found here (https://osf.io/nbrwz/), where also the visualized plots are readily viewable upon opening the code files in a matching coding environment (e.g., Visual Studio Code,

Jupyter Notebook). The audit trail of the thematic analysis can be found in Supplementary Document 2, which provides a transparent documentation of the analytical process in a step-by-step manner (Levitt et al., 2018; Steltenpohl et al., 2023). The ethical approval for the Norwegian dataset precludes submission of data to public repositories. The UK data is openly available and can be found at the online repository of the Center for Open Science (https://osf.io/v2zur/).

Results

Among the 2926 participants eligible for the study in the Norwegian sample, ages ranged from 18 to 86 years ($\overline{age}$ = 37.91), with 2332 (79.69%) being female, 179 of ethnic minority status (6.16% non-white Norwegian), and 1767 (60.39%) having a university degree. 248 (8.48%) reported hesitance to vaccinate. The distribution of these demographic variables in the vaccine hesitant group is reported in Supplementary Table S5. The percentage of participants reporting preexisting mental health conditions in this sample was 17.98%, representative of the known rate of psychiatric disorders in the Norwegian adult population, which is between 16.66% and 25.00% (Norwegian Institute of Public Health, 2016). The quota of participants sampled from each region of Norway was further proportional to the respective region's size, resulting in a geographically representative sample (Ebrahimi et al., 2021).

Predictive Performance

In the independent test set of the Norwegian sample, the best model achieved an AUC of 0.94, corresponding to a balanced accuracy (with an adjusted classification threshold, see Methods) of 86.27% (Table 1 and Figure 1). The model was able to correctly identify a large portion of those willing to vaccinate (*Specificity* = 0.98) and further a substantial amount of those hesitant (*Sensitivity* = 0.81). Partially due to the class imbalance, the positive predictive

value was low (*Precision* = 0.48). Overall, when using an appropriate classification threshold, the model performed well at discerning vaccine hesitant from willing participants.

Influential Predictors

The importance of the individual predictors were assessed based on the resulting information gain from their inclusion in the model (see Methods). The 10 variables most influential for predicting vaccine hesitancy are all listed in Figure 2a. The variable yielding the highest predictive utility was overconfidence in one's ability to avoid the COVID-19 virus. Post-hoc analyses (Supplementary Table S3 and Figure 2b) revealed that the distribution of responses for this variable in the willing subgroup of the test set was narrow ($mode_{willing}$ = 1, $SD_{willing}$ = 0.59), reflecting that most vaccine willing individuals showed similar extents and displayed low levels of this belief. In the vaccine hesitant group, responses were generally more spread out, with a greater proportion of subjects shifted toward having higher confidence in their ability to avoid the virus ($mode_{hesitant}$ = 3, $SD_{hesitant}$ = 1.22), as portrayed in Figure 2b. Examples of other influential predictors of hesitancy included fear of side-effects due to illness, distrust in information about vaccines from health officials, doubts about the efficacy of vaccination in both mitigating transmission in society and protecting against infection, believing in the superiority of natural immunity over vaccination, that vaccines were developed too fast to be safe, governmental distrust, and minimization of COVID-19 as a societal problem.

Notably, all 10 key predictors of vaccine hesitancy were cognition-related and attitudinal variables, and a similar patterns of differences in response distributions were identified for the 9 other key predictors (Figure 2a and 2b) as for the avoidance overconfidence variable, revealing more similar beliefs held in the willing group (typically towards one extreme of the scale) versus more scattered belief patterns in the hesitant group (often with a shifted mode towards either the center or the opposite extreme).

Independent Sample Replication Results

To further validate our findings, we replicated our predictive analysis in a representative sample from the UK which included a similar set of variables (n = 734, $\overline{age}$ = 40.34, n_{female} = 419 (57.08%), 14.47% being of non-white British or Irish ethnicity, 26.98% with an undergraduate university degree or higher, 18.12% reporting a pre-existing treatment history with mental health problems). 109 of the 734 participants (14.85%) reported hesitance to vaccinate. The distribution of demographic variables within the vaccine hesitant group can be found in Supplementary Table S5. We selected 18 variables which most closely resembled the top 10 influential variables identified in the predictive model of Norwegian population (Supplementary Table S4), and fit new models using an equivalent data splitting strategy (n_{train} = 489, n_{test} = 245) and inner cross-validation procedure. Here, the best performing model achieved an out-of-sample AUC of 0.98 and a balanced accuracy of 89.96% (Table 1), outperforming the model trained on Norwegian data. In general, beyond the availability of continuous versus ordinal predictors yielding greater statistical power, the differences between hesitant and willing subgroups was more clearly pronounced in the distributions of the predictors in UK versus the Norwegian sample (Supplementary Figure S1b), yielding an improved prediction. That is, greater polarization in the vaccine-related cognitions and attitudes between the hesitant and willing groups were observed in the UK versus Norway. Compared to the Norwegian model, the UK model had a similarly high specificity (0.98 vs 0.97, Norwegian vs UK sample, respectively) and recall (sensitivity; 0.81 vs 0.83), but significantly higher precision (0.48 vs 0.79). Overall, these results revealed that the predictors identified in the Norwegian population generalize to the UK population.

Thematic Analysis Results

To get more in-depth insight into the participants' underlying reasons and own explanations of their vaccine hesitancy, the open-ended survey responses (n = 52) available in

the Norwegian sample were analyzed using a template analysis approach to thematic analysis, resulting in six main themes. Respondents often provided multiple rather than a single reason underlying their vaccine hesitance (an overview of the responses with participant numbers is available in Supplementary Document 2). Figure 3 presents the full thematic map including the main themes and both levels of subthemes, with Table 2 embodying the main themes and the first subtheme level including illustrative quotes from the respondents as characteristic examples for each subtheme.

Theme 1, *Unnecessary vaccines* (*n* = 17, prevalence of explanation: 32.69% of the qualitative vaccine hesitant sample), consisted of statements where respondents did not view the vaccines as necessary, either since they did not perceive COVID-19 as dangerous or did not see themselves as being in the risk group. Theme 2, *Inefficient vaccines* (*n* = 19, explanation prevalence: 36.54%), concerned respondents' views of the vaccines as not offering much protection or mitigating transmission, as well as beliefs disqualifying the vaccines as real vaccines. Theme 3, *Frightening vaccines* (*n* = 32, 61.54%), was the most prevalent motive provided, where fear of the side-effects of the vaccines were emphasized, including worry about both potential short-term and long-term side-effects, fear of not getting healthcare follow-up if needed after vaccination, and worry about the approval of the vaccines as being too quick to be safe. Theme 4, *Distrust and polarization* (*n* = 18, 34.62%), characterized participants who displayed distrust in the authorities' vaccine handling, in the pharmaceutical industry, or the media's portrayal of the vaccines. Here, some respondents mentioned that the polarizing and stigmatizing portrayals of the unvaccinated was also a reason for becoming more reluctant to get vaccinated. Theme 5, *Alternatives to getting vaccinated* (*n* = 15, 28.85%), included the beliefs and measures participants listed in handling COVID-19 instead of getting vaccinated, including perceiving natural immunity as a better alternative than vaccination, seeing potential deaths from COVID-19 as the course of nature,

taking own health measures such as vitamin supplements, avoiding potential sites of infection such as crowded places, as well as getting overwhelmed with the choice of vaccination and the information overload surrounding the topic and thus mentioning that it was easier to not take a stance about vaccination. Finally, theme 6, *Reasons related to vaccines in general* (*n* = 12, 23.08%), included a general fear of needles, aversion of vaccination related to an existing disease such as eating disorders, and negative experiences with previous vaccinations. Combined, these results map out different underlying motives that were provided related to why hesitant adults did not wish to get vaccinated.

Discussion

Using two independent analytical teams and datasets across different countries (Ebrahimi et al., 2021; McBride et al., 2020), this study aimed to identify a core set of modifiable predictors of vaccine hesitancy in comparison with other variables previously found to be related to this phenomenon. Extending beyond traditional limitations of the predictive paradigm, we further conducted independent qualitative investigations aimed at understanding hesitant individuals' articulated motives underlying their vaccine hesitance.

Among the 111 variables investigated, our predictive analyses in a sample of Norwegian adults revealed that a highly accurate prediction of vaccine hesitancy can be achieved through 10 key factors. In line with calls in the literature underscoring the importance of investigating the utility of identified predictors across independent samples (Obermeyer & Emanuel, 2016), we successfully replicated this high prediction accuracy using a comparable set of variables in an independent sample from the United Kingdom (McBride et al., 2020). This supports the robustness and generalizability of the findings across European populations, and corroborates the utility of the identified variables as influential predictors of vaccine hesitance.

Overall, our mixed methods investigation identified five common domains of modifiable risk factors and psychological processes tied to vaccine hesitancy that overlapped across the predictive and qualitative approach.

The first domain concerned an illusion of invulnerability, where overconfidence in one's ability to avoid the COVID-19 virus and belief in the superiority of natural immunity above vaccination were identified as prominent predictors of hesitancy. This is consistent with a recent study identifying a desire to develop natural immunity and perceiving not needing vaccination as key predictors of hesitancy (Steinmetz, 2022). The thematic analysis further identified beliefs of not being in a risk group as a motive for not getting vaccinated. This identified underestimation of the risks the SARS-CoV-2 virus suggests a need for targeted communication strategies, which could address these misconceptions by emphasizing the unpredictable nature of novel emergent viruses and the potential for severe outcomes, even for healthy individuals (Poudel et al., 2021; Tenforde et al., 2020). We also identified hesitant adults reporting that they had taken their own alternative health measures against vaccination, including uptake of vitamin supplements as an alternative to vaccination. This finding highlights the need for evidence-based health literacy campaigns to dispel misconceptions, ensuring that individuals are equipped with accurate knowledge about effective preventive measures and the limitations of alternative health practices in conferring immunity against the virus (Zhang et al., 2022).

The second domain of influential processes concerned health-related fears tied to vaccination, with stronger fear of side-effects, greater perceived risk of vaccination, and believing that the vaccines were developed too fast to be safe (Paul et al., 2021) predicting hesitancy. This underscores the need for public health measures communicating the rigorous systematic procedures underlying vaccine development, even when expedited (Wong et al., 2022). Perceiving the vaccines as potentially harmful was the most frequently listed motive

underlying hesitance among the participants. Our open-ended analyses further extend the literature by revealing that some participants were unwilling to vaccinate due to fears of not getting healthcare follow-up if needed after vaccination. This finding reflects a need for communication about the presence of post-vaccination support systems to address and alleviate such concerns, with the potential of fostering a greater sense of safety to mitigate these health-related fears.

The third influential domain of factors underlying hesitance concerned doubts about the efficacy of vaccination, including doubting vaccines' ability to protect individuals against infection, and the efficacy of the vaccine in mitigating societal transmission overall (Phillips et al., 2022; Troiano & Nardi, 2021). Given the evidence of the significant role that vaccination plays in reducing diseases and deaths worldwide (e.g., Andre et al., 2008; Haas et al., 2021), this finding highlights the gap between scientific evidence and public perception of vaccinations across certain subgroups of the population. Future studies would do well to combine strategies both aimed at identifying these subgroups and tailored ways to minimize this gap in these respective groups of individuals toward increasing vaccine uptake.

Fourth, degree of trust in official entities was identified as an important predictor of hesitancy. Here, both lower levels of governmental trust in general, and less trust in the information provided by health-care officials about vaccination specifically, predicted vaccination hesitance, highlighting the important role of systemic trust in fostering large-scale behavior change (Fieselmann et al., 2022; Pertwee et al., 2022). One strategy that has previously been identified as effective in building trust includes engaging with subgroups through community leaders (Bavel et al., 2020; Burgess et al., 2021; Ojikutu et al., 2021), in addition to setting up arenas of dialogue allowing for open discussion and addressal of raised concerns (Burgess et al., 2021).

