Modeling prevention behaviors during the COVID-19 pandemic using Bayesian belief networks and protection motivation theory

Journal:	*Risk Analysis*
Manuscript ID	RA-00286-2023.R2
Wiley - Manuscript type:	Original Article
Key Words:	COVID-19, Bayesian belief network, protection motivation theory

Modeling prevention behaviors during the COVID-19 pandemic using Bayesian belief networks and protection motivation theory

Abstract

Prevention behaviors are important in mitigating the transmission of COVID-19. The protection motivation theory (PMT) links perceptions of risk and coping ability with the act of adopting prevention behaviors. The goal of this research is to test the application of the PMT in predicting adoption of prevention behaviors during the COVID-19 pandemic. Two research objectives are achieved to explore motivating factors for adopting prevention behaviors. (1) The first objective is to identify variables that are strong predictors of prevention behavior adoption. A data-driven approach is used to train Bayesian belief network (BBN) models using results of a survey of N=7,797 participants reporting risk perceptions and prevention behaviors during the COVID-19 pandemic. A large set of models are generated and analyzed to identify significant variables. (2) The second objective is to develop models based on the PMT to predict prevention behaviors. BBN models that predict prevention behaviors were developed using two approaches. In the first approach, a data-driven methodology trains models using survey data alone. In the second approach, expert knowledge is used to develop the structure of the BBN using PMT constructs. Results demonstrate that trust and experience with COVID-19 were important predictors for prevention measure adoption. Models that were developed using the PMT confirm relationships between coping appraisal, threat appraisal, and protective behaviors. Data-driven and PMT-based models perform similarly well, confirming the use of PMT in this context. Predicting adoption of social distancing behaviors provides insight for developing policies during pandemics.

Keywords: COVID-19, Bayesian belief network, protection motivation theory, prevention behaviors

1 INTRODUCTION

Prevention behaviors, such as wearing masks, social distancing, and proper handwashing hygiene are used as a first line of defense in epidemics and pandemics to reduce the transmission of infectious diseases (Bavel et al., 2020; Epstein et al., 2008; Funk et al., 2009; Reluga, 2010). During the COVID-19 pandemic, prevention behaviors were mandated by governments around the world, and communities across the globe relied on behavioral methods to manage transmission of the disease (Chu et al., 2020; Ma et al., 2020). The coronavirus was discovered in late 2019; by January 2021, more than 80% of governments instituted mask mandates, and approximately 50% of governments mandated one or more prevention behaviors (Hale et al., 2022; Koh, 2020).

Though prevention behaviors can be effective at mitigating disease transmission, the efficacy of behavioral measures is governed by the uptake of these behaviors within communities. Individuals make decisions to follow guidelines for adopting prevention behaviors, and a growing body of literature has identified that risk perception, which is a perceived probability and severity of a threat, directly affects the adoption of prevention behaviors. A study performed in the Netherlands, Germany, and Italy found that prevention behavior adoption was highest in Italy, where COVID-19 cases were highest and individuals likely had a high risk perception (Meier et al., 2020). Further studies quantified the connection between risk perception and the adoption of prevention behaviors in Switzerland (Siegrist et al., 2021), the US (Bruin & Bennett, 2020), and the Netherlands (Botzen et al., 2022). Botzen et al. (2022) found that experience with COVID-19, including direct personal experience and indirect experience through the exposure of family or friends, is an important factor in prevention behavior adoption. Their research identified that the adoption of prevention behaviors is strongly associated with risk perceptions and feelings toward the risk.

Protection motivation theory (PMT) is a cognitive model that links an individual's perception of risk and coping ability with the act of adopting coping behaviors (Rogers, 1975). PMT is used in this research to describe and predict the adoption of prevention behaviors based on perceptions of risk and specifically attributes decisions to appraisal of coping abilities and the threat. PMT has been used to predict prevention behavior adoption in the context of health threats, including viral disease transmission (Fisher et al., 2018; Timpka et al., 2014). PMT has also been applied to characterize protective actions during COVID-19 (Nudelman et al., 2022; Qazi et al., 2022). PMT was used to model COVID-19 prevention behavior adoption using frequentist regression techniques and statistical regression that quantified the strength of relationships between variables (Nudelman et al., 2022). Nudelman et al. (2022) concluded that PMT could be a useful tool in the context of COVID-19, but their research did not develop models that could be used to predict adoption decisions.

The goal of the research described in this manuscript is to test the application of the PMT in developing models to predict adoption of prevention behaviors during the COVID-19 pandemic. Two research objectives are achieved to contribute to a broader understanding of the motivating factors for adopting prevention behaviors. **Objective 1 of this research is to identify significant variables that lead to the adoption of prevention behaviors.** A data-driven approach is used to develop Bayesian belief network (BBN) models for predicting prevention behavior adoption (Heckerman & Wellman, 1995). BBNs provide a flexible approach to construct relationships between variables. This research uses a dataset that was published in April 2020 and reports results of a survey of N=7,797 participants from 11 countries across four continents (Dryhurst et al., 2020). Respondents reported demographic data, perception of COVID-19 as a risk, and adoption of social distancing behaviors such as working from home, wearing a face-covering, and avoiding public transit. Models are generated to predict the adoption of prevention behaviors across a wide set of behaviors, including washing hands more often, using alcohol-based hand sanitizer more often, wearing a face mask, avoiding social events, avoiding public transport, eating out less, touching your face less, shopping for groceries less, cooking at home more, staying home from work, and purchasing extra supplies. Models are analyzed to identify significant variables, and this insight can be used to better inform policy for future pandemics around specific prevention behaviors.

Objective 2 of this research is to develop BBN models using the PMT to predict prevention behaviors. Prevention behaviors are grouped into two composite prevention behavior variables, personal protection and limiting exposure, to reduce the complexity of predicting individual decision-making and improve model predictability. BBN models are developed using two approaches to allow comparison and analysis of the PMT as a construct. The first approach uses a data-driven methodology that fits model structure using survey data alone. The second approach applies expert knowledge to develop the structure of the BBN using PMT variables and relationships. Survey data included core PMT tenets, including questions that assess response efficacy, threat appraisal, and threat severity, and BBN models were created to link these variables using relationships described by the PMT. By comparing data-driven models with PMT-based models, this research tests the efficacy of PMT in the context of COVID-19. Modeling methods and behavioral theories that predict actions are important tools in virus suppression and pandemic mitigation. This research develops new insight into the factors that affect behaviors and generates new models that predict individual decisions to adopt protective behaviors. New knowledge and models can be used to guide development of public health policies during pandemics.

2 BACKGROUND

2.1 Protection Motivation Theory

This research applies PMT as a framework for creating BBN models. PMT is a social cognition model, which is a theoretical construct that describes the influence of individual experiences, the actions of others, and environmental factors on individual behaviors (Conner, 2010). PMT explains how individuals are motivated to react with self-protective behaviors in response to health threats (Rogers, 1975). Two key elements of PMT are threat appraisal and coping appraisal. Threat appraisal is based on threat severity, which is the individual's assessment of the severity of COVID-19, and threat susceptibility, which is the individual's perception of the likelihood of contracting COVID-19. Coping appraisal refers to the ability of the individual to cope with the perceived threat and is based on self efficacy and response efficacy. Self efficacy is an individual's ability to take action, such as wearing a mask properly, and response efficacy, which is the perception that a prevention behavior is effective. PMT posits that if an individual perceives each of these constructs positively, then the individual will adopt the behavior. For example, if an individual perceives that contracting COVID-19 is dangerous and likely, and wearing a mask is achievable and effective, then the individual will engage in wearing a mask for disease mitigation.

PMT has been used to predict prevention behavior adoption in various scenarios. In healthcare, PMT has been applied to predict risk-based health behaviors such as smoking cessation, skin cancer prevention, and malaria prevention (Ghahremani et al., 2014; Moeini et al., 2019; Xu & Chen, 2016). PMT has also been used to predict intentions to adopt protection behaviors in response to viral diseases, including the adoption of hand washing to prevent norovirus on cruise ships and protective behaviors to prevent flu infection in Sweden (Fisher et al., 2018; Timpka et al., 2014). In the context of COVID-19, PMT has been used to characterize COVID-19 protection behaviors using a moderately sized sample (N=1,299) from three countries (Germany, India, Israel) (Nudelman et al., 2022). Nudelman et al. (2022) used correlations between survey response variables to conclude that PMT could be an effective tool to predict COVID-19 protection behaviors. Furthermore, they concluded that in the countries they surveyed, response efficacy, an individual's perception that a protection behavior is effective, was the most important PMT variable. A global study, including data from 191 countries, identified "lack of coping capacity" as a significant driver for COVID-19 risk perception (Qazi et al., 2022). Qazi et al. (2022) links coping capacity, the ability of an individual to undertake coping behaviors, to COVID-19 risk perception, and COVID-19 risk perception to an individual's desire to adopt prevention behaviors. These studies demonstrate the strength of PMT as a theoretical framework in predicting the adoption of prevention behaviors in the context of COVID-19.

2.2 Bayesian Belief Networks

Bayesian belief networks (BBNs) provide an approach to characterize relationships among variables that are characterized by high uncertainty (Kelly (Letcher) et al., 2013). BBNs consist of directed acyclic graphs (DAGs) and conditional probability tables (CPTs). The DAG shows relationships between variables, while the CPTs show the strength of those relationships. The DAG and CPT for a BBN model can be determined using expert knowledge of the system or collected data that represents predictive variables (Heckerman & Wellman, 1995). Specific to COVID-19, BBNs have been used to predict patient movement patterns, give more clarity to rapid test results, and predict an individual's likelihood of infection (Aliyu et al., 2021; Arrizza & Caimo, 2021; Wu et al., 2021). The study conducted by Aliyu et al. (2021) used a dataset of N=5,434 participants and achieved a maximum accuracy of 98% in predicting COVID-19 infection. High accuracy of prediction demonstrates the power of data-driven BBN models to predict outcomes based on collected survey results. BBN modeling has not been used to predict prevention behavior adoption during the COVID-19 pandemic to date.

3 MATERIALS AND METHODS

This research used a BBN modeling framework to develop models that predict prevention behaviors in response to the COVID-19 pandemic. Models were developed using two approaches. The data-driven method develops DAGs and CPTs based on statistical inference, using classic BBN and model building techniques, including the Naïve Bayes classifier and forward selection. A hybrid approach was developed to test the use of PMT in guiding model development. The hybrid approach integrated BBN modeling with an expert knowledge approach, which uses PMT to describe decision-making about prevention behaviors. The data-driven approach was applied to achieve Objective 1 by creating a large set of BBNs and analyzing the models for significant variables (Figure 1). To achieve Objective 2, the data-driven and hybrid methods were applied to develop BBNs (Figures 2 and 3), and BBN models that were developed using the two alternative approaches were compared to assess the application of PMT in predicting prevention behaviors. The dataset, BBN framework, and modeling methods are described as follows.

3.1 COVID-19 Risk Perception Dataset and Data Cleaning

This study used data generated through an international survey that explored how individuals around the world responded to the coronavirus and perceived information about protection behaviors (Dryhurst et al., 2020). Dryhurst et al. (2020) collected responses ($N = 7,797$) between mid-March and mid-April, 2020,

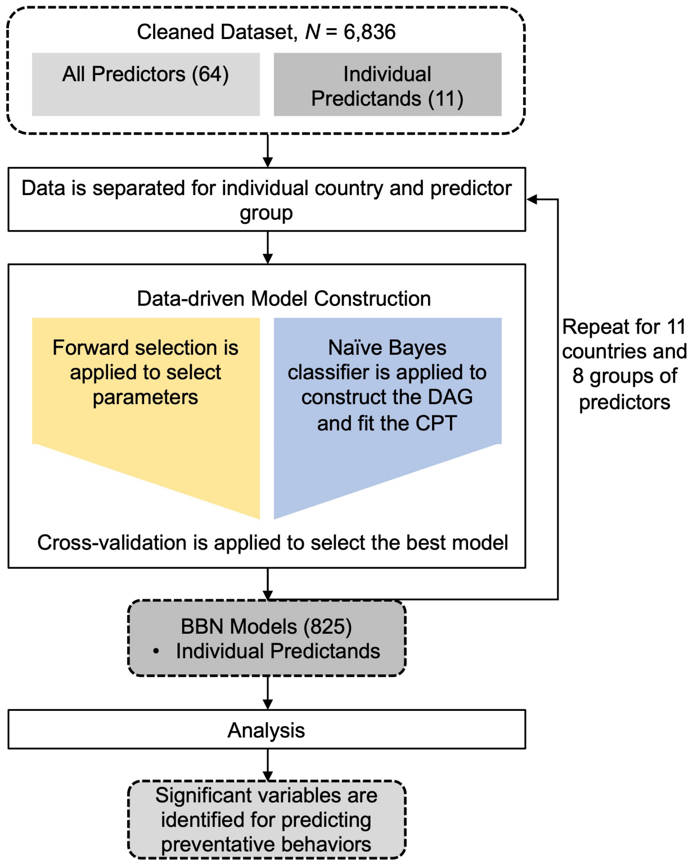

Figure 1: A data-driven approach applies Naïve Bayes classifier and forward selection to a large set of models that predict individual preventative behaviors to achieve Objective 1.