The fifth domain of psychological processes strongly tied to vaccine hesitancy concerned minimization and denial of COVID-19 as a societal problem (Aw et al., 2021; Troiano & Nardi, 2021). While this highlights a general need to combat pandemic-related myths, a recent randomized controlled trial (RCT) identified that information-based interventions emphasizing personal benefits of vaccination had a greater effect in changing hesitance than information provision pertaining to the societal level, such as the collective benefit of vaccination (Freeman et al., 2021). Combined with the identified predictors of hesitancy in the present study, these results imply that directing public health information toward individually-oriented dimensions such as the illusion of invulnerability against the virus, ways to address health-related fears tied to vaccination, and doubts concerning vaccination efficacy to protect oneself against infection, may prove most efficient in modifying hesitance and increasing vaccine uptake.

Beyond these predictors, a range of novel processes underlying hesitancy were identified through the analysis of open-ended responses. Several adults noted the media's impact on their vaccine hesitancy, pointing at cases in the mainstream media portraying individuals developing illnesses following vaccination. Given the known cognitive biases tied to vaccine hesitancy which can easily be amplified by media presentations, including availability bias where more vivid and emotionally memorable events are attributed higher weight and are more easily recalled (Azarpanah et al., 2021), this finding highlights the critical role and responsibility of the media as mediators of the public's perception of vaccination. Moreover, important nuances emerged for the unwillingness to vaccinate among certain individuals, including a general fear of needles, negative experiences with previous vaccinations, and aversions to vaccination related to underlying illness (e.g. eating disorder). This finding suggests the importance of screening for underlying psychiatric conditions and fears which may relate to hesitance, including fear of needles, with a range of effective

psychological interventions available for such phobias (e.g., Feitosa et al., 2013; Mackereth et al., 2012).

The open-ended analysis also identified a hesitant subgroup of adults perceiving disease, infection, and death as part of "the course of nature". Moreover, adults reported being overwhelmed by the hassle and overload of information around vaccination as a reason for their hesitance, highlighting the importance of identifying the appropriate dosage of information delivery in public health interventions (Honora et al., 2022; Voils et al., 2014). Notably, some individuals reported polarizing and stigmatizing portrayals of unvaccinated individuals as a key reason for becoming increasingly unwilling to get vaccinated. These findings highlight how stigmatizing portrayals can create an unfavorable “us vs. them” mentality, alienating unvaccinated individuals and making them more resistant to vaccination campaigns, underscoring the necessity for the media, governments, and healthcare communicators to avoid polarizing presentations (Marhánková et al., 2024).

Strengths and Limitations

The present study includes several strengths, including its mixed methods approach using two independent set of researchers who were blind to the investigation procedures and analytical results of one-another, the large-scale investigation of a topic of concurrent and future relevance, and replication of the results in a large independent sample from another nation. This replication and the high prediction accuracy across independent samples further corroborates the identification of central mechanisms related to vaccine hesitancy pertaining to the adult population. A notable strength of this study includes that the identified processes related to hesitancy were all factors that are loanable to manipulation by public health interventions. Unlike demographic risk factors such as one’s age group, these modifiable risk factors and processes can be actively modified through public health campaigns, providing key actionable steps for governments and health officials on the route to combating this

health-deteriorating global problem. Moreover, in addition to open documentation and code of the full workflow for the machine learning model, we adhered to open science and transparency recommendations in qualitative research (Levitt et al., 2018) by including an audit trail of the thematic analysis to document the analytical process in a step-by-step manner.

This study also includes several limitations. First, the recruitment procedure could have led to self-selection of specific subgroups above others, such as older adults with greater computer use experience. Additional efforts were undertaken to reduce this bias by also recruiting through platforms more accessible through the elderly population. The Norwegian sample was skewed toward oversampling of females and individuals with higher education. Our results were nonetheless replicated in an independent representative sample from the United Kingdom, highlighting the limited impact of this skewed sample on our results. The sole use of self-report measures is another limitation, calling for future studies based on electronic health records and objective vaccination records. Finally, while we investigated out-of-sample replication in an independent sample from another European country, the extent to which these results can be replicated in non-Western cultures and low- and middle-income countries remains to be investigated (Kola et al., 2021).

Concluding Remarks

This study identified modifiable psychological processes underlying vaccine hesitance which could function as targets in evidence-based health literacy campaigns aiming to address misconceptions and fears related to vaccination toward increasing vaccine uptake. Investigating adults across Norway and the United Kingdom, we identified five domains of psychological processes predicting vaccine hesitancy, including illusion of invulnerability, vaccine efficacy doubts, mistrust in health officials, minimization of COVID-19's societal significance, and vaccine-related health fears. Additionally, the media's portrayal of rare

cases amplifying fear, in addition to, stigmatization of unvaccinated individuals, were highlighted as other underlying motives leading to vaccine reluctance and polarization. These findings, consistent across two independent European samples, underscore the importance of addressing these modifiable factors in targeted public health campaigns to mitigate vaccine hesitance.

Acknowledgements

We wish to thank our participants in the MAP-19 study for their valuable time and efforts in contributing to more than two years of data collection. We further wish to express our gratitude to the investigators of the COVID-19 Psychological Research Consortium for their major service to the scientific community in providing an open and accessible dataset on a critical topic of investigation.

Author Contributions

OVE: Conceptualization, investigation, project administration, funding acquisition, methodology, data curation, formal analysis, visualization, writing - original draft, writing - review and editing. EMS: Methodology, formal analysis, writing - original draft, writing - review and editing. SMM: Methodology, formal analysis, writing - review and editing. SUJ: Investigation, project administration, funding acquisition, writing - review and editing. AH: Investigation, writing - review and editing. SB: Writing - review and editing. OAS: Writing - review and editing. LTW: Writing - review and editing. EHL: Conceptualization, methodology, data curation, formal analysis, visualization, writing - original draft, writing - review and editing.

Conflict of Interest

The author(s) declare no conflicts of interest with respect to the authorship or the publication of this article.

Open Science Procedures and Code Availability Statement

All code and supplementary material necessary to inspect the conducted analysis is uploaded at the online repository of the Center for Open Science (https://osf.io/nbrwz/), including the full workflow of the machine learning algorithm along with open code to reproduce the analysis. All plots in the code are in readily viewable form upon opening the code files in a matching coding environment (e.g., Visual Studio Code, Jupyter Notebook). An audit trail documenting the qualitative analysis in a transparent step-by-step manner is provided in Supplementary Document 2.

Data Availability Statement

The UK data is openly available and can be found at the online repository of the Center for Open Science (https://osf.io/v2zur/). For the Norwegian data, the received ethical approval from the Norwegian Centre for Research Data (NSD) precludes submission of raw data to public repositories. Access to the data can be granted from the principal investigators Omid V. Ebrahimi and Sverre Urnes Johnson following ethical approval of a suggested project plan for the use of data granted by NSD and REK.

Figure Captions and Supplementary Documents

Figure 1

Predictive performance of the best model in the test fold of the Norwegian sample. ROC curve: Receiver Operating Characteristic curve; AUC: Area under the ROC curve; t: Classification threshold.

Figure 2

Role of the key predictors in the model predicting vaccine hesitancy in the Norwegian sample. (a) Variable importances were calculated based on information gain in the training set. (b) The distribution for each predictor is visualized based on the hold-out test set. The first nine

variables in Figure 2b are ordinal variables, ranging from 0-5 (Completely Disagree to Completely Agree for variable 1-3 and 5-8; and Not at all probable to Highly probable for variable 4; and Not at all serious to Very serious for variable 9), whereas variable 10 was continuous, ranging from 0 to 100 (Not risky at all to Maximally risky).

Figure 3

Results of the thematic analysis including the main themes and two levels of sub-themes identified through analysis of the participants who also provided open-ended (n = 52) explanations underlying their vaccine hesitance.

Supplementary Figure S1

Role of the key predictors for predicting vaccine hesitancy in the UK sample. (a) The variable importances were calculated based on information gain in the training set. (b) The distribution for each predictor is visualized based on the hold-out test set. All top ten predictors were continuous, ranging from 0 to 100 (Completely disagree to Completely agree for variables 1-2, 6 and 9-10; Not at all to Completely for variables 3-5 and 7; and Not at all threatened to Extremely threatened for variable 8).

Supplementary Table S1

Variables used in the data splitting procedure.

Supplementary Table S2

Hyperparameters and the values tested during tuning. The highlighted values are those that produced the best performing model.

Supplementary Table S3

Summary of the distribution of the responses for the 10 most important variables for the vaccine hesitant and willing group in the independent test set.

Supplementary Table S4

The variables used in the independent sample replication in the UK dataset along with their corresponding variables in the Norwegian sample.

Supplementary Table S5

Proportion of vaccine hesitant individuals (n = 248 in Norway and n = 109 in UK) within different demographic subgroups.

Supplementary Document 1

List of variables used in the machine learning model.

Supplementary Document 2

Audit trail of the analytic procedure in the thematic analysis.

References

Anderson, R. M., Heesterbeek, H., Klinkenberg, D., & Hollingsworth, T. D. (2020). How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet, 395*(10228), 931–934.

Anderson, R. M., Vegvari, C., Truscott, J., & Collyer, B. S. (2020). Challenges in creating herd immunity to SARS-CoV-2 infection by mass vaccination. *The Lancet, 396*(10263), 1614–1616.

Andre, F. E., Booy, R., Bock, H. L., Clemens, J., Datta, S. K., John, T. J., Lee, B. W., Lolekha, S., Peltola, H., Ruff, T., et al. (2008). Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bulletin of the World Health Organization, 86*, 140–146.

Aw, J., Seng, J. J. B., Seah, S. S. Y., & Low, L. L. (2021). COVID-19 vaccine hesitancy – a scoping review of literature in high-income countries. *Vaccines, 9*(8), 900.

Azarpanah, H., Farhadloo, M., Vahidov, R., & Pilote, L. (2021). Vaccine hesitancy: Evidence from an adverse events following immunization database, and the role of cognitive biases. *BMC Public Health, 21*(1), 1–13.

Bavel, J. J. V., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M. J., Crum, A. J., Douglas, K. M., Druckman, J. N., et al. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour, 4*(5), 460–471.

Bovin, M. J., Marx, B. P., Weathers, F. W., Gallagher, M. W., Rodriguez, P., Schnurr, P. P., & Keane, T. M. (2016). Psychometric properties of the PTSD checklist for diagnostic and statistical manual of mental disorders–fifth edition (PCL-5) in veterans. *Psychological Assessment, 28*(11), 1379.

Braitman, A. L., Strowger, M., Shipley, J. L., Ortman, J., MacIntyre, R. I., & Bauer, E. A. (2022). Data Quality and Study Compliance Among College Students Across 2 Recruitment Sources: Two Study Investigation. *JMIR Formative Research, 6*(12), e39488.

Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology, 9*(1), 3.

Brooks, J., McCluskey, S., Turley, E., & King, N. (2015). The utility of template analysis in qualitative psychology research. *Qualitative research in psychology, 12*(2), 202–222.

Burgess, R. A., Osborne, R. H., Yongabi, K. A., Greenhalgh, T., Gurdasani, D., Kang, G., Falade, A. G., Odone, A., Busse, R., Martin-Moreno, J. M., et al. (2021). The COVID-19 vaccines rush: Participatory community engagement matters more than ever. *The Lancet, 397*(10268), 8–10.

Bzdok, H., Altman, N., & Krzywinski, M. (2018). Statistics versus machine learning. *Nature Methods, 15*(4), 233.

Calvert, C. L., & Khoshgoftaar, T. M. (2019). Threshold based optimization of performance metrics with severely imbalanced big security data. *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, 1328–1334.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.

Chowdhury, M. Z. I., & Turin, T. C. (2020). Variable selection strategies and its importance in clinical prediction modelling. *Family Medicine and Community Health, 8*(1).

Davis, C. J., Golding, M., & McKay, R. (2022). Efficacy information influences intention to take COVID-19 vaccine. *British Journal of Health Psychology, 27*(2), 300–319.

Dinga, R., Penninx, B. W., Veltman, D. J., Schmaal, L., & Marquand, A. F. (2019). *Beyond accuracy: Measures for assessing machine learning models, pitfalls and guidelines*. bioRxiv.

Dubé, È., Ward, J. K., Verger, P., & MacDonald, N. E. (2021). Vaccine hesitancy, acceptance, and anti-vaccination: Trends and future prospects for public health. *Annual Review of Public Health, 42*(1), 175–91.

Ebrahimi, O. V., Bauer, D. J., Hoffart, A., & Johnson, S. U. (2022). A critical period for pandemic adaptation: The evolution of depressive symptomatology in a representative sample of adults across a 17-month period during COVID-19. *Journal of Psychopathology and Clinical Science, 131*(8), 881.

Ebrahimi, O. V., Hoffart, A., & Johnson, S. U. (2021). Physical distancing and mental health during the COVID-19 pandemic: Factors associated with psychological symptoms and adherence to pandemic mitigation strategies. *Clinical Psychological Science, 9*(3), 489–506.

Ebrahimi, O. V., Hoffart, A., & Johnson, S. U. (2023). Viral mitigation and the COVID-19 pandemic: Factors associated with adherence to social distancing protocols and hygienic behaviour. *Psychology & Health, 38*(3), 283–306.

Feitosa, A. C. R., Sampaio, L. N., Batista, A. G. L., & Pinheiro, C. B. (2013). Frequency of fear of needles and impact of a multidisciplinary educational approach towards pregnant women with diabetes. *Revista Brasileira de Ginecologia e Obstetrıcia, 35*, 111–116.

Fieselmann, J., Annac, K., Erdsiek, F., Yilmaz-Aslan, Y., & Brzoska, P. (2022). What are the reasons for refusing a covid-19 vaccine? A qualitative analysis of social media in Germany. *BMC Public Health, 22*(1), 1–8.

Freeman, D., Loe, B. S., Yu, L.-M., Freeman, J., Chadwick, A., Vaccari, C., Shanyinde, M., Harris, V., Waite, F., Rosebrock, L., et al. (2021). Effects of different types of written vaccination information on COVID-19 vaccine hesitancy in the UK (OCEANS-III): A single-blind, parallel-group, randomised controlled trial. *The Lancet Public Health, 6*(6), e416–e427.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. The *Annals of Statistics, 29*(5), 1189–1232.

Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology, 13*(1), 1–8.

Haas, E. J., McLaughlin, J. M., Khan, F., Angulo, F. J., Anis, E., Lipsitch, M., ... & Alroy-Preis, S. (2022). Infections, hospitalisations, and deaths averted via a nationwide vaccination campaign using the Pfizer–BioNTech BNT162b2 mRNA COVID-19 vaccine in Israel: a retrospective surveillance study. *The Lancet Infectious Diseases, 22*(3), 357-366.

Honora, A., Wang, K.-Y., & Chih, W.-H. (2022). How does information overload about COVID-19 vaccines influence individuals' vaccination intentions? The roles of cyberchondria, perceived risk, and vaccine skepticism. *Computers in Human Behavior, 130*, 107176.

Khairat, S., Zou, B., & Adler-Milstein, J. (2022). Factors and reasons associated with low COVID-19 vaccine uptake among highly hesitant communities in the us. *American Journal of Infection Control, 50*(3), 262–267.

King, N. (2012). Doing template analysis. *Qualitative Organizational Research: Core Methods and Current Challenges, 426*, 426–450.

Kola, L., Kohrt, B. A., Hanlon, C., Naslund, J. A., Sikander, S., Balaji, M., Benjet, C., Cheung, E. Y. L., Eaton, J., Gonsalves, P., et al. (2021). COVID-19 mental health impact and responses in low-income and middle-income countries: Reimagining global mental health. *The Lancet Psychiatry, 8*(6), 535–550.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613.

Lazarus, J. V., Ratzan, S. C., Palayew, A., Gostin, L. O., Larson, H. J., Rabin, K., Kimball, S., & El-Mohandes, A. (2021). A global survey of potential acceptance of a COVID-19 vaccine. *Nature Medicine, 27*(2), 225–228.

Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal Article Reporting Standards for Qualitative Primary, Qualitative Meta-Analytic, and Mixed Methods Research in Psychology: The APA Publications and Communications Board Task Force Report. *American Psychologist, 73*(1), 26.

Lincoln, T. M., Schlier, B., Strakeljahn, F., Gaudiano, B. A., So, S. H., Kingston, J., Morris, E. M., & Ellett, L. (2022). Taking a machine learning approach to optimize prediction of vaccine hesitancy in high income countries. *Scientific Reports, 12*(1), 1–12.

MacDonald, N. E. et al. (2015). Vaccine hesitancy: Definition, scope and determinants. *Vaccine, 33*(34), 4161–4164.

Mackereth, P., Hackman, E., Tomlinson, L., Manifold, J., & Orrett, L. (2012). 'Needle with ease': Rapid stress management techniques. *British Journal of Nursing, 21*(Sup14), S18–S22.

Magnúsdóttir, I., Lovik, A., Unnarsdóttir, A. B., McCartney, D., Ask, H., Kõiv, K., Christoffersen, L. A. N., Johnson, S. U., Hauksdóttir, A., Fawns-Ritchie, C., et al. (2022). Acute COVID-19 severity and mental health morbidity trajectories in patient

populations of six nations: An observational study. *The Lancet Public Health, 7*(5), e406–e416.

Marhánková, J. H., Kotherová, Z., & Numerato, D. (2024). Navigating vaccine hesitancy: Strategies and dynamics in healthcare professional-parent communication. *Human Vaccines & Immunotherapeutics*, *20*(1), 2361943.

McBride, O., Butter, S., Hartman, T. K., Murphy, J., Hyland, P., Shevlin, M., Gibson-Miller, J., Levita, L., Mason, L., Martinez, A. P., et al. (2020). Sharing data to better understand one of the world's most significant shared experiences: Data resource profile of the longitudinal COVID-19 psychological research consortium (C19PRC) study. *International Journal of Population Data Science, 5*(4).

Mukumbang, F. C. (2023). Retroductive theorizing: A contribution of critical realism to mixed methods research. *Journal of Mixed Methods Research, 17*(1), 93–114.

Neumann-Böhme, S., Varghese, N. E., Sabat, I., Barros, P. P., Brouwer, W., van Exel, J., Schreyögg, J., & Stargardt, T. (2020). Once we have it, will we use it? A european survey on willingness to be vaccinated against COVID-19. *The European Journal of Health Economics, 21*, 977-982.

Norwegian Institute of Public Health. (2021). Public health after COVID-19. in: Public health report – October 2021. Retrieved June 18, 2024, from https://www.fhi.no/ss/korona/koronavirus/folkehelserapporten-temautgave-2021/del-1-9/tillit-og-vaksinasjon-i-norge

Norwegian Institute of Public Health. (2016). Mental illness among adults in Norway. in: Public health report - health status in Norway. Retrieved November 1, 2023, from https://www.fhi.no/en/op/hin/mental-health/psykisk-helse-hos-voksne/

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future—big data, machine learning, and clinical medicine. *The New England Journal of Medicine, 375*(13), 1216.

Obregon, R., Mosquera, M., Tomsa, S., & Chitnis, K. (2020). Vaccine hesitancy and demand for immunization in Eastern Europe and Central Asia: implications for the region and beyond. *Journal of Health Communication, 25*(10), 808–815.

Ojikutu, B. O., Stephenson, K. E., Mayer, K. H., & Emmons, K. M. (2021). Building trust in COVID-19 vaccines and beyond through authentic community investment. *American Journal of Public Health, 111*(3), 366–368.

Pallesen, S., Bjorvatn, B., Nordhus, I. H., Sivertsen, B., Hjørnevik, M., & Morin, C. M. (2008). A new scale for measuring insomnia: The Bergen Insomnia Scale. *Perceptual and Motor Skills, 107*(3), 691–706.

Paul, E., Steptoe, A., & Fancourt, D. (2021). Attitudes towards vaccines and intention to vaccinate against COVID-19: Implications for public health communications. *The Lancet Regional Health–Europe, 1*.

Pertwee, E., Simas, C., & Larson, H. J. (2022). An epidemic of uncertainty: Rumors, conspiracy theories and vaccine hesitancy. *Nature Medicine, 28*(3), 456–459.

Phillips, R., Gillespie, D., Hallingberg, B., Evans, J., Taiyari, K., Torrens-Burton, A., Cannings-John, R., Williams, D., Sheils, E., Ashfield-Watt, P., et al. (2022). Perceived threat of COVID-19, attitudes towards vaccination, and vaccine hesitancy: A prospective longitudinal study in the UK. *British Journal of Health Psychology, 27*(4), 1354–1381.

Poudel, A. N., Zhu, S., Cooper, N., Roderick, P., Alwan, N., Tarrant, C., Ziauddeen, N., & Yao, G. L. (2021). Impact of Covid-19 on health-related quality of life of patients: A structured review. *PloS one, 16*(10), e0259164.

Power, S. A., Velez, G., Qadafi, A., & Tennant, J. (2018). The SAGE model of social psychological research. *Perspectives on Psychological Science, 13*(3), 359–372.

Provost, F. (2008). Machine learning from imbalanced data sets 101. *AAAI Technical Report, WS-00-05*.

Ryan, J., & Malinga, T. (2021). Interventions for vaccine hesitancy. *Current Opinion in Immunology*, 71, 89–91.

Sachs, J. D., Karim, S. S. A., Aknin, L., Allen, J., Brosbøl, K., Colombo, F., Barron, G. C., Espinosa, M. F., Gaspar, V., Gaviria, A., et al. (2022). The lancet commission on lessons for the future from the COVID-19 pandemic. *The Lancet, 400*(10359), 1224–1280.

Salali, G. D., & Uysal, M. S. (2020). COVID-19 vaccine hesitancy is associated with beliefs on the origin of the novel coronavirus in the UK and Turkey. *Psychological Medicine*, 1–3.

Shen, A. K. (2020). Finding a way to address a wicked problem: Vaccines, vaccination, and a shared understanding. *Human Vaccines & Immunotherapeutics, 16*(5), 1030–1033.

Shen, Q., Joyce, E. E., Ebrahimi, O. V., Didriksen, M., Lovik, A., Sævarsdóttir, K. S., Magnúsdóttir, I., Mikkelsen, D. H., Unnarsdóttir, A. B., Hauksdóttir, A., et al. (2023). COVID-19 illness severity and 2-year prevalence of physical symptoms: An observational study in Iceland, Sweden, Norway and Denmark. *The Lancet Regional Health–Europe, 35*.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092–1097.

Steinmetz, L. (2022). Sociodemographic predictors of and main reasons for COVID-19 vaccine hesitancy in eastern oslo: A cross-sectional study. *BMC Public Health, 22*(1), 1–9.

Steltenpohl, C. N., Lustick, H., Meyer, M. S., Lee, L. E., Stegenga, S. M., Reyes, L. S., & Renbarger, R. (2023). Rethinking transparency and rigor from a qualitative open science perspective. *Journal of Trial & Error*, online first.