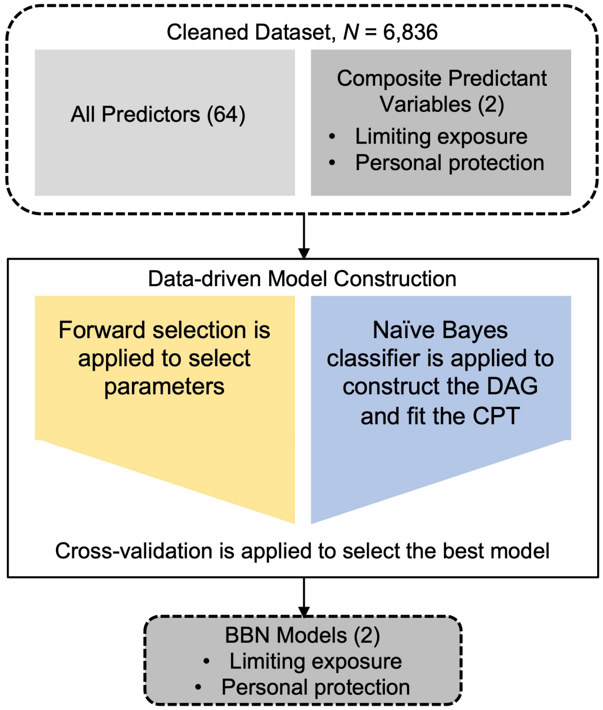

Figure 2: A data-driven approach applies Naïve Bayes classifier and forward selection to develop models that predict composite variables as part of Objective 2.

and published the dataset online for open access. Approximately 700 participants responded from each of 11 countries where the survey was administered, and participants are representative across age, sex, and ethnicity. The dataset consisted of responses to 103 questions.

The input data were cleaned prior to training by removing all variables that had more than 8% missing values. Responses that were missing data were removed subsequently. These steps ensured that the input data was free of missing values, which is a prerequisite of the modeling package, without removing an excessive number of responses. This process removes 27 variables and reduces the data set to 6,836 responses. Sixty-nine questions were included in the cleaned dataset.

3.1.1 Predictors

A total of 69 questions (listed in Table 5) were used to create variables, including 60 variables described in this section, and nine PMT variables described in the following section. Variables were used to create predictors, which were used in constructing BBNs.

Fifty of the survey questions were grouped into seven variable groups, including *Demographics* (5 questions), *Trust in Groups* (13), *Media Exposure* (7), *Current Worry* (7), *Future Worry - Personal (ps)* (8) and *Future Worry - Friends (fs)* (8), and *Cultural Cognition* (7 variables, including *Cultural Cognition-1,-2,-3,-*

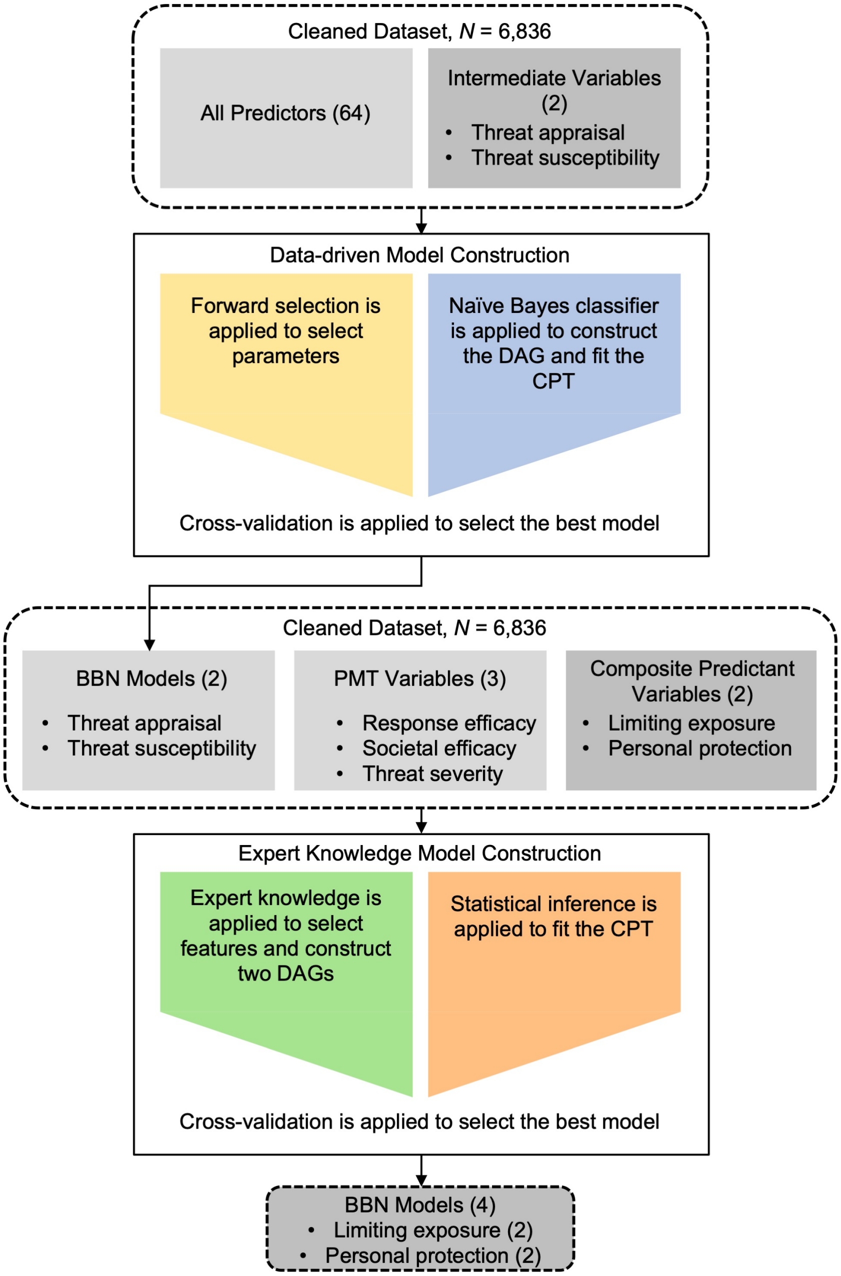

Figure 3: A hybrid approach applies data-driven and expert knowledge methods to develop models that predict composite variables as part of Objective 2. DAGs developed using expert knowledge are shown in Figure 4.

4,-5, and *-6* and *Prioritize Society*). Table 5 shows variables and questions associated with each variable group. Each group consolidates responses to a question that is asked about a range of prompts. The variable group *Cultural Cognition* was developed based on responses to a set of seven questions that explore opinions about the role of government and personal choice in society. These questions were characterized as cultural cognition variables based on the framework presented by Kahan (2012).

A set of six variables represent both direct and indirect experience with COVID-19 or a similar epidemic: *COVID-19 Experience*, *SARS*, and a group of three variables labeled *Effect*. *COVID-19 Experience* represents the direct experience with contracting COVID-19 or suspecting a COVID-19 infection. *SARS* represents indirect experience with a similar epidemic, such as SARS or MERS. The *Effect-Friends* variable represents personal impacts of the COVID-19 pandemic. The other three variables in the *Effect* group include *Effect-Social*, *Effect-Financial*, and *Effect-Mental health*. These three variables represent tangential impacts of the pandemic.

Seven other variables used in this analysis include *General Trust*, *Information*, *Scientist Understanding*, two *Certainty* variables, two *Vaccine* questions, and three *Longitude* variables. These variables were included due to their high response rate and pertinence to prevention behaviors and other variable groups.

3.1.2 PMT Variables

The PMT, as described in Section 2.1, requires four constructs, including *Self Efficacy*, *Response Efficacy*, *Threat Severity*, and *Threat Susceptibility*. Because no survey question tested concepts that matched the definition of *Self Efficacy*, it was replaced with a surrogate, societal efficacy, which tests the capacity of society to limit exposure. *Societal Efficacy* was tested with the question, "To what extent do you feel the actions that your country is taking to limit the spread of coronavirus make a difference?". *Response Efficacy* was mapped to the question, "To what extent do you feel that the personal actions you are taking to try to limit the spread of coronavirus make a difference?" and *Threat Severity* was mapped to the question, "How much do you agree or disagree with the following statement? Getting sick with the coronavirus/COVID-19 can be serious." Each of these questions are presented in Table 5.

Threat appraisal and threat severity were explored using a set of six risk perception variables, identified by Dryhurst et al. (2020), including the variables *Current Worry-Coronavirus*, *Future Worry(ps)-Coronavirus*, *Future Worry(fs)-Coronavirus*, *Country-Affect*, *Personal-Sick*, and *Threat Severity*. An exhaustive search of the six risk perception variables was conducted to identify a set of variables with the highest internal consistency. The group containing *Future Worry(ps)-Coronavirus* and *Future Worry(fs)-Coronavirus* ($\alpha =$ 0.86) had the highest internal consistency and theoretically represents perceptions that the individual and their friends and family will be affected by catching the coronavirus. This conforms to the definition of

Table 1: PMT composite variables are calculated as the average of input variables

Composite Variable	Input Variables
Threat Susceptibility	Future Worry (ps)-Coronavirus, Future Worry (fs)-Coronavirus
Threat Appraisal	Current Worry-Coronavirus, Future Worry(ps)-Coronavirus, Future Worry(fs)-Coronavirus, Country-Affect, Personal-Sick, and Threat Severity

threat susceptibility, therefore, the *Threat Susceptibility* variable was represented as the average value of the *Future Worry(ps)-Coronavirus* and *Future Worry(fs)-Coronavirus* variables. The six risk perception variables identified by Dryhurst et al. (2020) were used to represent a composite *Threat Appraisal* variable, based on the internal consistency of responses (Chronbach's $\alpha = 0.72$), and *Threat Appraisal* is calculated as the average of the six risk perception variables. Composite variables are summarized in Table 1.

3.1.3 Predictands

Predictands were developed to include 11 individual variables and two composite variables. Eleven prevention behavior variables were identified based on responses to one question: "Which of the following steps, if any, have you taken in the last month to prepare for the possibility of many cases of the coronavirus/COVID-19 in your community? Select all that apply." Participants could choose from the following responses: washing hands more often, using alcohol-based hand sanitizer more often, wearing a face mask, avoiding social events, avoiding public transport, eating out less, touching your face less, shopping for groceries less, cooking at home more, staying home from work, and purchasing extra supplies. Each response was used as a predictand representing one COVID-19 prevention behavior.

Composite behavioral variables were created to reduce the complexity of predicting multiple behavioral changes. The World Health Organization (WHO) recommends three categories of prevention behaviors, including increased hygiene, wearing a mask properly, and reducing environmental exposure (WHO, 2023). Previous work grouped prevention behaviors in the context of COVID-19 (Meier et al., 2020; Siegrist et al., 2021). Both Meier et al. (2020) and Siegrist et al. (2021) grouped a large set of prevention behaviors (16 and 13, respectively) into groups including personal protection (termed "personal protective behaviors" and "hygienic behaviors," respectively) and limiting exposure (termed "limiting interactions with people" and "physical distancing," respectively). (Meier et al., 2020) created one additional group titled "avoiding travel," that could be added to the limiting exposure group for simplicity. In this research, two prevention behavior composite variables were developed from the 11 prevention behaviors by first calculating the reliability of exhaustive combinations, then using expert judgement to define useful groups. *Limiting Exposure* is a composite variable that includes avoiding social events, avoiding public transport, eating out less, shopping for groceries less, cooking at home more, staying home from work, and purchasing extra supplies ($\alpha = 0.43$). *Personal Protection* is a composite variable that includes washing hands more often, using alcohol-based hand sanitizer more often, wearing a face mask, and touching your face less ($\alpha = 0.77$). Personal protection and limiting exposure represent important categories of prevention behaviors as they encompass the ways individuals protect themselves from COVID-19 and provide insight to decisions made to adopt prevention behaviors.

3.2 Bayesian Belief Network Framework

3.2.1 Naïve Bayes Classifier

The Bayes classifier specifies the construction of the DAG and affects the complexity of the DAG and the required computational time. The Naïve Bayes classifier is a simple classifier that includes only one parent node, and all other nodes are added as children nodes. The children nodes are independent variables, and connections between children nodes are not allowed. Naïve Bayes classifier is one of several Bayesian classifiers and was chosen because its performance is similar to other classifiers and requires less computational time to develop models (Fasaee et al., 2021).

3.2.2 Forward Selection

Forward selection was chosen as a method to select variables, based on previous work, which demonstrated that forward selection and backward elimination perform similarly well in developing BBNs to fit survey data (Fasaee et al., 2021). The forward selection process initializes an empty set of variables (feature set) and adds variables one at a time, calculating the performance of each set at each step. The set with the highest performance metric is selected and used to construct the BBN model. During forward selection, a series of cross-validation steps were performed to reduce systematic error in the selection of responses used for the training and validation datasets. Cross-validation was completed with 10 runs of 10 folds with each run using nine folds for training and one fold for validation. The fold used for validation was changed each run, and performance was averaged over the 10 runs. The Naïve Bayes classifier and forward selection were coded in R using the *bnlearn* library (Scutari, 2010).

3.2.3 Performance Metrics

BBN models were evaluated using accuracy and the F_1 metric. Accuracy is defined as the ratio of the number of true predictions made to the total number of predictions:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

The F_1 metric is the harmonic mean of both the recall and precision metrics, defined as the proportion of true positives to the total correct values and the total true values, respectively.