Stenwig, E., Salvi, G., Rossi, P. S., & Skjærvold, N. K. (2022). Comparative analysis of explainable machine learning prediction models for hospital mortality. *BMC Medical Research Methodology, 22*(1), 1–14.

Tenforde, M. W., Kim, S. S., Lindsell, C. J., Rose, E. B., Shapiro, N. I., Files, D. C., Gibbs, K. W., Erickson, H. L., Steingrub, J. S., Smithline, H. A., et al. (2020). Symptom duration and risk factors for delayed return to usual health among outpatients with COVID -19 in a multistate health care systems network—United States, March–June 2020. *Morbidity and Mortality Weekly Report, 69*(30), 993.

Trogen, B., & Pirofski, L.-a. (2021). Understanding vaccine hesitancy in covid-19. *Med, 2*(5), 498–501.

Troiano, G., & Nardi, A. (2021). Vaccine hesitancy in the era of COVID-19. *Public Health, 194*, 245–251.

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics, 7*, 1-8.

Varoquaux, G., & Colliot, O. (2023). Evaluating machine learning models and their diagnostic value. In O. Colliot (Ed.), *Machine learning for brain disorders* (pp. 601–630). Springer US.

Voils, C. I., King, H. A., Maciejewski, M. L., Allen, K. D., Yancy Jr, W. S., & Shaffer, J. A. (2014). Approaches for informing optimal dose of behavioral interventions. *Annals of Behavioral Medicine, 48*(3), 392–401.

WHO. (2019). *Ten threats to global health in 2019*. Retrieved November 22, 2023, from https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019

Wiltshire, G., & Ronkainen, N. (2021). A realist approach to thematic analysis: Making sense of qualitative data through experiential, inferential and dispositional themes. *Journal of Critical Realism, 20*(2), 159–180.

Wiysonge, C. S., Ndwandwe, D., Ryan, J., Jaca, A., Batouré, O., Anya, B.-P. M., & Cooper, S. (2022). Vaccine hesitancy in the era of COVID-19: could lessons from the past help in divining the future? *Human Vaccines & Immunotherapeutics, 18*(1), 1–3.

Wong, J. C., Lao, C. T., Yousif, M. M., & Luga, J. M. (2022). Fast tracking—vaccine safety, efficacy, and lessons learned: A narrative review. *Vaccines, 10*(8), 1256.

Zhang, H., Li, Y., Peng, S., Jiang, Y., Jin, H., & Zhang, F. (2022). The effect of health literacy on COVID-19 vaccine hesitancy among community population in China: The moderating role of stress. *Vaccine, 40*(32), 4473–4478.

Table 1. Classification metrics for the best performing model in the independent test set of the Norwegian sample and the test portion of the independent sample from the United Kingdom.

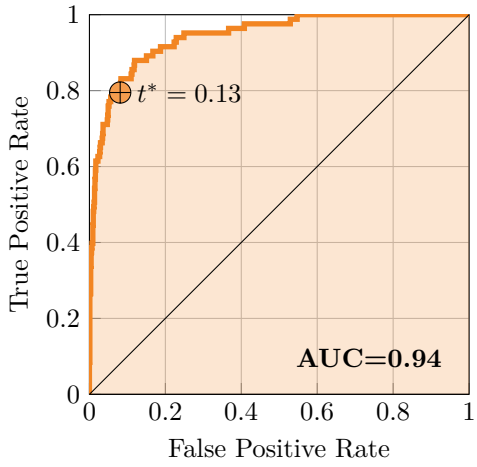
	Default classification threshold	Optimal classification threshold
	Norwegian sample	
AUC	0.94	
Brier score	0.04	
Accuracy	94.77%	90.87%
Balanced accuracy	75.83%	86.27%
Recall (sensitivity)	0.53	0.81
Precision (PPV)	0.79	0.48
Specificity	0.96	0.98
	UK sample	
AUC	0.98	
Brier score	0.04	
Accuracy	94.69%	94.28%
Balanced accuracy	89.09%	89.96%
Recall (sensitivity)	0.81	0.83
Precision (PPV)	0.83	0.79
Specificity	0.96	0.97

Table 2. Thematic map including the identified main and submotives underlying vaccine hesitancy along with illustrative quotes from the vaccine hesitant individuals.

Theme	Subtheme	Example extract	*n*
Theme 1: Unnecessary vaccines (*n*=17)	1.1 Covid-19 in general not dangerous	«(...) Just the concept of walking around being afraid of a cold/flu is absurd» (#35, Woman, 34 years)	12
	1.2 Perceiving oneself as not personally being at risk	«The recommendations for taking the vaccine are based on the average Norwegian. I am younger than the average and much more fit (...)» (#48, woman, 33 years)	9
Theme 2: Inefficient vaccines (*n*=19)	2.1 Does not offer much protection	«It doesn't seem like the vaccine has such a good effect, and mutations haven't been taken into account, as evidenced by the new lockdown.» (#1, man, 20 years)	11
	2.2 Does not hinder transmission	«A vaccine that doesn't help against transmission, it's too early to say anything about side-effects.» (#14, woman, 45 years)	7
	2.3 Disqualifying these as *real* vaccines	«These are NOT vaccines. A vaccine gives immunity and hinders transferal. What is being injected into people does neither and it is misleading people to call it vaccine.» (#13, woman, 50 years)	4
Theme 3: Frightening vaccines (*n*=32)	3.1 Fear of side-effects	«I'm skeptical about long-term side-effects that may arise.» (#36, woman, 33 years)	26
	3.2 Fear of not getting healthcare follow-up if needed	«I don't have confidence that I'll receive help from the healthcare system if I experience vaccine side-effects, and I consider it likely that I will get side-effects.» (#43, woman, 26 years)	3
	3.3 Vaccine approvals too quick	«I'm skeptical to that the vaccine had not been tested but was approved without further evaluation.» (#9, woman, 45 years)	8

Theme 4: Distrust and polarization (*n*=18)	4.1 Authorities' vaccine handling	«I'm losing trust in everything politicians say because I know they withhold information when looking at what's happening abroad.» (#31, woman, 40 years)	10
	4.2 Pharmaceutical industry	« (...) It's especially bad that this is one of the most lucrative things for pharmaceutical companies.» (#38, woman, 32 years)	3
	4.3 Mainstream media's vaccine presentation	«I read other sources than VG and Dagbladet [mainstream media] and see how many injuries this so-called 'vaccine' is causing people." (#5) (woman, 22 years)	7
	4.4 Portrayal of the unvaccinated	«Public shaming of the unvaccinated leads to a greater degree of reluctance.» (#40, man, 64 years)	6
Theme 5: Alternatives to getting vaccinated (*n*=15)	5.1 Natural immunity	«Natural immunity has always been the only realistic way out of the pandemic phase.» (#4, woman, 46 years)	4
	5.2 The course of nature	«In the animal world, disease and infection occur when there are too many in the herd, and nature's way is to sort it out on its own (...). If it's meant that my time has come, it will come.» (#8, man, 64 years)	5
	5.3 Own health measures	«I want to belong to the control group, emphasizing a strong immune system (vitamin D, vitamin C, and zinc).» (#40, man, 64 years)	2
	5.4 Avoid infections	«I easily have the opportunity to choose or avoid places where the infection is highest.» (#44) (Woman, 66 years)	3
	5.5 Avoidance as easier than taking a stance	«I can't deal with all the 'nagging'; the more 'hassle' there is around me, the less I can handle it. So for me, it's easiest (least	2

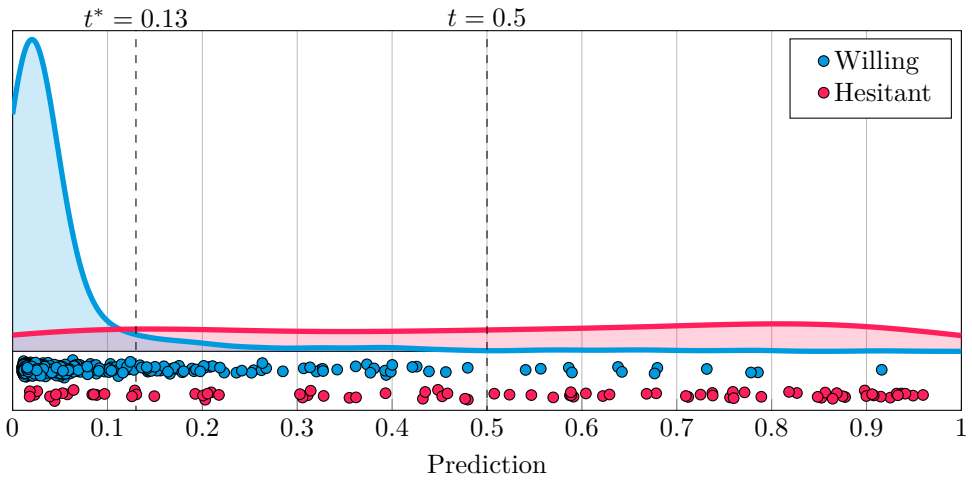
		psychologically painful) to just do nothing.» (#2, woman, 32 years)	
Theme 6: Reasons related to vaccines in general (*n*=12)	6.1 Fear of needles	«I have a fear of needles and don't dare to get vaccinated.» (#6, woman, 36 years)	4
	6.2 Aversion related to own underlying disease	«It's related to eating disorder issues. I can't handle putting something into my body that I can't control.» (#34, woman, years anonymized)	4
	6.3 Negative experiences with previous vaccination	«I have experienced poor health after the swine flu vaccine, I'm afraid, and don't want to risk worse health due to another vaccine.» (#50, woman, years anonymized)	4