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + TN} \tag{3}$$

$$F_1 = \frac{2 \cdot Recall \cdot Precision}{Recall + Precision} \tag{4}$$

The performance metric used for variable selection, or the performance used to determine the best variable to add to the model at each iteration, was selected based on the proportion of positive responses in the predicted variable. When the proportion of positive responses was within 20% of parity, the dataset was labeled balanced, and accuracy was used as the measure. For datasets where the proportion of positive responses was skewed high or low (0-30% or 70%-100% positive responses), the F_1 metric was used for variable selection, as it better characterizes the performance of a model for imbalanced datasets.

3.3 Modeling Methods

This research developed and applied two modeling methods to construct BBN models to predict prevention behaviors. The data-driven method and the hybrid method are described as follows.

3.3.1 Data-driven Method

The data-driven method applied the Naïve Bayes classifier and forward selection to develop DAGs and CPTs (Figures 1 and 2). Each data-driven model was developed to predict one predictand variable, either an individual prevention behavior or a prevention behavior composite variable (described above in Section 3.1.3).

3.3.2 Hybrid Method

The hybrid method used an expert knowledge approach, based on PMT, to develop a DAG to predict prevention behaviors (Figure 3). In the first step, a model for intermediate PMT variables, *Threat Appraisal*

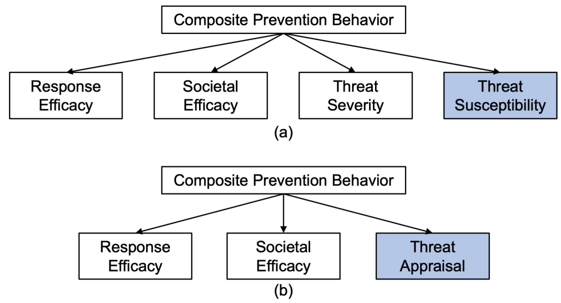

Figure 4: The hybrid approach was applied to develop two DAGs to predict composite prevention behaviors, *Limiting Exposure* and *Personal Protection*. (a) A DAG for *Threat Susceptibility* was developed using the data-driven method and inserted into the DAG that characterizes protection behaviors (Threat Susceptibility Model). (b) A DAG for *Threat Appraisal* was developed using the data-driven method and inserted into the DAG that characterizes protection behaviors (Threat Appraisal Model).

and *Threat Susceptibility* were developed using the data-driven approach. Predictors that are used to develop DAGs for *Threat Appraisal* and *Threat Susceptibility* include the list of all other variables listed in Table 5. The second step uses expert knowledge to develop DAGs for predictands *Limiting Exposure* and *Personal Protection* using the PMT variables *Response Efficacy*, *Societal Efficacy*, *Threat Susceptibility*, *Threat Severity*, and *Threat Appraisal*. Figure 4a shows the Threat Susceptibility Model and Figure 4b shows the Threat Appraisal Model. CPTs for each DAG are developed using statistical inference for the dataset of responses.

4 RESULTS

4.1 Objective 1: Identifying Significant Variables in the Adoption of Prevention Behaviors

The data-driven approach was applied to explore the relative importance of groups of variables and individual variables on predicting 11 individual prevention behaviors. Individual models were developed using variables from one variable group as predictors and one prevention behavior as the predictand for each model. Seven variable groups are shown in Table 5 as *Demographics*, *Trust in Groups*, *Media Exposure*, *Current Worry*, *Future Worry - Personal*, *Future Worry - Friends*, and *Cultural Cognition*. The variable group *All Predictors* includes all 64 variables listed in Table 5 (not including PMT variables). Each model was developed for data collected at one country, and the dataset includes 11 countries. Twenty-two models from the cultural cognition group could not be constructed and analyzed due to insufficient data from two countries (China and UK). A total of 825 models were created to predict individual prevention behaviors.

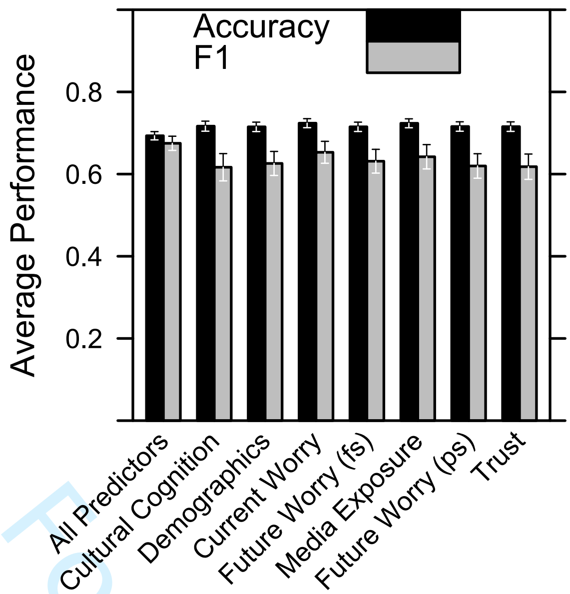

Figure 5: Mean accuracy and F_1 for data-driven models that were created using different variable groups. Mean is calculated over 121 models for each variable group. Mean for cultural cognition variable group is calculated over 99 models. Error bars represent standard error.

4.1.1 Model Performance

The performance of individual prevention behavior models is shown in Fig. 5, reported as the mean accuracy and F_1 across 121 models for each predictor group (11 countries and 11 predictands), except cultural cognition, which reports the mean performance over 99 models. Model performance was similar across all variable groups, with accuracy values ranging from 69.3-72.4% and F_1 values ranging from 61.7-67.5%. *Media Exposure* and *Current Worry* reported the highest predictive performance with accuracy values of 72.4%. The models built with all predictors (all predictors from Table 5 except PMT variables) show a smaller accuracy and a larger F_1 than the other groups, leading to a smaller gap between accuracy and F_1. The *All Predictors* group also has a much smaller error, likely caused by the larger starting pool of variables leading to similar models across the countries and prevention behaviors.

4.1.2 Variable Importance

The importance of individual variables within the set of 825 models was explored by evaluating the frequency that variables were selected in constructing DAGs. Fig. 6 shows the frequency that each variable in variable groups was selected as the most important variable through variable importance (black square). Fig. 6 also shows the frequency that each variable was included in DAGs (grey circle). *Age* and *Sex* were selected more frequently than the other variables in the demographics group, and *Age* was selected as the best predictor more often than any other predictor (40.0%). Worry about catching the coronavirus was an important variable in all three worry groups (*Current Worry*, *Future Worry - Personal*, and *Future Worry - Friends*). In models created using the *Current Worry* group, *Coronavirus* was selected as the best variable more often than any other variable was selected in all models. This shows that current worry about catching the coronavirus is an important predictor of prevention behavior. The *Cultural Cognition* group had relatively equal occurrence proportions across the seven questions, but *Prioritize Society* had the greatest proportion of best predictor and total selections. Both *Future Worry* groups exhibited similar trends with the same top five performing variables. Each variable in the *Trust* group performed similarly with the most included variable being trust in family members. These results show general importance of individual variables compared to others in each group and show an expected connection between worry about contracting COVID-19 and the prediction of prevention behaviors.

4.1.3 Model Performance for Individual Countries

The set of 825 individual prevention behavior models were analyzed to explore performance across 11 countries. Figs. 7 and 8 show the average accuracy of each variable group and prevention behavior for each country in the dataset. Models developed for the United States reported the highest average accuracy, 76.4%, and models for Australia reported the lowest average accuracy, 66.8%. Overall, Fig. 7 shows small differences between variable groups in most countries and small variations between countries. Larger differences are observed between prevention behaviors, with less outliers, which is due to the variations in the number of positive responses among prevention behaviors. Table 2 shows the reported deaths and cumulative cases per million people in each country at the end of the survey period. No visible correlation between case number and model performance was shown, potentially caused by selection bias in the survey respondents.

Experience with COVID-19 has been shown to be an important factor when predicting risk perception (Botzen et al., 2022; Dryhurst et al., 2020). Models that were created using the variable group *All Predictors* were analyzed to identify important variables. The proportion of models that include one or more of three COVID-19 experience variables (*COVID-19 Experience*, *Effect-Friends*, and *SARS*) is presented for each country in Table 3. These data show high importance of COVID-19 experience in predicting prevention behavior in all countries except the US and Spain. Italy and Spain had the highest number of cases per million people, exceeding the next closed country (Sweden) by approximately 350 cases. The two countries with the highest case counts per million people are the in the top 50% of country specific performance (Fig. 7). Australia, US, Japan, China, and the UK have similar, relatively low case counts and Mexico had the lowest case count at two cases per million people.

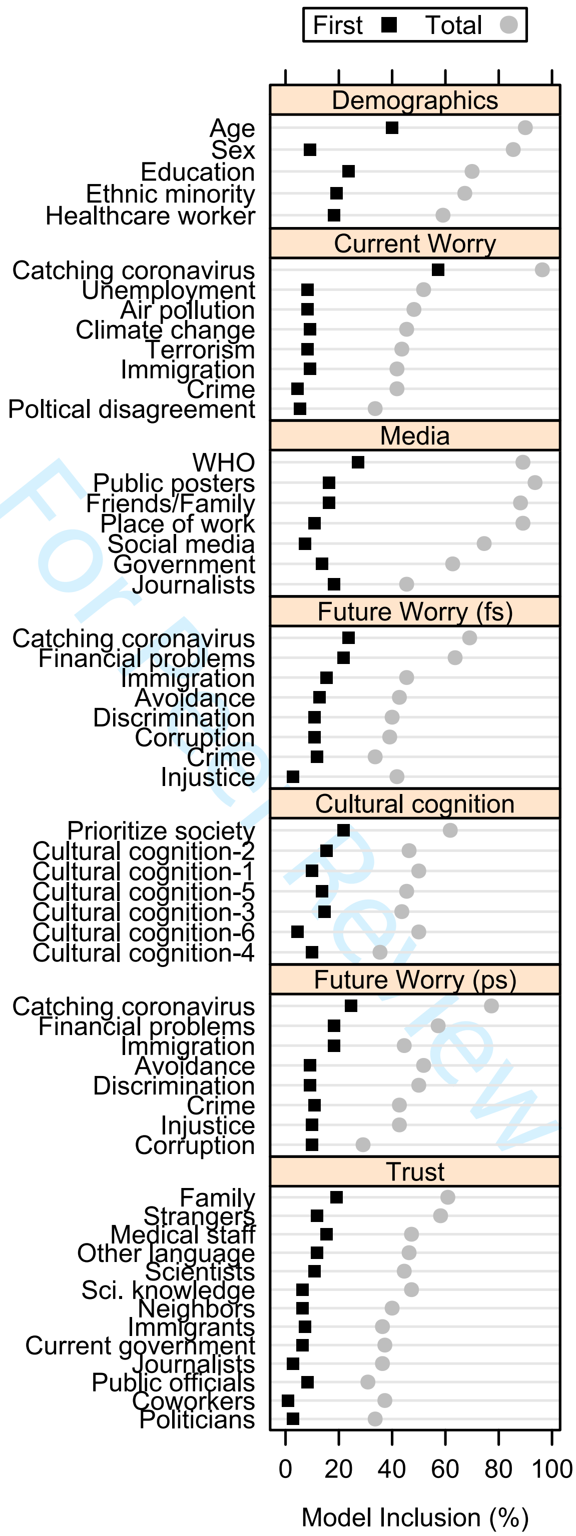

Figure 6: Frequency of individual predictors in model sets. Black squares refer to the percentage of models the specific predictor was selected as the best predictor in a model. Grey circles refer to the percentage of models the specific question was selected.

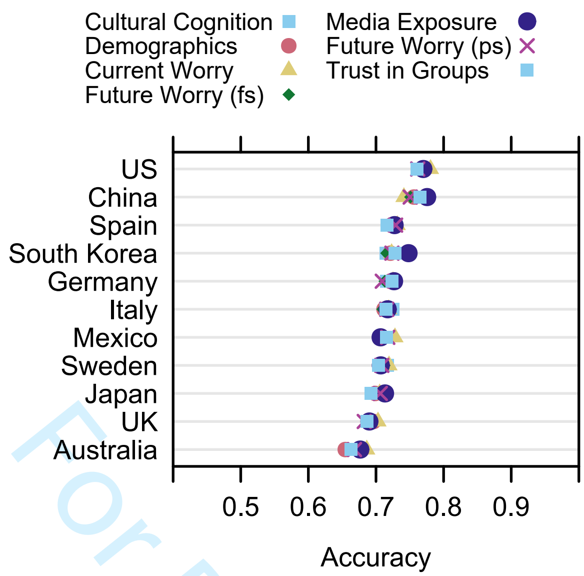

Figure 7: Average model accuracy for each variable group and each country. Countries are sorted based on average model accuracy.

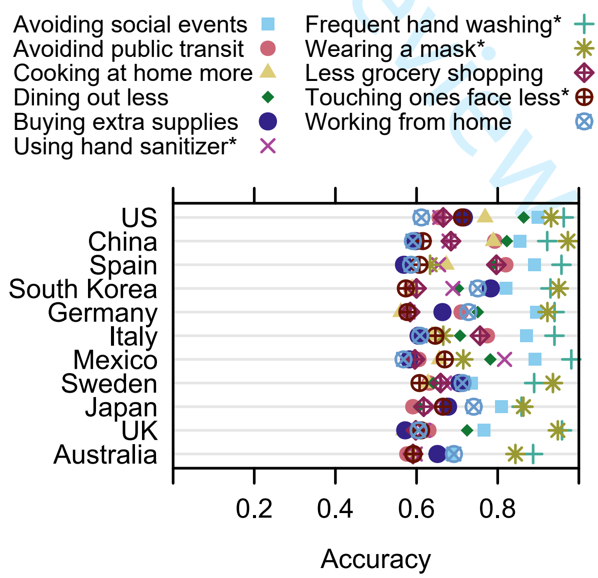

Figure 8: Average model accuracy for each prevention behavior and country. Countries are sorted from largest to smallest average accuracy, top to bottom.

Table 2: Reported deaths and cumulative cases per million people in each country on last date of survey collection. Data is sorted based on cumulative cases.

Country	Survey Dates	Reported Deaths	Cumulative Cases per Million People
Italy	22-24 March	5,476	1,002
Spain	22-25 March	1,772	605
Sweden	28-30 March	239	347
Germany	23-24 March	86	272
South Korea	9-11 April	204	201
UK	19 March	368	96
China	9-11 April	3,368	59
Japan	10-12 April	125	50
US	19-21 March	266	40
Australia	20 & 23 March	7	30
Mexico	21-27 March	2	2

Table 3: Proportion of models that included one or more of the three experience with COVID-19 variables (*COVID-19 Experience*, *Effect-Friends*, and *SARS*).

Country	Proportion of models with COVID-19 Experience
Italy	100.0%
Australia	81.8%
Germany	81.8%
South Korea	81.8%
Sweden	81.8%
Mexico	81.8%
UK	72.7%
Japan	72.7%
China	72.7%
Spain	63.6%
US	54.5%

4.2 Objective 2: Developing BBN Models to Predict Prevention Behaviors

The data-driven and hybrid methods were applied to develop models that predict composite predictands, *Limiting Exposure* and *Personal Protection*. Two models were developed for each predictand using the hybrid method, and one model was developed for each predictand using the data-driven method. Models that predict *Limiting Exposure* report similar accuracy and F_1 values, with a maximum accuracy of 69.8% (Table 4). Models that predict *Personal Protection* report higher performance at a maximum of 81.2% (Table 4). The data-driven models performed marginally better than the PMT based models for both predictands (4.5% improvement for *Limiting Exposure* and 2.9% improvement for *Personal Protection*).

The data-driven models for *Personal Protection* and *Limiting Exposure* are shown in Figs. 9 and 10, respectively. Many variables were selected for both models, including two *Longitude* variables, *SARS*, and *Effect-Friends*. These variables continue to show the importance of both direct and indirect experience with COVID-19 and general interest and worry about COVID-19. Fig. 11 shows the performance of data-driven models with each individual prevention behavior as the predictand. The four starred prevention behaviors correspond to the personal protection variable. The *frequent hand washing* variable explains most of the increased performance of the *Personal Protection* variable from Table 4. The *avoiding social events* variable had the best performance in the *Limiting Exposure* group, but still had an accuracy of 10% less than *frequent hand washing* in the *Personal Protection* group. Both the *frequent hand washing* and *avoiding social events* had datasets skewed heavily towards the positive category (93% and 82% yes, respectively), which aided in their increased accuracy and F_1 scores.

In executing the hybrid approach, two sub-BBNs were created using the data-driven approach for intermediate variables, *Threat Susceptibility* and *Threat Appraisal*. Sub-BBNs were used in models to predict *Limiting Exposure* and *Personal Protection*, as shown in Figure 4. Hybrid BBN models predicting *Personal Protection* are shown in Fig. 12 and 13. The sub-BBNs for intermediate variables *Threat Susceptibility* and *Threat Appraisal* shown in these figures reported accuracy values of 93.5% and 84.7%, respectively (Table 7). Both sub-BBNs include the predictor *COVID-19 Experience*, which refers to an individual either contracting COVID-19 or suspecting an infection. This shows the importance of experience with COVID-19 to predicting threat appraisal and corroborates previous studies linking personal experience of a threat with high risk perception (Botzen et al., 2022; Dryhurst et al., 2020).

The DAGs shown in Figures 12-15 demonstrate the use of the BBN structure to represent the PMT to predict adoption of *Personal Protection* and *Limiting Exposure* behaviors. In Figures 12 and 14, the PMT is represented by elements threat appraisal and coping appraisal, where threat appraisal is based on threat severity and threat susceptibility, and coping appraisal refers to societal efficacy and response

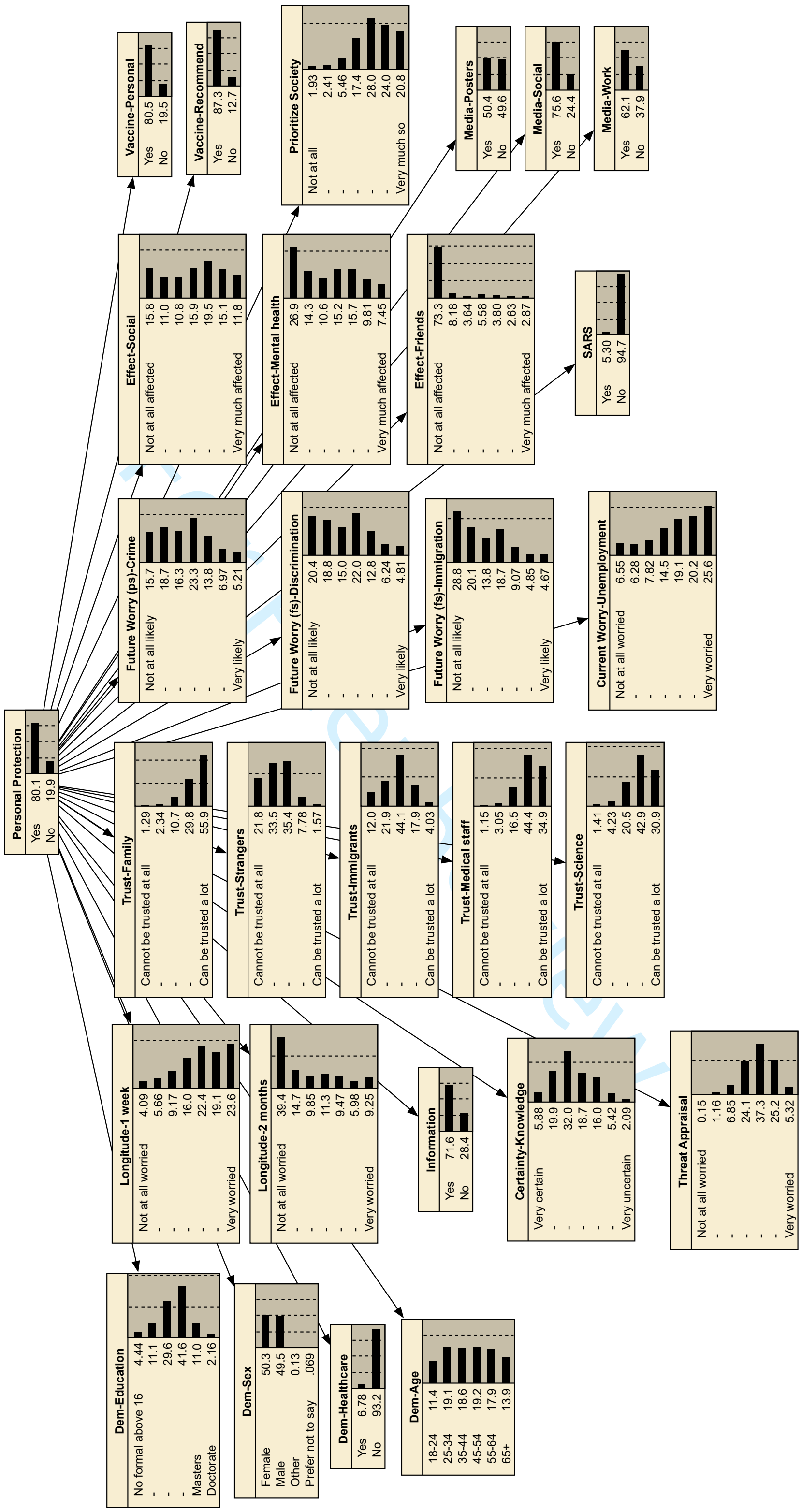

Figure 9: DAG representing the data-driven model to predict *Personal Protection*. Bars represent the responses present in the original dataset.

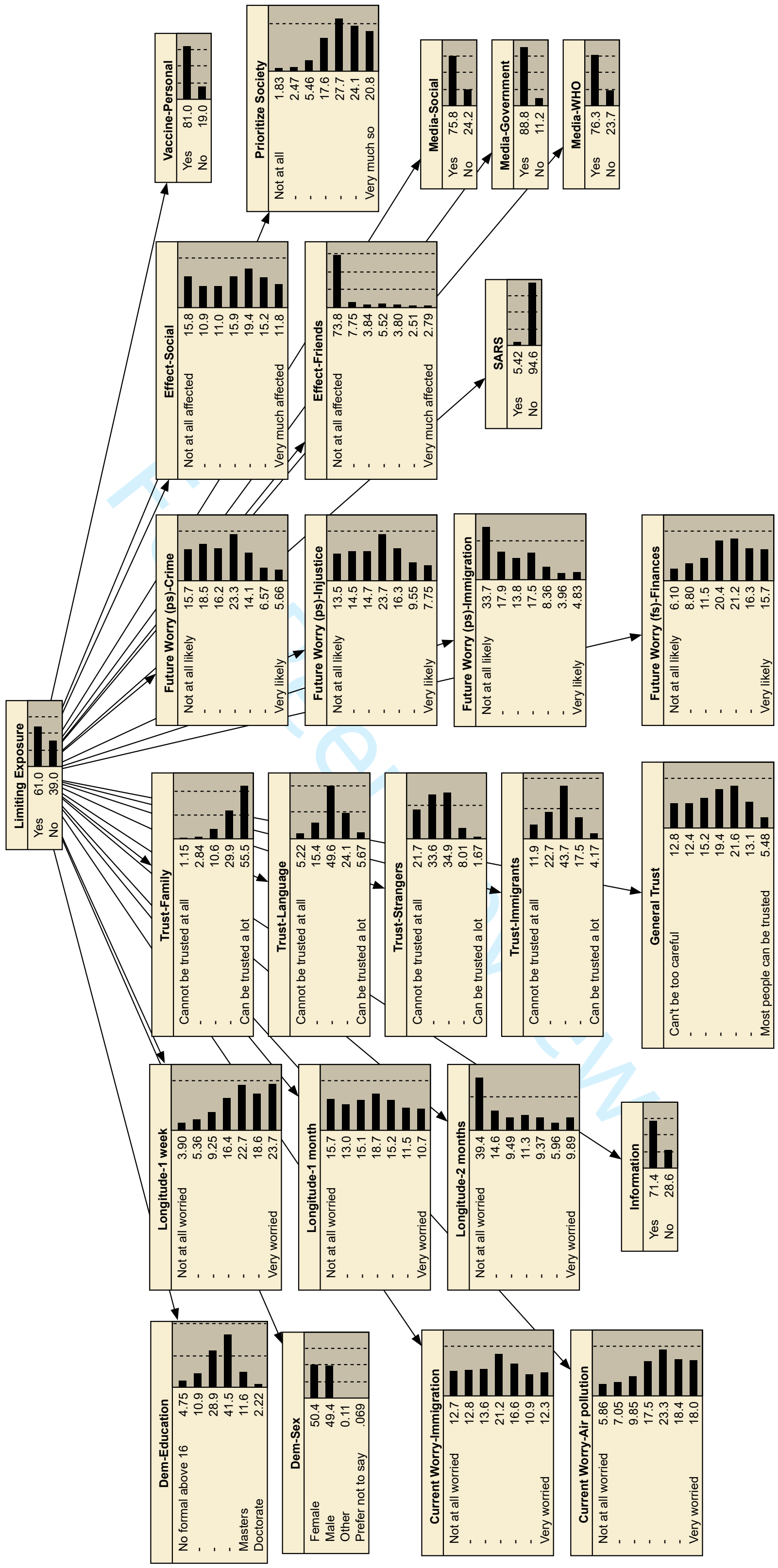

Figure 10: DAG representing the data-driven model to predict *Limiting Exposure*. Bars represent the responses present in the original dataset.