(a) ROC curve

		Predicted	
		0	1
Observed	0	819	73
	1	16	67

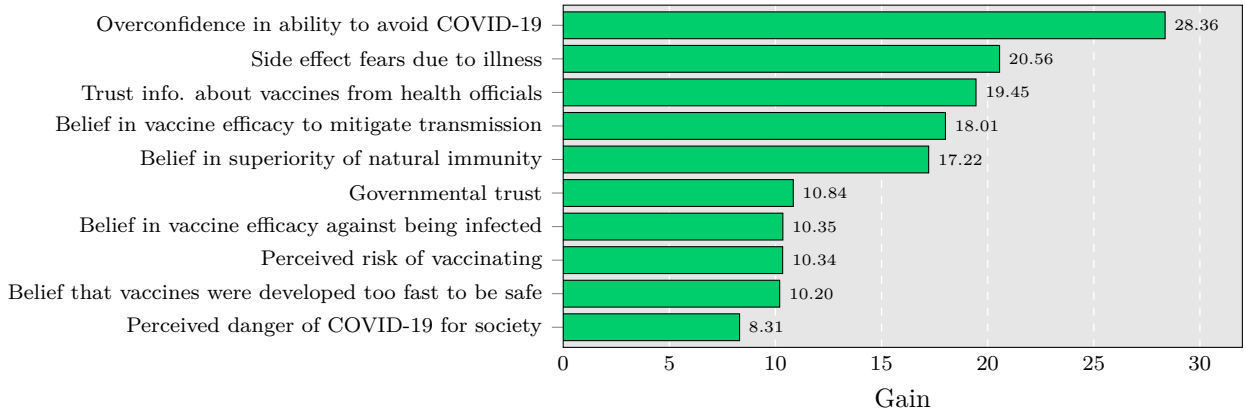
(b) Confusion matrix



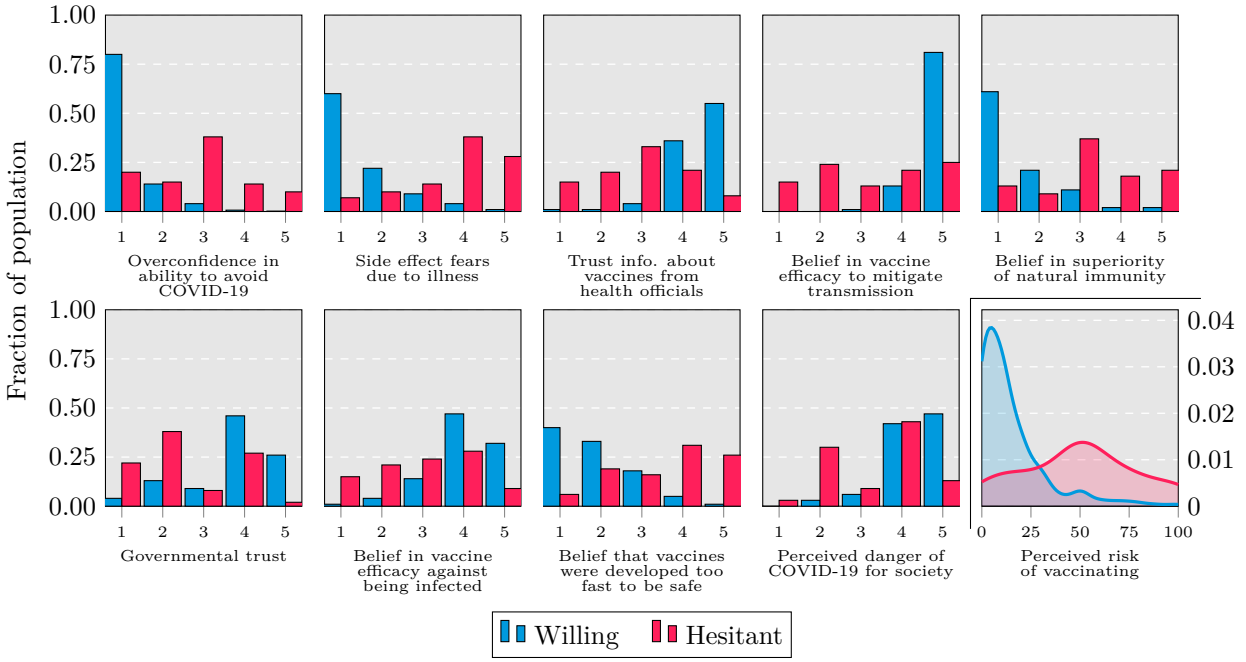
(c) Out-of-sample predictions

Figure 1

Predictive performance of the best model in the test fold of the Norwegian sample. ROC curve: Receiver Operating Characteristic curve; AUC: Area under the ROC curve; t: Classification threshold.



(a) Ten most important variables and their contribution to the model



(b) Distributions of responses for the top 10 predictors in the two groups

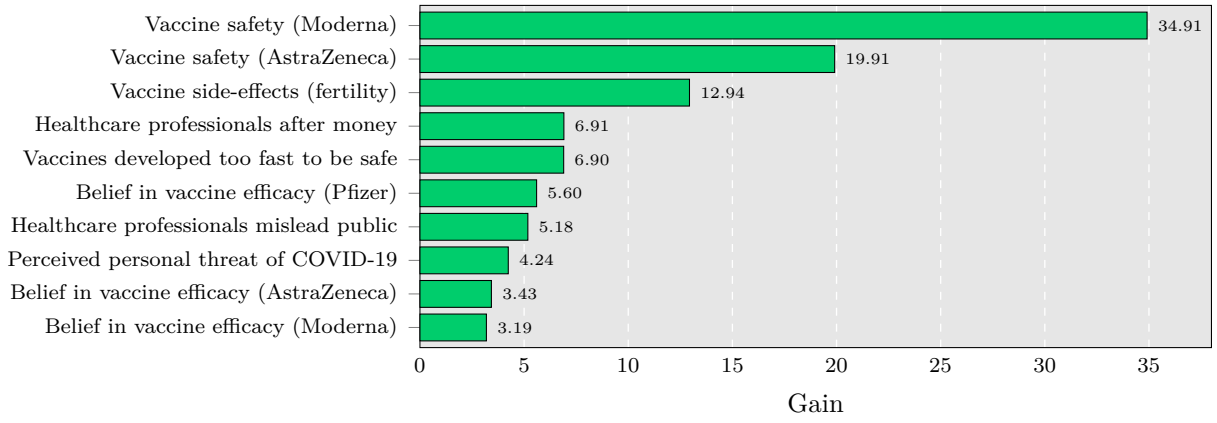
Figure 2

Role of the key predictors in the model predicting vaccine hesitancy in the Norwegian sample. (a) Variable importances were calculated based on information gain in the training set. (b) The distribution for each predictor is visualized based on the hold-out test set. The first nine variables in Figure 2b are ordinal variables, ranging from 0-5 (Completely Disagree to Completely Agree for variable 1-3 and 5-8; and Not at all probable to Highly probable for variable 4; and Not at all serious to Very serious for variable 9), whereas variable 10 was continuous, ranging from 0 to 100 (Not risky at all to Maximally risky).

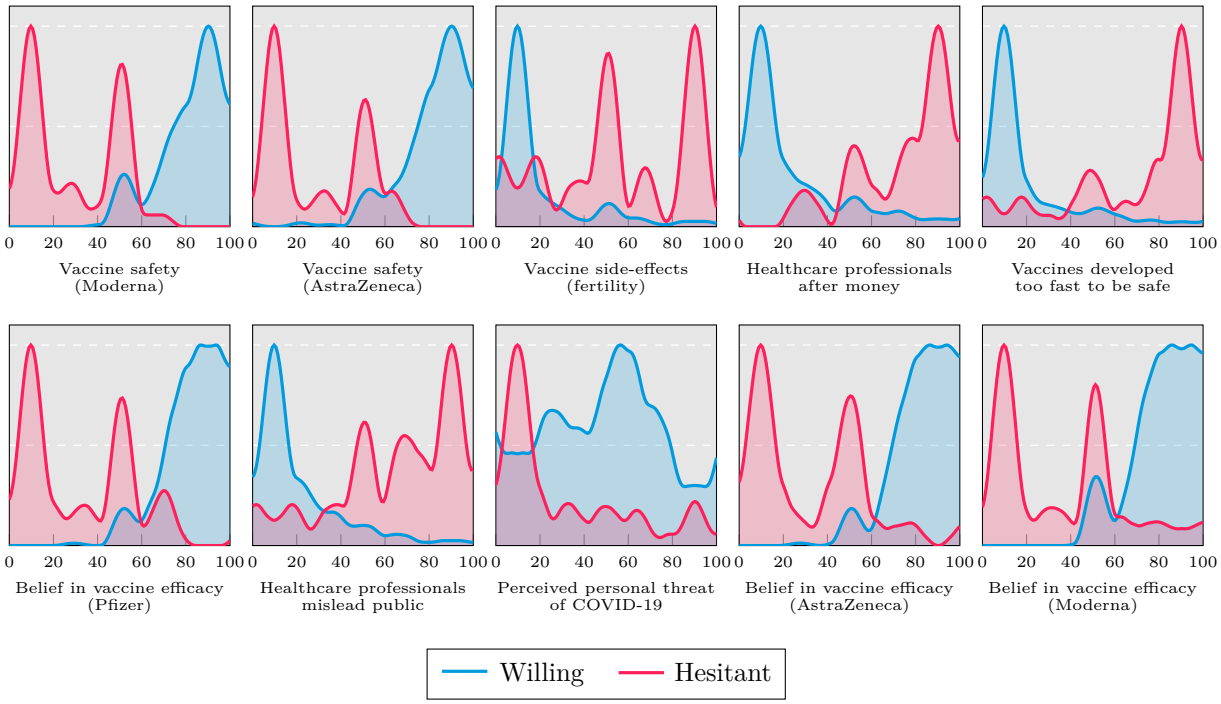
- Vaccine hesitant (n = 52)
  - Unnecessary vaccines (n = 17)
    - COVID-19 in general not dangerous (n = 12)
      - Not at all dangerous (n = 3)
      - Not as dangerous as it is portrayed to be (n = 7)
      - Participants or acquaintances have only had mild symptoms (n = 3)
    - Perceiving oneself as not personally being at risk (n = 9)
  - Distrust and polarization (n = 18)
    - Authorities' vaccine handling (n = 10)
      - Not providing sufficient and/or correct information (n = 7)
      - Distrusting authorities' vaccine handling (n = 7)
    - Pharmaceutical industry (n = 3)
      - Do not want to support pharmaceutical industry (n = 1)
      - Distrusting pharmaceutical industry (n = 3)
    - Portrayal of the unvaccinated (n = 6)
      - Polarizing portrayals contributing to unwillingness (n = 3)
      - Pressure to vaccinate (n = 3)
    - Mainstream media's vaccine presentation (n = 7)
      - Misleading (n = 3)
      - Refers to vaccine critical sources (n = 6)
  - Reasons related to vaccines in general (n = 12)
    - Fear of needles (n = 4)
    - Fear related to underlying disease (n = 4)
    - Negative experiences with previous vaccination or medication (n = 4)
  - Frightening vaccines (n = 32)
    - Fear of not getting healthcare follow-up if needed (n = 3)
    - Vaccine approval too quick (n = 8)
    - Fear of side-effects (n = 26)
      - Generally fearing side-effects (n = 12)
      - Uncertain about possible long-term side-effects (n = 7)
      - Side-effects more risky than COVID-19 (n = 6)
      - Personally know others that experienced side-effects (n = 4)
      - Worry about menstruation changes (n = 4)
      - Pregnancy or fertility reasons (n = 3)
  - Inefficient vaccines (n = 19)
    - Does not hinder transmission (n = 7)
    - Disqualifying these as *real* vaccines (n = 4)
    - Does not offer much protection (n = 11)
      - Does not have a good effect (n = 5)
      - Effect does not last (n = 6)
  - Alternatives to getting vaccinated (n = 15)
    - Avoidance easier than taking a stance (n = 2)
    - Own health measures (n = 2)
      - Strengthen immune system (n = 1)
      - Staying physically fit (n = 1)
    - Avoid infections (n = 3)
      - Avoid places where one can get infected (n = 2)
      - Avoid people in vulnerable groups (n = 1)
    - Natural immunity (n = 4)
      - Natural herd immunity as the solution (n = 1)
      - Natural immunity better than vaccine (n = 3)
    - The course of nature (n = 5)
      - Own potential death is natural (n = 3)
      - Natural for old people to pass away (n = 1)
      - Pandemic is nature's way of handling overpopulation and misuse (n = 2)

Figure 3

Results of the thematic analysis including the main themes and two levels of sub-themes identified through analysis of the participants who also provided open-ended (n = 52) explanations underlying their vaccine hesitance.



(a) *Ten most important variables and their contribution to the model*



(b) *Distribution of responses for the top ten predictors in the two groups*

Supplementary Figure S1

Role of the key predictors for predicting vaccine hesitancy in the UK sample. (a) The variable importances were calculated based on information gain in the training set. (b) The distribution for each predictor is visualized based on the hold-out test set. All top ten predictors were continuous, ranging from 0 to 100 (Completely disagree to Completely agree for variables 1-2, 6 and 9-10; Not at all to Completely for variables 3-5 and 7; and Not at all threatened to Extremely threatened for variable 8).

Supplementary Tables S1 – S5

Supplementary Table S1. Variables used in the data splitting procedure.