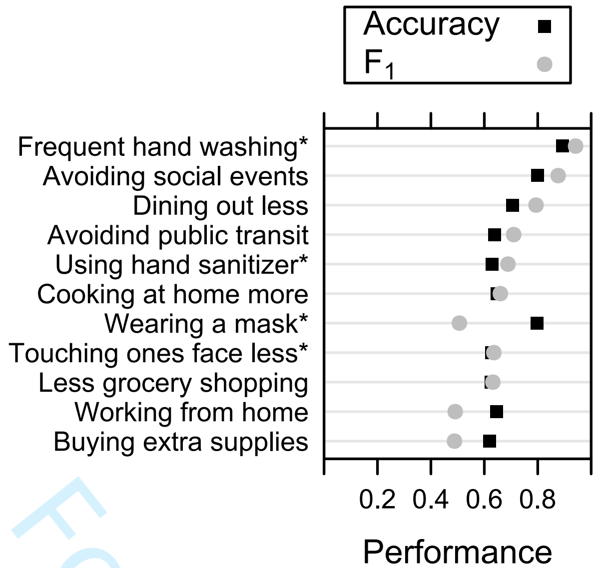

Figure 11: Average accuracy and F_1 of each prevention behavior using data-driven models. The four starred prevention behaviors correspond to the *Personal Protection* composite variable, and the remaining seven correspond to the *Limiting Exposure* composite variable.

efficacy. Variables that contribute to threat susceptibility, including experience with COVID-19, trust, worry, certainty, media sources, and experience with SARS, capture perceptions of risk as they are integrated within the PMT framework. In Figures 13 and 15, threat appraisal is not separated into the components threat severity and threat susceptibility. Threat appraisal comprises a larger group of variables because it must account for more variability in responses, due to the removal of the variable threat severity, compared with the BBN shown in Figures 12 and 14. The DAGs reflect relationships in the PMT which pose that individuals take actions if they perceive that action is useful (response and societal efficacy) for preventing severe consequences (threat appraisal).

The six BBN models generated using the data-driven and hybrid methods selected trust variables including trust in family, trust in strangers, and trust in immigrants. This implies general importance of trust in predicting both risk perception and prevention behaviors. Four of the five demographic variables were selected for the *Personal Protection* model, whereas only two were selected for the *Limiting Exposure* model. This shows increased importance of demographic variables in predicting *Personal Protection*. One difference between the two models is the selection of the *Threat Appraisal* variable for the *Personal Protection* model, which was included in both the hybrid and data-driven models.

The CPTs for the four hybrid BBNs were generated using statistical inference and are shown in Figures 16 and 17. The two forms of the hybrid models can be compared through Figures 16 and 17. The relationship between coping appraisal variables (response efficacy and societal efficacy) and adoption behaviors is similar

Table 4: Accuracy and F_1 values for the six prevention behavior prediction models.

Predictand	Predictors	Method	Prediction Accuracy	Prediction F_1
Limiting Exposure	Threat susceptibility, Threat severity, Response efficacy, Societal efficacy	Hybrid/PMT	66.8%	75.3%
	Threat appraisal, Response efficacy, Societal efficacy	Hybrid/PMT	65.4%	73.9%
	All predictors	Data-driven	69.8%	77.0%
Personal Protection	Threat susceptibility, Threat severity, Response efficacy, Societal efficacy	Hybrid/PMT	78.8%	88.0%
	Threat appraisal, Response efficacy, Societal efficacy	Hybrid/PMT	78.9%	88.1%
	All predictors	Data-driven	81.2%	89.3%

for both forms of the model. To demonstrate, the CPTs for *Response Efficacy* and *Societal Efficacy* are nearly identical for both the Threat Severity/Susceptibility model (Figure 16) and the Threat Appraisal model (Figure 17), when compared for each predictand. The trends that emerge in the CPTs provide a quantitative confirmation of the PMT: the CPTs show that as individuals perceive increasing levels of coping and threat appraisal, they increase their propensity to adopt, rather than do not adopt, across the four models. CPTs for the predictands *Limiting Exposure* and *Personal Protection* are similar but not identical, indicating that individuals may take similar decision-making processes for the two different prevention behaviors. There is one notable difference between CPTs when comparing *Limiting Exposure* with *Personal Protection*, as follows. For *Limiting Exposure*, perceptions of coping appraisal and threat appraisal variables need to reach higher levels before individuals are more likely to adopt rather than do not adopt, when compared with *Personal Protection*. This indicates that individuals must perceive higher levels of coping ability and threat before they will take actions around social distancing, as compared with actions that are protective, such as washing hands and wearing masks. This finding contributes to an understanding of the types of protective actions individuals are willing to take and their motivation around those actions during the COVID-19 pandemic.

5 DISCUSSION

This research addresses two objectives through the development of a BBN framework that is applied to generate models for predicting prevention behavior adoption during the COVID-19 pandemic.

5.1 Objective 1: Identifying Significant Variables in the Adoption of Prevention Behaviors

In achieving Objective 1, this research identified significant variables associated with the adoption of prevention behaviors. A set of 825 Naïve Bayes models were generated and analyzed for their ability to predict behaviors for individual preventative actions. These models were developed using a data-driven approach and were compared across country, prevention behavior, and group of predictands. Variables are assessed based on their ability to characterize prevention measure adoption to expand on past work in this space (Botzen et al., 2022; Bruin & Bennett, 2020; Dryhurst et al., 2020; Meier et al., 2020; Nudelman et al., 2022; Siegrist et al., 2021).

Results of this research demonstrated that experience with COVID-19 and trust were important predictors, and these findings are supported by published research. Trust has been linked to an individual's perception of and behavior during pandemic events and specifically during the COVID-19 pandemic (Fong & Chang, 2011; Siegrist et al., 2021; Siegrist & Zingg, 2014). Fong and Chang (2011) reported no correlation between trust and perceived community action in cities that were affected by the SARS epidemic in Taiwan, but did see a positive correlation in cities not affected. Siegrist et al. (2021) reported, on the other hand, that general trust is an important indicator of prevention behavior adoption during the COVID-19 pandemic. Results presented in this research show that trust is an important predictor for prevention behavior adoption and general risk perception in each country surveyed. This shows the importance of trust for predicting prevention behavior adoption in countries regardless of case count, supporting the claim that trust is an important predictor in both countries with and without widespread COVID-19 infection. Results also demonstrate that trust in strangers, immigrants, and family were the most important trust variables. These trust variables show no specific discretion between positive, negative, and neutral trust variables, as trust in strangers, immigrants, and family are negative, neutral, and positive, respectively (Fig. 9). This lack in specificity likely hints that the selection of these trust variables may be affected by confounding variables, such as the type of information that is being communicated and the effect of mis-information. The effects of confounding variables may lessen the intrinsic importance of these specific trust variables. Survey questions were designed without testing for the type of information that is communicated by information sources and create some limitations in the application of the modeling approach and analysis presented in this research.

Botzen et al. (2022) concluded that the demographic variables sex and age were important for prevention behavior prediction and that education status was not significant. This analysis showed a similar trend with age and sex being the top two most important variables in the demographics group, and education the third best variable. Education was included in both data-driven models and the threat appraisal model showing importance.

Previous experience with COVID-19 was identified as a strong indicator of prevention behavior adoption in three COVID-19 related studies (Botzen et al., 2022; Dryhurst et al., 2020; Meier et al., 2020). Results presented here also show similar results for COVID-19 experience, with the inclusion of at least one experience item (personal experience with COVID-19, family/friends experience with COVID-19, or previous experience with SARS/MERS) in all prevention behavior prediction models including the best performing model. This result applies for each country except Mexico and the US. Only 54.5% of models built with US data (6/11) contained one or more of the experience predictors, and 63.6% of models built with Mexico data (7/11) contained one or more experience predictors. The number of cases and deaths in Mexico was extremely low at the time the survey was administered, which may be the cause for low inclusion of COVID-19 experience items. The number of cases and deaths in the US was not as low as Mexico and were at similar levels to Australia, Japan, and China, each of which had high levels of COVID-19 experience inclusion. This research leveraged data collected during the initial wave of the COVID-19 pandemic, with case counts ranging from 2-1,002 per million individuals for different countries. Individuals from 11 countries were surveyed, providing a robust comparison of contextual responses to the pandemic. The studies by Botzen et al. (2022), Bruin and Bennett (2020), Jaspal et al. (2020), Meier et al. (2020), and Siegrist et al. (2021) isolated between one and three countries for survey and analysis with data from Germany, Italy, the Netherlands, Switzerland, the UK, and the US. Each of these studies were conducted independently, with a different set of survey questions but with similar research objectives, making comparison difficult. These previous studies showed high-level results such as the importance of trust and COVID-19 experience on predicting the adoption of prevention behaviors, but contained minor differences such as the strength of demographic predictors. Results presented here corroborate these findings and further explore the impact of case count on strength of predictors. This research identified that case count did not impact the performance of individual models or on the importance of variable group performance. No discernible pattern was found between case count or death count at survey collection and model performance, but the overall performance of models developed with data from individual countries differed by up to 14%. This suggests that prediction performance is dependent on the country and not the case count. An alternative hypothesis suggests that collection bias led to similar country-level results irrespective of case or death count. Future work could explore new frameworks to reduce collection bias in causal networks (Deffner et al., 2022).

Understanding the propensity of individuals to adopt prevention behaviors is needed to inform public policy during pandemics. This research found that trust, experience with COVID-19, age, sex, and education were among the most significant predictors. This insight can be used to guide the development of interventions and public messaging around preventative behaviors.

5.2 Objective 2: Developing BBN Models to Predict Prevention Behaviors

Predictive models were developed for composite prevention behavior variables to acheive Objective 2. Six BBN models were generated: three models predict the composite variable, personal protection, and three models predict the composite variable, limiting exposure. Of the three models built for each composite prevention behavior, one model was built with the data-driven approach and two were developed using the PMT framework to guide the development of the model structure. Results demonstrated that prevention behaviors can be predicted with accuracy using both PMT and data-driven approaches. Variables including trust, both direct and indirect experience with COVID-19, and general interest and worry about COVID-19 were important predictors for the BBNs. Demographics were more important variables in predicting personal protection behaviors when compared with limiting exposure behaviors. These insights can be used to develop messaging and interventions that are intended to target these variables, such as trust and demographics, or overcome the effects of these variables, such as lack of experience with COVID-19. Further research is needed to explore why these variables continue to appear as important across the objectives of this research and to apply the insight that was gained through this research to further refine the PMT and risk assessment framework.

A few limitations are present in the findings for Objective 2. First, the dataset that was developed by Dryhurst et al. (2020) was constructed to explore risk perceptions, but the dataset was not developed specifically to represent the constructs of the PMT. No questions from the dataset represented personal efficacy, and personal efficacy was excluded from the models presented in this research. Development of questionnaires in further research can explore questions that better test personal efficacy and explicitly represent PMT constructs. In addition, this research applies the Naïve Bayes classifier to develop DAGs because previous research identified that methods such as Tree Augmented Naïve (TAN) BBN performed similarly well but required a significant increase in the amount of computing time to train models (Fasaee et al., 2021). Using the Naïve Bayes classifier allows ready understanding of the contribution of variables to prevention behaviors, as shown in CPT and DAG figures. TAN includes additional relationship among predictors and may result in better predictive accuracy than Naïve Bayes; however, complex trees may be difficult to analyze and limit new insight that is gained to refine knowledge about the PMT. Further research can explore and assess other modeling methods, such as TAN, for accurately predicting behaviors and concisely representing PMT constructs.

Previous studies explored the prediction of social distancing behaviors, and several studies reported the application of PMT and BBN models to different facets of the COVID-19 pandemic; however, none have combined these techniques. Botzen et al. (2022) concluded that typical risk based heuristics and biases are applicable to COVID-19 risk perception, showing that previously defined risk cognition theories could be used to describe COVID-19 risk perception. This research builds on previous research by developing predictive models through integrating data-driven methods and expert knowledge of the PMT. PMT-based models are compared with models developed using a data-driven approach alone. Data-driven models provide a baseline for comparison by demonstrating the predictability that can be achieved through data collection and statistical inference alone. BBN models that were developed using expert knowledge about the PMT perform similarly to data-driven models, which strengthens the assumption that previously vetted risk cognition theories such as PMT can be successfully deployed in predicting prevention behavior adoption based on perceived risk of COVID-19. Data-driven models can be developed easily, but they rely on the relationships present in the data and may not generalize well to new datasets. This research develops an approach to develop BBN models based on the PMT; these structured models may perform more rigorously when they are applied to new datasets because they capture relationships that are inherent in selecting protective behaviors based on perceptions of risk. Further testing for new datasets is needed to demonstrate the power of the theoretical and data-driven approaches.

This research develops new methods for translating the PMT, which does not incorporate mathematics in its formulation, to a computational model. Individuals responded to survey questions that tested components of the PMT using qualitative response options (e.g., Likert scale items), and these responses were transposed and coded into quantitative data for analysis and modeling purposes. The methodology developed in this research provides an approach to integrate PMT within quantitative risk assessment frameworks, providing a new approach to understand and predict individual behavior in response to potential threats. In this research, new insight was gained to characterize threats of COVID-19 and subsequent evaluation of motivated responses around social distancing behavior. For example, DAGs demonstrated the relationships that lead to the adoption of prevention behaviors. This research used a quantitative approach to confirm that high perceptions of coping appraisal and threat appraisal contribute to the adoption of prevention behaviors, and higher levels of these perceptions are needed to adopt preventative behaviors that limit exposure through social distancing, compared with personal protective behaviors, such as hand-washing and wearing masks. Further research can explore the numeric relationships among variables and translate findings to refine the PMT. The integration of the PMT with classical risk assessment methodologies provides a nuanced approach for evaluating risk perception and behavior adoption. By considering individual assessment of threat severity, susceptibility, and perceived efficacy in coping with the threat, PMT introduces a psychological dimension to traditional risk assessment. This psychological perspective complements the quantitative analysis of a classical risk assessment, providing insight to the motivations that form individual decisions to adopt prevention behaviors. PMT is a valuable tool for capturing the complexities of human behavior in the face of risks because PMT is adaptable to new contexts and focuses on perception and motivation. PMT can bridge the gap between quantitative risk assessment and the intricacies of human decision-making positions, enhancing the predictive power of risk assessment and fostering a more comprehensive understanding of risk-related behaviors.