Variable name	Description	Coding
VaccineHesitancy	Planning to vaccinate against COVID-19 upon availability of a vaccine	Yes; No
Sex	Sex assigned at birth	Females: 0; Males: 1
AgeGroup	Age group	18-30 years; 31-44 years; 45-64 years; and 65 years and above
Education	Highest education to date	Compulsory school: 0; Upper secondary High school: 1; Student: 2; Any university degree: 3
PsychiatricDiagnosisBaseline	Psychiatric diagnosis at baseline	Yes; No
CityResidency	City (versus rural) residency	Yes; No
LivesWithChild	Lives with children (< 18 years)	Yes; No
NotWorking	Unemployed	Yes; No
GovernmentalTrust	I trust in the government and their decision making regarding the pandemic	Completely disagree: 1 to Completely agree: 5
TrustInfoAboutVaccinesHealthOfficials	I trust the information I receive from health officials about vaccines	Completely disagree: 1 to Completely agree: 5

Supplementary Table S2. Hyperparameters and the values tested during tuning. The highlighted values are those that produced the best performing model.

Parameter	**Values**
max_depth	{**3**, 6}
learning_rate	{0.01, **0.1**, 0.3}
n_estimators	{100, 250, **500**}
colsample_bytree	{**0.3**, 0.7}
gamma	{0, 1, **10**}

Supplementary Table S3. Summary of the distribution of the responses for the 10 most important variables for the vaccine hesitant and willing group in the independent test set.

		Willing		**Hesitant**	
Variable	**Type**	***mode***	***SD***	***mode***	***SD***
OverconfidenceInC19Avoidance	Ordinal	1	0.59	3	1.22
SideEffectFearsDueToIllness	Ordinal	1	0.96	4	1.19
TrustInfoAboutVaccinesHealthOfficials	Ordinal	5	0.74	3	1.16
BeliefVaccinationEfficacyMitigateTransmission	Ordinal	5	0.62	5	1.43
BeliefInSuperiorityOfNaturalImmunity	Ordinal	1	0.94	3	1.26
GovernmentalTrust	Ordinal	4	1.10	2	1.18
BeliefVaccinationEfficacyAgainstBeingInfected	Ordinal	4	0.86	4	1.23
PerceivedRiskOfVaccinating	Continuous	14 (mean)	17	48 (mean)	29
BeliefVaccinesDevelopedToofastToBeSafe	Ordinal	1	0.98	4	1.23
PerceivedDangerOfCOVIDForSociety	Ordinal	5	0.77	4	1.14

Supplementary Table S4. The variables used in the independent sample replication in the UK dataset along with their corresponding variables in the Norwegian sample.

Variable name in UK dataset	Description	Coding	Corresponding variable in the Norwegian dataset
COVID19_threat	Perceived personal threat of COVID-19	Not at all threatened: 0 to Extremely threatened: 100	OverconfidenceInC19Avoidance
Vax_conspiracy_4	Vaccines are not safe because they were rapidly developed and tested	Not at all: 0 to Completely: 100	BeliefVaccinesDevelopedToofastToBeSafe
Vax_conspiracy_5	Vaccines can damage fertility	Not at all: 0 to Completely: 100	SideEffectFearsDueToIllness
Trust_Body_1	How much trust do you have in the UK Parliament	Complete trust: 1 to Do not trust at all: 5	GovernmentalTrust
Trust_Body_2	How much trust do you have in the UK Government	Complete trust: 1 to Do not trust at all: 5	GovernmentalTrust
Trust_Body_6	How much trust do you have in your local government (Council or Local authority)	Complete trust: 1 to Do not trust at all: 5	GovernmentalTrust
Science_conspiracy_1	Healthcare professionals and scientists often deceive or mislead the public	Not at all: 0 to Completely: 100	TrustInfoAboutVaccinesHealthOfficials
Science_conspiracy_2	Healthcare professionals and scientists often cover up their mistakes	Not at all: 0 to Completely: 100	TrustInfoAboutVaccinesHealthOfficials
Science_conspiracy_3	Healthcare professionals and scientists are more concerned with making money than taking care of people	Not at all: 0 to Completely: 100	TrustInfoAboutVaccinesHealthOfficials
Science_conspiracy_4	Healthcare professionals and scientists don’t know what they are doing	Not at all: 0 to Completely: 100	TrustInfoAboutVaccinesHealthOfficials

Trust_Body_10	How much trust do you have in the scientists	Complete trust: 1 to Do not trust at all: 5	TrustInfoAboutVaccinesHealthOfficials
Trust_Body_11	How much trust do you have in doctors and other health professionals	Complete trust: 1 to Do not trust at all: 5	TrustInfoAboutVaccinesHealthOfficials
Vaccine_beliefs_1	The Pfizer/BioNTech vaccine is effective	Completely disagree: 0 to Completely agree: 100	BeliefVaccinationEfficacyAgainstBeingInfected
Vaccine_beliefs_3	The Oxford/AstraZeneca vaccine is effective	Completely disagree: 0 to Completely agree: 100	BeliefVaccinationEfficacyAgainstBeingInfected
Vaccine_beliefs_5	The Moderna vaccine is effective	Completely disagree: 0 to Completely agree: 100	BeliefVaccinationEfficacyAgainstBeingInfected
Vaccine_beliefs_2	The Pfizer/BioNTech vaccine is safe	Completely disagree: 0 to Completely agree: 100	PerceivedRiskOfVaccinating
Vaccine_beliefs_4	The Oxford/AstraZeneca vaccine is safe	Completely disagree: 0 to Completely agree: 100	PerceivedRiskOfVaccinating
Vaccine_beliefs_6	The Moderna vaccine is safe	Completely disagree: 0 to Completely agree: 100	PerceivedRiskOfVaccinating

Supplementary Table S5. Proportion of vaccine hesitant individuals (n = 248 in Norway and n = 109 in UK) within different demographic subgroups.

Variable	Norwegian sample $N_{hesitant}$ of $N_{subgroup}$ (%)	UK sample $N_{hesitant}$ of $N_{subgroup}$ (%)
Sex		
Female	205 of 2332 (8.79%)	58 of 419 (13.84%)
Male	43 of 572 (7.52%)	50 of 312 (16.03%)
Missing[a]	0 of 248 (0%)	1 of 109 (0.92%)
Age (years)		
18-30	95 of 1131 (8.40%)	23 of 148 (15.54%)
31-44	93 of 907 (10.25 %)	46 of 312 (14.74%)
45-64	49 of 716 (6.84%)	36 of 259 (13.90%)
65+	10 of 153 (6.54%)	4 of 15 (26.67%)
Missing[a]	1 of 248 (0.40%)	0 of 109 (0%)
University degree		
Yes	149 of 1767 (8.43%)	25 of 198 (12.63%)
No	98 of 1140 (8.60%)	28 of 142 (19.72%)
Missing[a]	1 of 248 (0.40%)	56 of 109 (51.38%)
Residency		
Urban	131 of 1844 (7.10%)	15 of 87 (17.24%)
Rural area	26 of 306 (8.50%)	38 of 253 (15.02%)
Missing[a]	91 of 248 (36.69%)	56 of 109 (51.38%)
Ethnic status		
Ethnic majority[b]	235 of 2727 (8.62%)	44 of 292 (15.07%)
Ethnic minority	12 of 179 (6.70%)	9 of 48 (18.75%)
	1 of 248 (0.40%)	56 of 109 (51.38%)
Psychiatric diagnosis		
Yes	46 of 526 (8.75%)	23 of 182 (12.64%)
No	201 of 2380 (8.45%)	78 of 517 (15.09%)
Missing[a]	1 of 248 (0.40%)	8 of 109 (7.34%)

Note. [a]Missing responses within the vaccine hesitant group. [b]White British or Irish (UK sample) or white Norwegian (Norwegian sample) ethnicity.

Supplementary Document 1. List of variables used in the machine learning model

List of the 111 variables used in the machine learning model.

Variable name	Description	Coding
OverconfidenceInC19Avoidance	I do not need to vaccinate as I have managed to avoid the coronavirus so far	Completely disagree: 1 to Completely agree: 5
SideEffectFearsDueToIllness	I am afraid of vaccinating because of possible side-effects or related to underlying disease	Completely disagree: 1 to Completely agree: 5
TrustInfoAboutVaccinesHealthOfficials	I trust the information I receive from health officials about vaccines	Completely disagree: 1 to Completely agree: 5
BeliefVaccinationEfficacyMitigateTransmission	I believe vaccination can help mitigate transmission of infectious disease	Not at all probable: 1 to Highly probable: 5
BeliefInSuperiorityOfNaturalImmunity	It is less risky for me to obtain natural immunity rather than vaccinating against the coronavirus	Completely disagree: 1 to Completely agree: 5
GovernmentalTrust	I trust in the government and their decision making regarding the pandemic	Completely disagree: 1 to Completely agree: 5
BeliefVaccinationEfficacyAgainstBeingInfected	I believe the COVID-19 vaccines are effective in protecting against infection from the coronavirus	Completely disagree: 1 to Completely agree: 5
PerceivedRiskOfVaccinating	How risky or dangerous do you perceive it to vaccinate against the coronavirus?	Not risky at all: 0 to Maximally risky: 100
BeliefVaccinesDevelopedToofastToBeSafe	I believe the vaccines have been developed too fast to be safe.	Completely disagree: 1 to Completely agree: 5
PerceivedDangerOfCOVIDForSociety	Perceived danger of the SARS-CoV-2 virus for society	Not at all serious: 0 to Very serious: 5
Sex	Sex assigned at birth	Females: 0; Males: 1

Age	Age in years	Integer, number of years
AgeGroup	Age group	18-30 years; 31-44 years; 45-64 years; and 65 years and above
Education	Highest education to date	Compulsory school: 0; Upper secondary High school: 1; Student: 2; Any university degree: 3
EmploymentStatus	Employment status	Full time job; Part time; Sick leave; Work clearance allowance; Disability benefits; Retired; Student; Homemaker; Parental leave; Other
County	County of residence	Agder; Innlandet; Møre og Romsdal; Nordland; Oslo; Rogaland; Troms og Finnmark; Trøndelag; Vestfold og Telemark; Vestland; Viken
Ethnicity	Ethnic background	Norwegian; European; Middle-Eastern or North-African; Middle-Asian; East- and South-East Asian; Central and South African; North-American Background; Latin-, Caribbean, and South-American; Oseanic background