6 CONCLUSIONS

This research developed both data-driven and PMT based Bayesian belief network models to characterize and predict prevention measure adoption in response to the COVID-19 pandemic. Data-driven models performed marginally better than models developed using the PMT framework in predicting the adoption of prevention behavior, but all models performed well. Variable groups were tested for both group importance and variable importance. All groups showed equal importance in prediction power, but questions surrounding worry about catching the coronavirus/COVID-19 and experiencing financial problems were shown to be important individual variables. Direct or indirect experience with COVID-19 was also an important variable, but exhibited country-specific differences irrespective of case counts. In general, the prediction performance changed little by country and did not follow obvious trends based on case or death count at the time the survey was administered. This research developed models that generate heterogeneous decisions among individuals characterized by diverse demographics and experiences. Future research can operationalize diverse decision-making in a population of agents to simulate disease transmission with realistic simulation of the selection of prevention behaviors. An agent-based modeling approach that simulates decision-making using PMT and BBN models can simulate the expected uptake of prevention behaviors and impacts on disease transmission for scenarios that explore population characteristics and mitigation strategies.

The COVID-19 pandemic engendered a major push to collect and publish accessible datasets to expedite research seeking new ways to respond to the virus and mitigate future pandemics. This study leveraged recent literature explaining risk perception related to the COVID-19 pandemic and survey data collected from several countries to apply PMT to the adoption of prevention behaviors in response to the COVID-19 pandemic. This research applied the PMT to characterize the adoption of prevention behaviors during the COVID-19 pandemic, and future research can address existing limitations. First, data used to train these models were not collected within the construct of a PMT framework. As a result, not all PMT variables were represented in the dataset, and this research developed composite variables to represent the threat appraisal variable. Future work should be tailored to align questions and responses with PMT constructs. Second, many prevention behaviors were not represented uniformly in the dataset, and responses created skewed datasets. Future surveys should be distributed to better represent a distribution of responses. Finally, survey responses were collected at the onset of the pandemic and contain no information about vaccine uptake or changes in responses over time. Future work can include other risk perception models (e.g. secondary risk theory) to incorporate decisions around alternative prevention options, such as vaccines, and perceived tradeoffs (Cummings et al., 2021). This work developed models to characterize adoption of prevention behaviors and can be used within decision support systems to predict and mitigate disease transmission during pandemics.

DATA AVAILABILITY STATEMENT

Select models are available at: 10.5281/zenodo.10562900. Other models are available from the corresponding author upon reasonable request.

APPENDIX

Table 5: List of predictors after data cleaning. Question responses are included in the question column in brackets and Likert scale answers are denoted by a number representing the number of available responses (Likert responses are presented in detail in Table 6).

Variable	Question
Trust-Family	How much do you trust people in your family? [5]
Trust-Neighbors	How much do you trust people in your neighbourhood? [5]
Trust-Coworkers	How much do you trust people you work or study with? [5]
Trust-Language	How much do you trust people who speak a different language from you? [5]
Trust-Strangers	How much do you trust strangers? [5]
Trust-Immigrants	How much do you trust immigrants? [5]
Trust-Medical staff	How much do you trust medical doctors and nurses? [5]
Trust-Scientists	How much do you trust scientists? [5]
Trust-Politicians	How much do you trust politicians in the country you are living in? [5]
Trust-Journalists	How much do you trust journalists? [5]

Continued on next page

Table 5: List of predictors after data cleaning. Question responses are included in the question column in brackets and Likert scale answers are denoted by a number representing the number of available responses (Likert responses are presented in detail in Table 6). (Continued)

Variable	Question
Trust-Government	How much do you trust the current government of the country you are living in? [5]
Trust-Science	How much do you trust scientific knowledge? [5]
Trust-Public officials	How much do you trust civil servants or public officials in the country you are living in? [5]
Media-Posters	Have you come across information about coronavirus or COVID-19 from: official public posters. [1: yes, 2: no]
Media-Social	Have you come across information about coronavirus or COVID-19 from: social media or online blogs from individuals. [1: yes, 2: no]
Media-Journalist	Have you come across information about coronavirus or COVID-19 from: journalists and commentators in the media (TV, radio, newspapers). [1: yes, 2: no]
Media-Government	Have you come across information about coronavirus or COVID-19 from: government or official sources such as websites or public speeches/broadcasts within the country you are living in. [1: yes, 2: no]
Media-Work	Have you come across information about coronavirus or COVID-19 from: official messages from your place of work or education. [1: yes, 2: no]
Media-Friends	Have you come across information about coronavirus or COVID-19 from: friends and family. [1: yes, 2: no]
Media-WHO	Have you come across information about coronavirus or COVID-19 from: World Health Organisation. [1: yes, 2: no]
Current Worry-Climate	How worried are you personally about climate change at present? [7]
Current Worry-Politics	How worried are you personally about political disagreement at present? [7]

Continued on next page

Table 5: List of predictors after data cleaning. Question responses are included in the question column in brackets and Likert scale answers are denoted by a number representing the number of available responses (Likert responses are presented in detail in Table 6). (Continued)

Variable	Question
Current Worry-Immigration	How worried are you personally about immigration at present? [7]
Current Worry-Unemployment	How worried are you personally about unemployment at present? [7]
Current Worry-Terrorism	How worried are you personally about terrorism at present? [7]
Current Worry-Air pollution	How worried are you personally about air pollution at present? [7]
Current Worry-Crime	How worried are you personally about crime at present? [7]
Current Worry-Coronavirus (TA)	How worried are you personally about coronavirus/COVID-19 at present? [7]
Future Worry-Crime	How likely do you think it is that [you (ps) OR your friends and family in the country you are currently living in (fs)] will be directly affected by crime in the next 6 months? [7]
Future Worry-Finances	How likely do you think it is that [you (ps) OR your friends and family in the country you are currently living in (fs)] will be directly affected by financial problems in the next 6 months? [7]
Future Worry-Discrimination	How likely do you think it is that [you (ps) OR your friends and family in the country you are currently living in (fs)] will be directly affected by discrimination in the next 6 months? [7]
Future Worry-Avoidance	How likely do you think it is that [you (ps) OR your friends and family in the country you are currently living in (fs)] will be directly affected by antisocial behavior by others in the next 6 months? [7]
Future Worry-Injustice	How likely do you think it is that [you (ps) OR your friends and family in the country you are currently living in (fs)] will be directly affected by injustice in the next 6 months? [7]
Future Worry-Corruption	How likely do you think it is that [you (ps) OR your friends and family in the country you are currently living in (fs)] will be directly affected by corruption or fraud in the next 6 months? [7]

Continued on next page

Table 5: List of predictors after data cleaning. Question responses are included in the question column in brackets and Likert scale answers are denoted by a number representing the number of available responses (Likert responses are presented in detail in Table 6). (Continued)

Variable	Question
Future Worry-Immigration	How likely do you think it is that [you (ps) OR your friends and family in the country you are currently living in (fs)] will be directly affected by immigration in the next 6 months? [7]
Future Worry-Coronavirus (TS, TA)	How likely do you think it is that [you (ps) OR your friends and family in the country you are currently living in (fs)] will be directly affected by catching the coronavirus/COVID-19 in the next 6 months? [7]
Dem-Sex	What is your sex? [1: female, 2: male, 3: other, 4: prefer not to say]
Dem-Age	What is your age? [1: 18-24, 2: 25-34, 3: 35-44, 4: 45-54, 5: 55-64, 6: 65+]
Dem-Minority	Do you consider yourself to be part of a minority group within the country you are currently living in? [1: yes, 2: no, 4: prefer not to answer]
Dem-Healthcare	Are you a healthcare provider (e.g. doctor, nurse, paramedic, pharmacist, carer)? [1: yes, 2: no]
Dem-Education	Please indicate your highest educational qualification: [1: no formal above 16 to 5: Masters, 9: doctorate]
Cultural Cognition-1	The government interferes far too much in our everyday lives. [6]
Cultural Cognition-2	Sometimes government needs to make laws that keep people from hurting themselves. [6]
Cultural Cognition-3	It's not the government's business to try to protect people from themselves. [6]
Cultural Cognition-4	The government should stop telling people how to live their lives. [6]
Cultural Cognition-5	The government should do more to advance society's goals, even if that means limiting the freedom and choices of individuals.[6]
Cultural Cognition-6	Government should put limits on the choices individuals can make so they don't get in the way of what's good for society. [6]

Continued on next page

Table 5: List of predictors after data cleaning. Question responses are included in the question column in brackets and Likert scale answers are denoted by a number representing the number of available responses (Likert responses are presented in detail in Table 6). (Continued)

Variable	Question
Prioritize Society	To what extent do you think it's important to do things for the benefit of others and society even if they have some costs to you personally? [7]
COVID-19 Experience	Have you ever had, or thought you might have, the coronavirus/COVID-19? [9: unsure, 8: no, 3: I think I might have had it but am recovered, 2: I think I might have it now but not tested, 1: doctor suspected but tested negative, 4: doctor suspected but no test yet, 5: doctor confirmed and am still infected, 6: doctor confirmed but now test negative, 7: doctor confirmed but not been tested again]
Longitude-1 week	How worried were you about coronavirus 1 week ago? [7]
Longitude-1 month	How worried were you about coronavirus 1 month ago? [7]
Longitude-2 months	How worried were you about coronavirus 2 months ago? [7]
Effect-Financial	To what extent have you been affected by the coronavirus/COVID-19 in the following ways? - I have experienced financial difficulties as a result of the pandemic [7]
Effect-Social	To what extent have you been affected by the coronavirus/COVID-19 in the following ways? - I have experienced social difficulties as a result of the pandemic [7]
Effect-Mental health	To what extent have you been affected by the coronavirus/COVID-19 in the following ways? - I have experienced mental health difficulties as a result of the pandemic (e.g. increased anxiety) [7]
Effect-Friends	To what extent have you been affected by the coronavirus/COVID-19 in the following ways? - I have friends/family who have tested positive or died from the virus [7]

Continued on next page

Table 5: List of predictors after data cleaning. Question responses are included in the question column in brackets and Likert scale answers are denoted by a number representing the number of available responses (Likert responses are presented in detail in Table 6). (Continued)

Variable	Question
SARS	Have you personally been affected by a previous similar epidemic such as SARS (Severe Acute Respiratory Syndrome), MERS (Middle East Respiratory Syndrome) or Ebola? [1: yes, 2: no]
General Trust	Generally speaking, would you say most people can be trusted, or that you can't be too careful in dealing with people? [7]
Information	Have you sought out information specifically about coronavirus/COVID-19? [1: yes, 2: no]
Scientist Understanding	To what extent do you think scientists have a good understanding of the coronavirus/COVID-19? [7]
Certainty-Knowledge	How certain or uncertain do you think the following are: - The current scientific knowledge about the coronavirus/COVID-19? [7]
Certainty-Cases	How certain or uncertain do you think the following are: - The estimates of the number of cases of coronavirus/COVID-19 worldwide [7]
Vaccine-Personal	If a vaccine were to be available for the coronavirus/COVID-19 now: - Would you get vaccinated yourself? [1: yes, 2: no]
Vaccine-Recommend	If a vaccine were to be available for the coronavirus/COVID-19 now: - Would you recommend vulnerable friends/family to get vaccinated? [1: yes, 2: no]
Country-Affect (TA)	How much do you agree or disagree with the following statements? - The coronavirus/COVID-19 will NOT affect very many people in the country I'm currently living in [5]
Personal-Sick (TA)	How much do you agree or disagree with the following statements? - I will probably get sick with the coronavirus/ COVID-19 [5]
Threat Severity (PMT)	How much do you agree or disagree with the following statements? - Getting sick with the coronavirus/COVID-19 can be serious [5]

Continued on next page

Table 6: Likert scales used for selected questions.

Variable	Likert Scale
Trust	1 = Cannot be trusted at all to 5 = Can be trusted a lot
Current Worry	1 = not at all worried to 7 = very worried
Future Worry	1 = not at all likely to 7 = very likely
Cultural Cognition	1 = strongly disagree to 6 = Strongly agree
Prioritize Society	1 = not at all to 7 = very much so
Longitude	1 = not at all worried to 7 = very worried
Effect	1 = not at all affected to 7 = very much affected
General Trust	1 = Can't be too careful to 7 = Most people can be trusted
Scientist Understanding	1 = very limited understanding to 7 = very good understanding
Certainty	1 = very certain to 7 = very uncertain
Country-Affect and Personal-Sick	1 = strongly disagree to 5 = strongly agree
Threat Severity	1 = strongly disagree to 5 = strongly agree
Response Efficacy and Societal Efficacy	1 = not at all to 7 = very much

Table 7: Performance of threat appraisal and threat susceptibility models for training and testing data.

Predictand	Training					Testing				
	Accuracy	Recall	Precision	F_1	AUC	Accuracy	Recall	Precision	F_1	AUC
Threat susceptibility	93.8%	99.6%	94.1%	96.8%	66.9%	93.5%	99.6%	93.9%	96.6%	66.4%
Threat appraisal	86.3%	96.2%	88.2%	92.0%	82.9%	84.7%	96.8%	86.2%	91.2%	83.1%

Table 5: List of predictors after data cleaning. Question responses are included in the question column in brackets and Likert scale answers are denoted by a number representing the number of available responses (Likert responses are presented in detail in Table 6). (Continued)

Variable	Question
Response Efficacy (PMT)	To what extent do you feel that the personal actions you are taking to try to limit the spread of coronavirus make a difference? [7]
Societal Efficacy (PMT)	To what extent do you feel the actions that your country is taking to limit the spread of coronavirus make a difference? [7]

References

Aliyu, S., Zakari, A. S., Adeyanju, I., & Ajoge, N. S. (2021). A Bayesian Network Model for the Prognosis of the Novel Coronavirus (COVID-19). In *Computational Science and Its Applications – ICCSA 2021* (pp. 127–140). Springer International Publishing.

Arrizza, A. M., & Caimo, A. (2021). Bayesian dynamic network actor models with application to South Korean COVID-19 patient movement data. *Statistical Methods & Applications*, *30*(5), 1465–1483. https://doi.org/10.1007/s10260-021-00599-x

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., Drury, J., Dube, O., Ellemers, N., Finkel, E. J., Fowler, J. H., Gelfand, M., Han, S., Haslam, S. A., Jetten, J., … Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, *4*(5), 460–471. https://doi.org/10.1038/s41562-020-0884-z

Botzen, W. J. W., Duijndam, S. J., Robinson, P. J., & van Beukering, P. (2022). Behavioral biases and heuristics in perceptions of COVID-19 risks and prevention decisions. *Risk Analysis*, *n/a*(n/a). https://doi.org/10.1111/risa.13882

Bruin, W. B. d., & Bennett, D. (2020). Relationships Between Initial COVID-19 Risk Perceptions and Protective Health Behaviors: A National Survey. *American Journal of Preventive Medicine*, *59*(2), 157–167. https://doi.org/10.1016/j.amepre.2020.05.001

Chu, D. K., Akl, E. A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H. J., Chu, D. K., Akl, E. A., El-harakeh, A., Bognanni, A., Lotfi, T., Loeb, M., Hajizadeh, A., Bak, A., Izcovich, A., Cuello-Garcia, C. A., Chen, C., Harris, D. J., Borowiack, E., … Schünemann, H. J. (2020). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: A systematic review and meta-analysis. *The Lancet*, *395*(10242), 1973–1987. https://doi.org/10.1016/S0140-6736(20)31142-9

Conner, M. (2010). Cognitive Determinants of Health Behavior. In A. Steptoe (Ed.), *Handbook of Behavioral Medicine* (pp. 19–30). Springer New York.

Cummings, C. L., Rosenthal, S., & Kong, W. Y. (2021). Secondary Risk Theory: Validation of a Novel Model of Protection Motivation. *Risk Analysis*, *41*(1), 204–220. https://doi.org/10.1111/risa.13573

Deffner, D., Rohrer, J. M., & McElreath, R. (2022). A Causal Framework for Cross-Cultural Generalizability. *Advances in Methods and Practices in Psychological Science*, *5*(3).

Dryhurst, S., Schneider, C. R., Kerr, J., Freeman, A. L. J., Recchia, G., van der Bles, A. M., Spiegelhalter, D., & van der Linden, S. (2020). Risk perceptions of COVID-19 around the world. *Journal of Risk Research*, *23*(7-8), 994–1006. https://doi.org/10.1080/13669877.2020.1758193

Epstein, J. M., Parker, J., Cummings, D., & Hammond, R. A. (2008). Coupled Contagion Dynamics of Fear and Disease: Mathematical and Computational Explorations (A. P. Galvani, Ed.). *PLoS ONE*, *3*(12), e3955. https://doi.org/10.1371/journal.pone.0003955

Fasaee, M. A. K., Berglund, E., Pieper, K. J., Ling, E., Benham, B., & Edwards, M. (2021). Developing a framework for classifying water lead levels at private drinking water systems: A Bayesian Belief Network approach. *Water Research*, *189*, 116641. https://doi.org/10.1016/j.watres.2020.116641

Fisher, J. J., Almanza, B. A., Behnke, C., Nelson, D. C., & Neal, J. (2018). Norovirus on cruise ships: Motivation for handwashing? *International Journal of Hospitality Management*, *75*, 10–17. https://doi.org/10.1016/j.ijhm.2018.02.001

Fong, E., & Chang, L.-y. (2011). Community Under Stress: Trust, Reciprocity, and Community Collective Efficacy During SARS Outbreak. *Journal of Community Health*, *36*(5), 797–810. https://doi.org/10.1007/s10900-011-9378-2

Funk, S., Gilad, E., Watkins, C., & Jansen, V. (2009). The spread of awareness and its impact on epidemic outbreaks.

Ghahremani, L., Faryabi, R., & Kaveh, M. H. (2014). Effect of Health Education Based on the Protection Motivation Theory on Malaria Preventive Behaviors in Rural Households of Kerman, Iran. *International Journal of Preventive Medicine*, *5*(4), 9.

Hale, T., Cameron-Blake, E., Di Folco, M., Furst, R., Green, K., Philips, T., Sudarmawan, A., Tatlow, H., & Zha, H. (2022). What have we learned from tracking every government policy on COVID-19 for the past two years?

Heckerman, D., & Wellman, M. P. (1995). Bayesian networks. *Communications of the ACM*, *38*(3), 27–30. https://doi.org/10.1145/203330.203336

Jaspal, R., Lopes, B., & Lopes, P. (2020). Predicting social distancing and compulsive buying behaviours in response to COVID-19 in a United Kingdom sample. *Cogent Psychology*, *7*(1). https://doi.org/https://doi.org/10.1080/23311908.2020.1800924

Kahan, D. M. (2012). Cultural Cognition as a Conception of the Cultural Theory of Risk. In S. Roeser, R. Hillerbrand, P. Sandin, & M. Peterson (Eds.), *Handbook of Risk Theory* (pp. 725–759). Springer Netherlands.

Kelly (Letcher), R. A., Jakeman, A. J., Barreteau, O., Borsuk, M. E., ElSawah, S., Hamilton, S. H., Henriksen, H. J., Kuikka, S., Maier, H. R., Rizzoli, A. E., van Delden, H., & Voinov, A. A. (2013). Selecting among five common modelling approaches for integrated environmental assessment and management. *Environmental Modelling & Software*, *47*, 159–181. https://doi.org/10.1016/j.envsoft.2013.05.005

Koh, D. (2020). COVID-19 lockdowns throughout the world. *Occupational Medicine (Oxford, England)*, kqaa073. https://doi.org/10.1093/occmed/kqaa073

Ma, Q.-X., Shan, H., Zhang, H.-L., Li, G.-M., Yang, R.-M., & Chen, J.-M. (2020). Potential utilities of mask-wearing and instant hand hygiene for fighting SARS-CoV-2. *Journal of Medical Virology*, *92*(9), 1567–1571. https://doi.org/10.1002/jmv.25805

Meier, K., Glatz, T., Guijt, M. C., Piccininni, M., van der Meulen, M., Atmar, K., Jolink, A.-T. C., Kurth, T., Rohmann, J. L., Zamanipoor Najafabadi, A. H., & on behalf of the COVID-19 Survey Study group. (2020). Public perspectives on protective measures during the COVID-19 pandemic in the Netherlands, Germany and Italy: A survey study (A. M. Samy, Ed.). *PLOS ONE*, *15*(8), e0236917. https://doi.org/10.1371/journal.pone.0236917

Moeini, B., Ezati, E., Barati, M., Rezapur-Shahkolai, F., Mohammad Gholi Mezerji, N., & Afshari, M. (2019). Skin Cancer Preventive Behaviors in Iranian Farmers: Applying Protection Motivation Theory. *Workplace Health & Safety*, *67*(5), 231–240. https://doi.org/10.1177/2165079918796850

Nudelman, G., Kamble, S. V., & Otto, K. (2022). Using Protection Motivation Theory to Predict Adherence to COVID-19 Behavioral Guidelines. *Behavioral Medicine*, 1–10. https://doi.org/10.1080/08964289.2021.2021383

Qazi, A., Simsekler, M. C. E., & Gaudenzi, B. (2022). Prioritizing Multidimensional Interdependent Factors Influencing COVID-19 Risk. *Risk Analysis*, *42*(1), 143–161. https://doi.org/10.1111/risa.13841

Reluga, T. C. (2010). Game Theory of Social Distancing in Response to an Epidemic (C. T. Bergstrom, Ed.). *PLoS Computational Biology*, *6*(5), e1000793. https://doi.org/10.1371/journal.pcbi.1000793

Rogers, R. W. (1975). A Protection Motivation Theory of Fear Appeals and Attitude Change1. *The Journal of Psychology*, *91*(1), 93–114. https://doi.org/10.1080/00223980.1975.9915803

Scutari, M. (2010). Learning Bayesian Networks with the bnlearn R Package. *arXiv:0908.3817 [stat]*.

Siegrist, M., Luchsinger, L., & Bearth, A. (2021). The Impact of Trust and Risk Perception on the Acceptance of Measures to Reduce COVID-19 Cases. *Risk Analysis*, *41*(5), 787–800. https://doi.org/10.1111/risa.13675

Siegrist, M., & Zingg, A. (2014). The role of public trust during pandemics: Implications for crisis communication. *European Psychologist*, *19*(1), 23–32. https://doi.org/10.1027/1016-9040/a000169

Timpka, T., Spreco, A., Gursky, E., Eriksson, O., Dahlström, Ö., Strömgren, M., Ekberg, J., Pilemalm, S., Karlsson, D., Hinkula, J., & Holm, E. (2014). Intentions to Perform Non-Pharmaceutical Protective Behaviors during Influenza Outbreaks in Sweden: A Cross-Sectional Study following a Mass Vaccination Campaign (J. W. Tang, Ed.). *PLoS ONE*, *9*(3), e91060. https://doi.org/10.1371/journal.pone.0091060

WHO. (2023). Advice for the public on COVID-19 – World Health Organization. Retrieved August 30, 2023, from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public

Wu, Y., Foley, D., Ramsay, J., Woodberry, O., Mascaro, S., Nicholson, A. E., & Snelling, T. (2021). Bridging the gaps in test interpretation of SARS-CoV-2 through Bayesian network modelling. *Epidemiology and Infection*, *149*, e166. https://doi.org/10.1017/S0950268821001357

Xu, Y., & Chen, X. (2016). Protection motivation theory and cigarette smoking among vocational high school students in China: A cusp catastrophe modeling analysis. *Global Health Research and Policy*, *1*(1), 3. https://doi.org/10.1186/s41256-016-0004-9

7 Supplemental Information

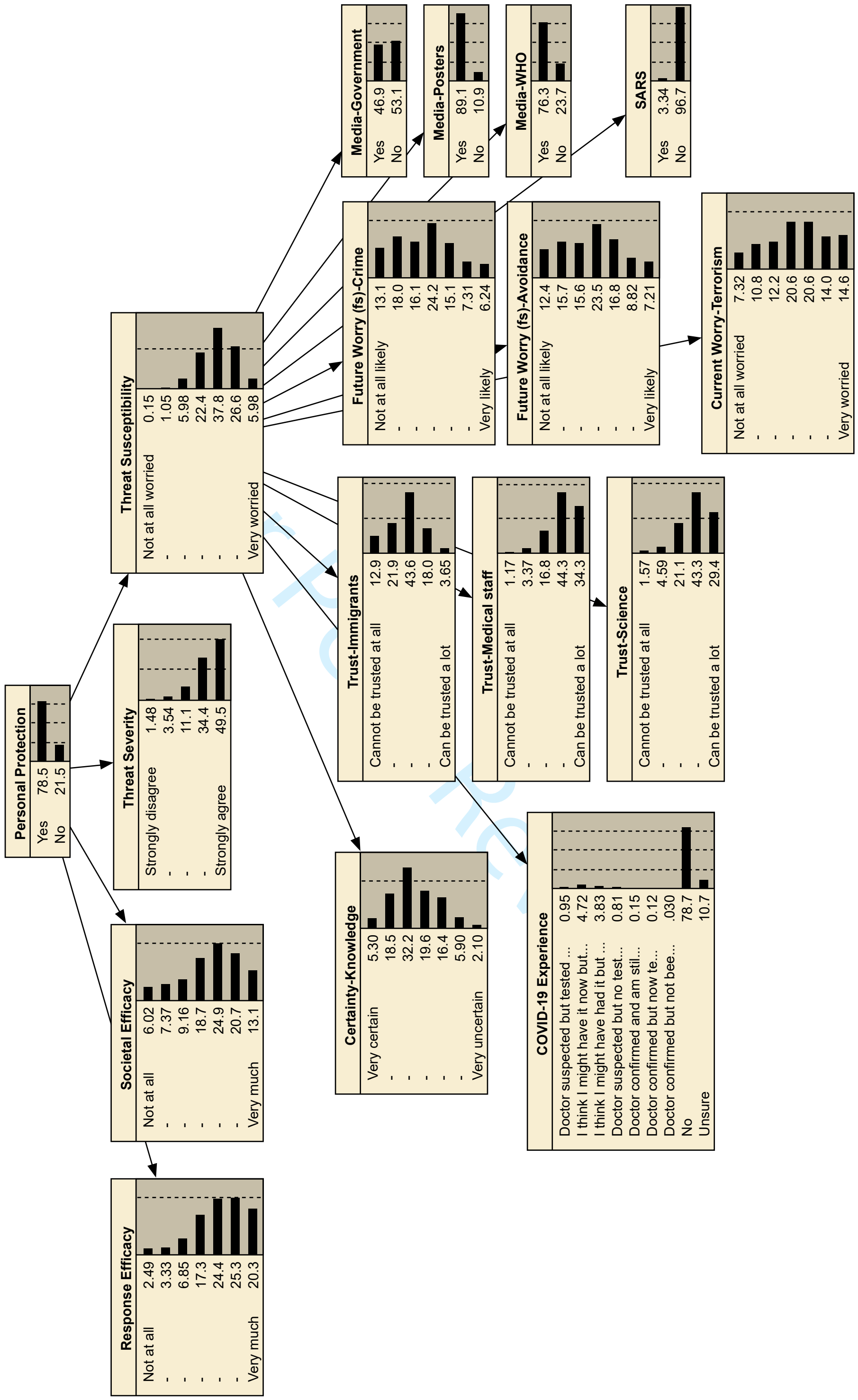

Figure 12: DAG representing the hybrid model for personal protection predictand. Threat appraisal is represented by threat two variables, threat susceptibility and threat severity. Bars represent the responses present in the original dataset.