ImmigrationStatus	Refugee or immigration status	No; I am a first generation immigrant/refugee; I am a second generation immigrant
Region		East Norway; West Norway; Middle Norway; North Norway
CityResidency	City (versus rural) residency	Yes; No
LivesWithChild	Lives with children (< 18 years)	Yes; No
BMI	Body Mass Index	Numeric
NumberOfChildrenCohabiting	Number of children (< 18 years) participant resides with	Integer, number of children
HealthProfessional	Heath professional (e.g., physician, nurse)	Yes; No
LoneResidency	Lives alone	Yes; No
CivilStatusChange	Civil status change during the pandemic	No; Yes, got in a relationship/civil partnership/married; Yes, became single/divorced; A combination of the above
NotWorking	Unemployed	Yes; No
CurrentPsychiatricTreatment	Receiving current psychiatric treatment	Not seeking any psychological treatment; Treatment related to anxiety; Treatment related to depressive symptoms; Treatment related to loneliness; Treatment related to stress and trauma-related problems; Treatment for loss

		and/or grief; Treatment for obsessive-compulsive problems; and Treatment for other psychological problems
PreviousPsychiatricTreatment	Receiving previous psychiatric treatment	No; For anxiety disorder; Depression; Bipolar or other mood disorder; Eating disorder; Obsessive-compulsive disorder; Trauma and stress disorder; Panic disorder; Insomnia or other sleep disorder; Alcohol or substance use disorder; Personality disorder; Psychotic disorder; Dissociative disorder; Attention-deficit disorder; Treatment for other psychiatric diagnoses
PsychiatricDiagnosisCurrent	Current psychiatric diagnosis	None; Anxiety disorder; Depression; Bipolar or other mood disorder; Eating disorder; Obsessive-compulsive disorder; Trauma and stress disorder; Panic disorder; Insomnia or other sleep disorder; Alcohol or substance use

		disorder; Personality disorder; Psychotic disorder; Dissociative disorder; Attention-deficit disorder; Other psychiatric diagnosis
MedicalComorbidityBeforePandemic	Medical illness from before the pandemic	None; Lung disease; Blood-related disease; Cancer; Heart-related disease; ME/Chronic Fatigue; Congenital disease; "Ear, nose, or throat (otolaryngology) disease; Eye disease; Infectious disease; Immune or inflammatory disease; Metabolic or endocrine-related disease; Musculoskeletal and joint disease; Neurological disease
MedicalComorbidityCurrent	Current medical illness	None; Lung disease; Blood-related disease; Cancer; Heart-related disease; ME/Chronic Fatigue; Congenital disease; "Ear, nose, or throat (otolaryngology) disease; Eye disease; Infectious disease; Immune or inflammatory disease; Metabolic or endocrine-related disease; Musculoskeletal

		and joint disease; Neurological disease
ChronicDiseaseCurrent	Current chronic illness	Yes; No
LearningDifficultiesCurrent	Learning difficulties	None; Dyslexia; Dysgraphia or dysorthographia; dyscalculia; General learning difficulties; Other learning difficulties
DepressionLevel	Patient Health Questionnaire-9 (PHQ-9), total score	Numeric, range: 0-29
AnxietyLevel	Generalized Anxiety Disorder-7 (GAD-7), total score	Numeric, range: 0-21
DepressionFunctioning	Depressive symptoms interfering with daily life functioning	Not at all: 0 to Nearly every day: 3
AnxietyFunctioning	Anxiety symptoms interfering with daily life functioning	Not at all: 0 to Nearly every day: 3
Loneliness	UCLA Loneliness Scale (ULS-8)	Numeric, range: 8-32
ObsessiveThoughtsC19	Obsession with COVID-19 Scale (OCS)	Numeric, range: 0-16
SleepDifficulties	Bergen Insomnia Scale (BIS)	Numeric, range: 0-42
WellBeing	Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS)	Numeric, range: 7-35
TraumaSymptoms	PTSD Checklist for DSM-5 (PCL-5)	Numeric, range: 0-80
ParentalStress	Danish Parental Stress Scale (PSS) – Short Form	Numeric, range: 3-15
SomatoSensoryAmplification	SomatoSensory Amplification Scale (SSAS) – Short Form	Numeric, range: 4-20
SelfEfficacy	General Self-Efficacy Scale (GSE) – Short Form	Numeric, range: 2-8

Fatigue	The Shortened Fatigue Questionnaire (SFQ)	Numeric, range: 4-28
SomaticSymptoms	Patient Health Questionnaire-15 (PHQ-15)	Numeric, range: 0-30
C19Infection	Infected by SARS-CoV-2	Yes; No
BedriddenC19Infection	Number of days bedridden during SARS-CoV-2 infection	Numeric, range: 0-30
VaccineAccessibility	Experience of vaccines being easy to access	Strongly Disagree: 1 to Strongly Agree: 5
BeliefinEfficacyOfMitigationProtocols	Social distancing protocols are effective in mitigating transmissions	Highly unlikely/Not at all: 0 to Highly likely: 5
PerceivedDangerOfCOVIDForSelf	Perceived danger of being infected by SARS-CoV-2	Not at all serious: 0 to Very serious: 5
PerceivedDangerOfCOVIDForOthers	Perceived danger of others being infected by SARS-CoV-2	Not at all serious: 0 to Very serious: 5
PerceivedDangerC19Overall	Perceived danger of the SARS-CoV-2 in general	Numeric, range: 0-15
AdherenceHygienicProtocols	Adherence to hygienic behavior recommendations	Numeric, range: 0-28
AdherenceSDPProtocols	Adherence to social distancing protocols	Numeric, range: 0-20
QuarantineExposureFrequency	Number of times quarantined	Integer
IsolationExposureFrequency	Number of times in isolation	Integer
PhysicalSocialContact	Extent of engagement in physical social contact during past two weeks	Not at all: 0 to Extremely: 5
DigitalSocialContact	Extent of engagement in digital social contact during past two weeks	Not at all: 0 to Extremely: 5
NotLeftOwnHomePast14Days	Number of days subject has not been outside their home during the past two weeks	Integer, range: 0-14

PhysicalActivity	Physical activity frequency during the past two weeks	Integer, range: 0-14
InformationNewspapers	Frequency of information acquisition from newspapers during the last month	Not at all: 0 to 7: Multiple times per hour
InformationTV	Frequency of information acquisition from television during the last month	Not at all: 0 to 7: Multiple times per hour
InformationSocialMedia	Frequency of information acquisition from social media during the last month	Not at all: 0 to 7: Multiple times per hour
InformationsForumsBlogs	Frequency of information acquisition from forums and/or blogs during the last month	Not at all: 0 to 7: Multiple times per hour
InformationPeers	Frequency of information acquisition from peers during the last month	Not at all: 0 to 7: Multiple times per hour
InformationOther	Frequency of information acquisition from other sources during the last month	Not at all: 0 to 7: Multiple times per hour
InformationActiveAvoidance	Frequency of active decision to avoid all information/news sources during the last month	Not at all: 0 to 7: Multiple times per hour
TotalInformationObtainmentTime	Mean score of information acquisition behavior across sources	Numeric, range: 0-7
FinancialDifficulties	Financial difficulties	Not at all: 0 to Very much: 4
FinancialOccupationalWorry	Extent of worry about losing job or personal economy	Not at all: 0 to Nearly every day: 3
FearOwnHealthC19	Fear of own health from SARS-CoV-2	Numeric, range: 0-6
FearSignificantOthersDyingC19	Worry about significant others dying from SARS-CoV-2	Not at all: 0 to Nearly every day: 3
FearOthersHealthC19	Worry about others' health related to SARS-CoV-2	Numeric, range: 0-6

FearOfTransmittingC19Others	Worrying about transmitting others with SARS-CoV-2	Not at all: 0 to Nearly every day: 3
HealthAnxietyAboutC19	Health anxiety related to SARS-CoV-2	Numeric, range: 0-6
FearDeath	Fear of dying from SARS-CoV-2	Not at all: 0 to Nearly every day: 3
IncreaseTobaccoUse	Increased usage of tobacco during the pandemic	No increase: 0 to Much more than before the pandemic: 3
IncreaseAlcConsumption	Increased consumption of alcohol during the pandemic	No increase: 0 to Much more than before the pandemic: 3
IncreaseFoodIntake	Increased food consumption during the pandemic	No increase: 0 to Much more than before the pandemic: 3
IncreaseSweetsConsumption	Increased sweets consumption during the pandemic	No increase: 0 to Much more than before the pandemic: 3
EmotionRegulationDifficulties	Difficulties in Emotion Regulation Scale (DERS) – Short Form	Numeric, range: 6-30
NegativeMetacognitions	Cognitive-attentional Syndrome Questionnaire (CAS-1) – Negative metacognitions subscale	Numeric, range: 0-400
MaladaptiveCopingStrategies	Cognitive-attentional Syndrome Questionnaire (CAS-1) – Maladaptive coping strategies subscale	Numeric, range: 0-64
IntoleranceOfUncertainty	Intolerance of Uncertainty Scale (IUS) – Short Form	Numeric, range: 3-15
ExecutiveFunctioning	Self-Report Measure of Executive Function for Administration via the Internet (WebEXEC)	Numeric, range: 1-24
Impulsivity	Impulsive behavior	No problem with this at all: 0 to Frequent problems with this: 4

CloseOthersInfectedC19	Significant other infected by SARS-CoV-2	Yes; No
WhichCloseOthersInfectedC19	Which significant other infected by SARS-CoV-2	Partner; Child; Parent; Sibling; Grandparent; Close Friend; Grandchild; Other family; Other
CloseOtherHospitalizedC19	Significant other hospitalized after SARS-CoV-2 infection	Yes; No
CloseOtherICUC19	Significant other sent to ICU after SARS-CoV-2 infection	Yes; No
CloseOtherDeathC19	Significant other deceased from SARS-CoV-2 infection	Yes; No
CloseOthersMaintainedSymptomatologyC19	Significant other suffers from sustained symptoms after COVID-19 infection	Yes; No
CloseOtherSymptomLengthC19	Length suffering from long-term symptom experience after SARS-CoV-2 infection by significant other	More than 1 month: 1; 1-3 months: 2; 4-6 months: 3; More than 6 months: 6
Extroversion	Brief Version of the Big Five Personality Inventory (BFI-10) – Extroversion subscale	Numeric, range: 2-10
Conscientiousness	Brief Version of the Big Five Personality Inventory (BFI-10) – Conscientiousness subscale	Numeric, range: 2-10
Neuroticism	Brief Version of the Big Five Personality Inventory (BFI-10) – Neuroticism subscale	Numeric, range: 2-10
Openness	Brief Version of the Big Five Personality Inventory (BFI-10) – Openness subscale	Numeric, range: 2-10
Agreeableness	Brief Version of the Big Five Personality Inventory (BFI-10) – Agreeableness subscale	Numeric, range: 2-10
AutonomyNeeds	Experience a need for autonomy	Not at all: 0 to Nearly every day: 3
AutonomyFrustration	Experience autonomy need deficit and frustration	Not at all: 0 to Nearly every day: 3

Altruism	Desire to help society and societal peers	Not at all: 0 to Nearly every day: 3
ContextualConsideration	Often think about the health and welfare of others around me	Not at all: 0 to Nearly every day: 3
PerceivedCompetence	Belief in own ability to deal with novel crises	Not at all: 0 to Nearly every day: 3
CivilStatus	Married or in a civil partnership	Yes; No
Mentalization	Others describing me as good of perceiving of others internal states	Does not apply to me at all: 1 to Applies to me nearly all the time: 5
AvailableSupport	Experience sufficient social support by others	Not at all: 0 to Nearly every day: 3
InterpersonalProblems	Inventory of Interpersonal Problems – Short Form	Numeric, range: 0 to 68
MediaTypePreference	Preferred media platform type	Recognized and source-checked national, regional and local platforms; Unmonitored information platforms

Supplementary Document 2. Audit trail of the analytic procedure in the thematic analysis

To analyze the 52 open-ended survey responses from the Norwegian adults who had reported that they had not taken and will not take the COVID-19 vaccines, the second and third authors of the study used the template analysis approach to thematic analysis (Brooks et al., 2015; King, 2012). This method was chosen because it involves a high degree of structure in the analysis of data while retaining flexibility by creating a coding template while working with the data. This template is then revised and refined as the analysis develops. The following supplementary document presents an audit trail of our analytical procedure to increase transparency (see Levitt et al., 2018; Steltenpohl et al., 2023). As per King (2012) and Brooks et al. (2015), our approach is divided into six steps.

1. Become familiar with the accounts to be analyzed

In the first step, both authors familiarized themselves with the data by reading and re-reading the dataset consisting of the 53 open-ended responses provided by those participants who had chosen to not get vaccinated to the question: "If you did not get vaccinated against COVID-19, could you please explain what contributed to you not wanting to get vaccinated?". One of the responses was excluded from further analysis due to the response merely stating: "No", thus, not providing any underlying reason for the unwillingness to vaccinate, leaving 52 responses. The full dataset cannot be publicly shared due to confidentiality.

2. Carry out preliminary coding of the data

The second step involves the first coding of the data that eventually results in the initial coding template (Brooks et al., 2015). Coding is done by highlighting anything in the text that can contribute to the understanding of the research question, which was: "What reasons do the participants provide for their choice of not wanting to get vaccinated?". Both authors independently coded the full dataset line-by-line in two columns: one for manifest meaning that was close to the participants' words, and one for potential latent meaning that allowed for the researchers' interpretation of the meanings in the dataset. Below is an example from one of the researchers' first coding of the data (Table S6):

Supplementary Table S6. Example from initial coding.

Respondents' answer	Initial codes (Manifest meaning)	Initial codes (Latent meaning)
I am young and healthy. I have little contact with older or vulnerable people. The vaccine provides temporary protection.	Young and healthy	Unnecessary if not in a high-risk group
	Limited contact with high-risk groups	Unnecessary if no contact with high-risk groups
	Vaccine provides only temporary protection	Lack of belief in the necessity of the vaccine
COVID-19 is here to stay. Given the known side-effects, I believe it has been unethical to include children and young people in the vaccination campaign, especially in the way it has been done.	COVID-19 here to stay	
	Unethical to vaccinate children due to side-effects	Political disagreement with the campaigns
Newspapers have recently published a lot of hatred and misleading information about the unvaccinated, among other things.	Media spreads hatred and misinformation about the unvaccinated	Dissatisfaction with the media's portrayal
		Non-vaccination politicized (i.e., as a protest)
My lack of vaccination is therefore partly a personal assessment of the personal health risk posed by the choices presented to me, and a protest against some of the discussed political points.	Assessment of my own health risk as low	
		Autonomy - an active choice
	A protest against the political points	

([Gender anonymized], [years anonymized])

3. Organize the emerging themes into meaningful clusters

After both authors independently coded the whole dataset for manifest and latent meaning, they jointly went through both fully coded versions. From this process, the authors clustered meaningful codes into hierarchical themes in the coding template, focusing on the research question of what reasons the participants provided for their choice of not wanting to get vaccinated, with the aim of obtaining a full mapping of the different reasons provided. Some participants gave single reasons, while others provided various reasons. When giving various reasons, all were coded, as illustrated in the above example (Table S6) of coding from the same participant.

4. Define an initial coding template

The clustering of codes from stage 3 in the analytical process resulted in the initial coding template that is shown in Table S7. Here we highlight that we thus slightly deviated from the standard approach in Template analysis, where you typically make a coding template from a few of the interviews. This has been pointed out as a potential weak spot of template analysis, as one could risk overlooking new aspects in the latter interviews (see Braun & Clarke, 2022). We thus chose to do a thorough initial coding of all the material first, to avoid this limitation, and subsequently build our coding template.

Supplementary Table S7. Initial coding template.

Superordinate codes (main theme level)	**Subordinate code (sub-theme level)**	**Codes at sub-sub level (sub-sub theme level)**
General reasons (non-related to COVID-19	Fear of needles	
	Fear related to own underlying disease	
	Skeptical to vaccines in general	

vaccines or COVID-19)		
Experience-based health-related reasons	Negative experiences with previous vaccinations	
	Fear that would not get healthcare follow up if needed based on previous experiences	
Side-effects of COVID-19 vaccines	Long-term side-effects not clear	
	Scared of concrete side-effects	Menstruation changes
		Personally know others who have had severe side-effects
		More harmful side-effects than previous debated vaccines (such as the swine flu vaccine)
		Fertility reasons
	The risk of side-effects perceived as greater than risk associated with COVID-19	Some on the fence: Might take vaccine if this changes
Aspects with COVID-19	Not so dangerous	I am not specifically in the risk group
		I am not in the risk of infecting others
		Themselves/acquaintances have only had mild symptoms of COVID-19
	Serves from human activity on nature, nature's way of "taking back"/"re-balancing"	
Distrust	Lack of trust, authorities	
	Lack of trust, pharmaceutical industry	

		Do not want to support pharmaceutical industry
	Mainstream media misleading people	
		"Uncensored" media telling the truth
	Censorship of alternative voices	
Alternatives to the vaccine	Avoiding crowds	
	The course of nature	Acceptance of own death as course of nature
		Natural for old people to pass away
		Nature's way of handling overpopulation
	Natural immunity	Natural herd immunity
	Avoiding taking a stance because overwhelming	
	Own control: Strengthen own immune system, taking vitamin supplements, exercise, control diet	
General about the nature of the vaccine	The vaccine should be given to those who want/need it more	
	Process too quick, a lot of uncertainty	
	Not offering much protection	Does not hinder the spread of COVID-19
		Alarming to include kids and youth
	Disqualifying these as real vaccines	
	Not interested in the repeated rounds of vaccinations that are required	

Portrayal of the unvaccinated	Pressure to get vaccinated
	Unwilling due to the portrayal of the unvaccinated

5. Apply the initial template to data and modify as necessary

In this phase, both authors independently applied the initial coding template to the entire dataset, while considering whether any of the themes defined in this template could represent each data extract. After this the authors investigated the degree of overlap in their use of codes from the template. Of 317 coding instances 282 instances overlapped at the main theme level (88.96%) and 260 at the sub-theme level (82.02%). In the instances where the authors had coded differently from the template or had discovered data instances where the template did not fit, codes were discussed, and the coding template revised. For example, at one instance one of the authors coded an extract under the superordinate level code "Aspects with COVID", while the other author coded the same under the subordinate code "Risk of side-effects greater than the risk of COVID", demonstrating that these codes needed refinement and clarification. As emphasized by Brooks et al. (2015), the initial template should be revised in an iterative process until a rich and comprehensive representation of the interpretation of the data is achieved.

6. Finalize the template and apply it to the full dataset

The authors finalized the coding template and applied it to the full dataset. This template is presented as the six themes that are presented in the article. Below (Table S8) is the final coding template consisting of all levels of final codes. This final coding template is an extended thematic map of the same version as the one presented in the paper, but including all sub-levels.

Supplementary Table S8: Extended thematic map developed from coding template.

Theme	**Subtheme**	***N***	**Sub-subtheme**	***n* = Total respondents providing this as a reason** **# = Respondent number**
Theme 1: Unnecessary Vaccines (*n*=17)	1.1 COVID-19 in general not dangerous	*n*=12	1.1.1 Not at all dangerous	*n*=3 (#19, #35, #39)
			1.1.2 Not as dangerous as it is portrayed to be	*n*=7 (#4, #10, #17, #33, #41, #47, #48)
			1.1.3 Participants/acquaintances have only had mild symptoms of COVID-19	*n*=3 (#10, #23, #51)
	1.2 Perceiving oneself as not personally being at risk	*n*=9		*n*=9 (#4, #7, #9, #10, #14, #23, #24, #42, #48)
Theme 2: Inefficient Vaccines (*n*=19)	2.1 Does not offer much protection	*n*=11	2.1.1 Does not have a good effect	*n*=5 (#1, #22, #23, #38, #47)
			2.1.2 Effect does not last	

				n=6 (#4, #7, #9, #20, #27, #40)
	2.2 Does not hinder transmission	*n*=7		*n*=7 (#4, #10, #11, #13, #19, #40, #42)
	2.3 Disqualifying these as *real* vaccines	*n*=4		*n*=4 (#5, #13, #29, #52)
Theme 3: Frightening Vaccines (*n*=32)	3.1 Fear of side-effects	*n*=26	3.1.1 Generally fearing side-effects of the vaccines	*n*=12 (#4, #5, #7, #13, #14, #20, #24, #27, #35, #38, #45, #49)
			3.1.2 Uncertainty about possible long-term side-effects	*n*=7 (#11, #12, #27, #28, #36, #37, #50)
			3.1.3 Perceiving risks of side-effects as larger than risks of infection with COVID-19	*n*=6 (#10, #19, #23, #38, #48, #51)

				n=4
			3.1.4 Personally know others who have experienced side-effects	(#10, #29, #38, #51)
				n=4
			3.1.5 Worry about menstruation changes	(#21, #48, #50, #51)
				n=3
			3.1.6 Pregnancy/fertility reasons for fear of side-effects	(#23, #32, #48)
	3.2 Fear of not getting healthcare follow-up if needed	*n*=3		*n*=3 (#20, #21, #43)
	3.3 Vaccine approvals too quick	*n*=8		*n*=8 (#4, #9, #13, #16, #17, #25, #27, #40)
Theme 4: Distrust and Polarization (*n*=18)	4.1 Authorities' vaccine handling	*n*=10	4.1.1 Not providing sufficient and/or correct information	*n*=7 (#13, #14, #16, #18, #19, #31, #40)
			4.1.2 Distrusting authorities' vaccine handling	*n*=7 (#3, #4, #16, #18, #31, #40, #52)

4.2 Pharmaceutical industry	*n*=3	4.2.1 Do not want to support pharmaceutical industry economically	*n*=1 (#38)
		4.2.2 Distrusting pharmaceutical industry	*n*=3 (#3, #16, #38)
4.3 Mainstream media's vaccine presentation	*n*=7	4.3.1 Mainstream media as misleading	*n*=3 (#7, #31, #35)
		4.3.2 Refers to vaccine critical sources	*n*=6 (#5, #19, #25, #31, #35, #48)
4.4 Portrayal of the unvaccinated	*n*=6	4.4.1 Polarizing portrayals contributing to unwillingness to get vaccinated	*n*=3 (#7, #20, #40)
		4.4.2 Pressure to get vaccinated	*n*=3 (#19, #22, #31)

Theme 5: Alternatives to Getting Vaccinated (*n*=15)	5.1 Natural immunity	*n*=4	5.1.1 Natural herd immunity as the solution to the pandemic	*n*=1 (#4)
			5.1.2 Own natural immunity through infection with COVID-19 gives better protection than the vaccine	*n*=3 (#9, #30, #42)
	5.2 The course of nature	*n*=5	5.2.1 Viewing own potential death as natural	*n*=3 (#8, #15, #39)
			5.2.2 Natural for old people to pass away	*n*=1 (#14)
			5.2.3 Pandemic as nature's way of handling overpopulation/misuse	*n*=2 (#8, #9)
	5.3 Own health measures	*n*=2	5.3.1 Strengthen own immune system	*n*=1 (#40)
			5.3.2 Staying physically fit	*n*=1 (#48)
	5.4 Avoid infections	*n*=3	5.4.1 Avoid places where one can get infected	*n*=2 (#24, #44)

			5.4.2 Avoid people in vulnerable groups	*n*=1 (#7)
	5.5 Avoidance as easier than taking a stance	*n*=2		*n*=2 (#2, #20)
Theme 6: Reasons Related to Vaccines in General (*n*=12)	6.1 Fear of needles	*n*=4		*n*=4 (#6, #20, #38, #46)
	6.2 Fear related to own underlying disease	*n*=4		*n*=4 (#21, #31, #34, #45)
	6.3 Negative experiences with previous vaccination/ medication	*n*=4		*n*=4 (#26, #29, #33, #50)

References

Braun, V. & Clarke, V. (2022). *Thematic analysis. A practical guide.* Sage.

Brooks, J., McCluskey, S., Turley, W. & King, N. (2015). The Utility of Template Analysis in Qualitative Psychology Research, *Qualitative Research in Psychology, 12*(2), 202-222.

King, N. (2012). Doing template analysis. In G. Symon & C. Cassell (Eds.) *Qualitative Organizational Research: Core Methods and Current Challenges*, (pp. 426-450). SAGE.

Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal article reporting standards for qualitative primary, qualitative meta analytic, and mixed methods research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 26-46.

Steltenpohl, C. N., Lustick, H., Meyer, M. S., Lee, L. E., Stegenga, S. M., Reyes, L. S., & Renbarger, R. (2023). Rethinking transparency and rigor from a qualitative open science perspective. *Journal of Trial & Error*, online first.