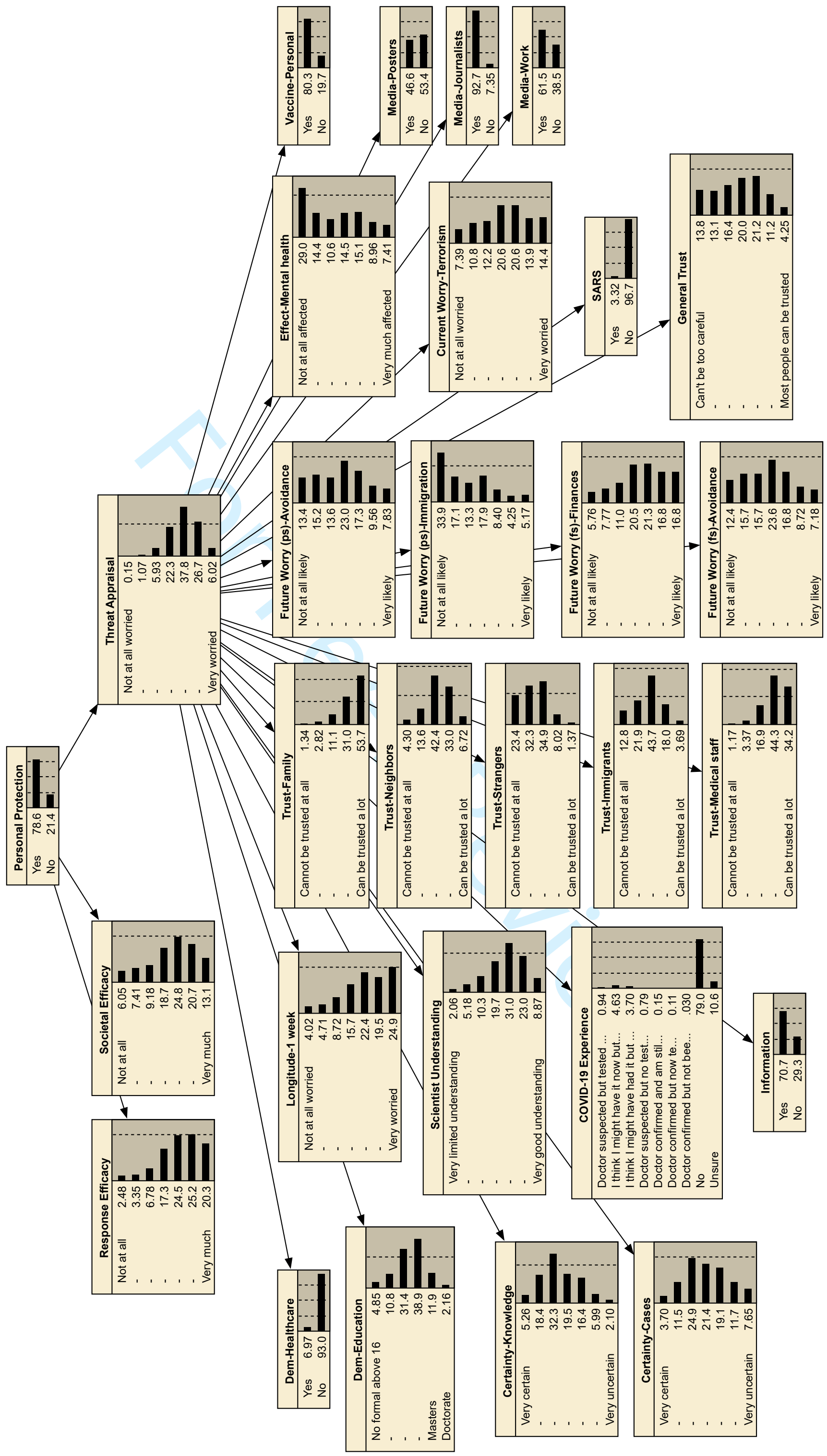

Figure 13: DAG representing the hybrid model for personal protection predictand. Threat appraisal is represented as one variable. Bars represent the responses present in the original dataset.

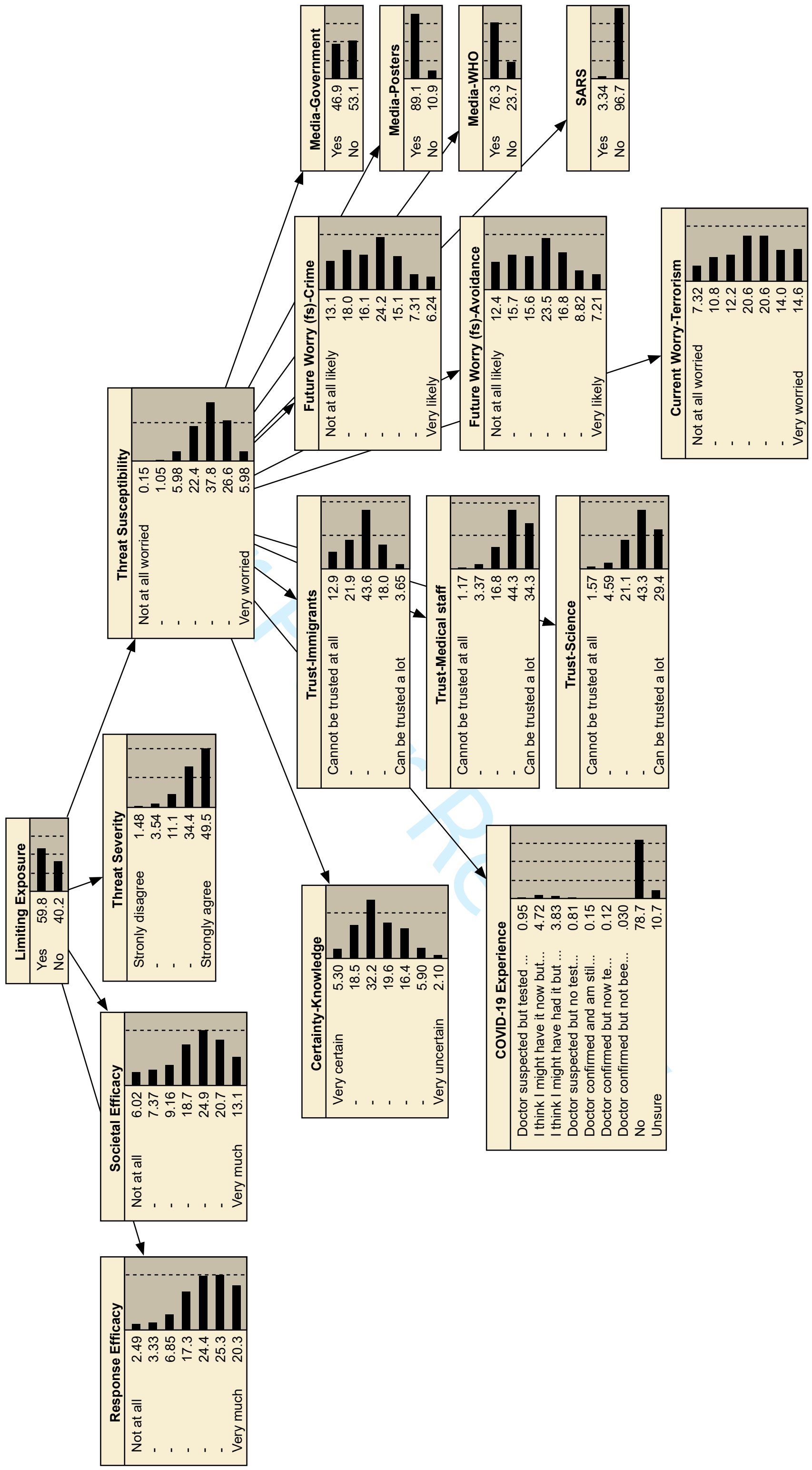

Figure 14: DAG representing the hybrid model for limiting exposure predictand. Threat appraisal is represented by threat two variables, threat susceptibility and threat severity. Bars represent the responses present in the original dataset. Bars represent the responses present in the original dataset.

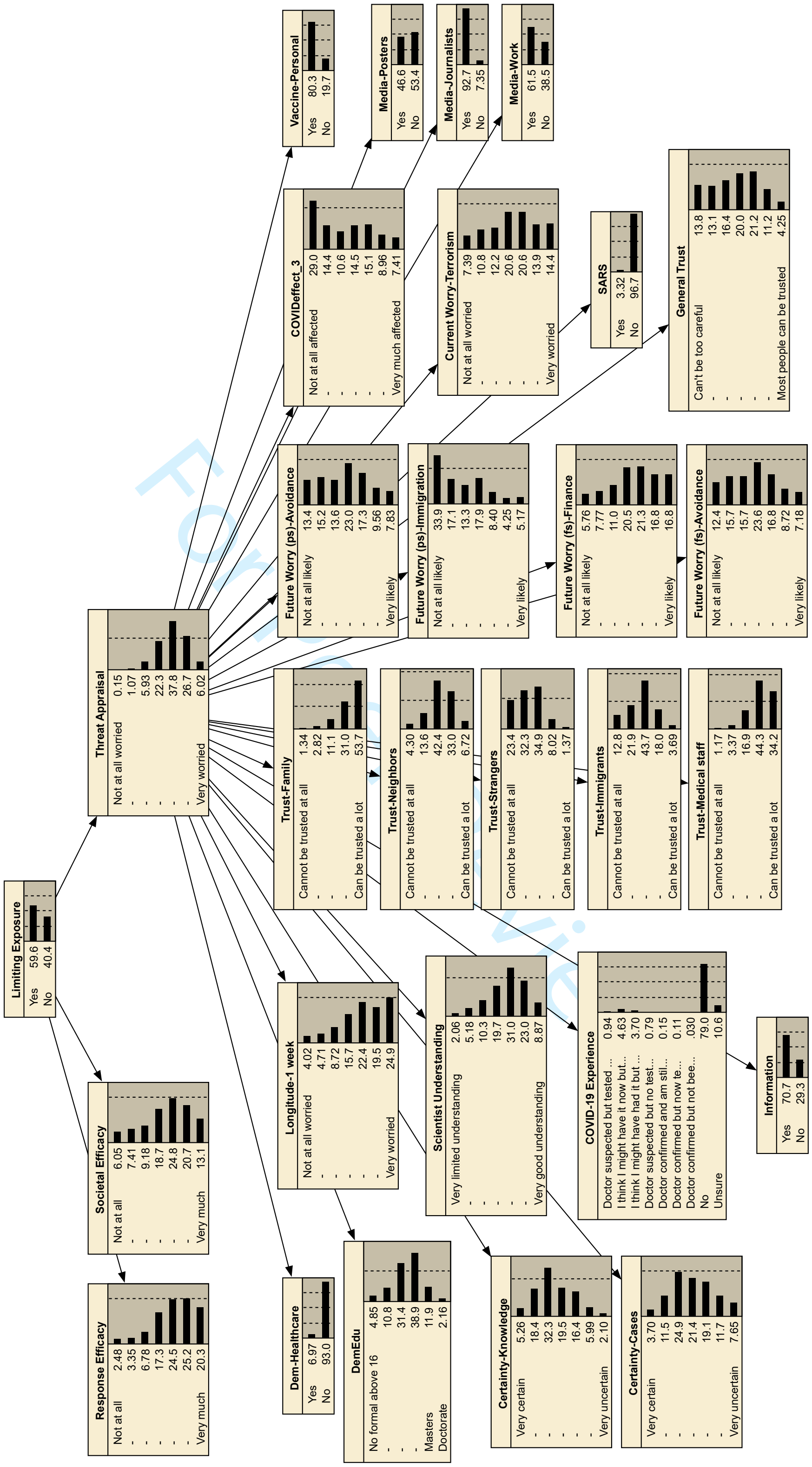

Figure 15: DAG representing the hybrid model for limiting exposure predictand. Threat appraisal is represented as one variable. Bars represent the responses present in the original dataset. Bars represent the responses present in the original dataset.

Personal Protection

Response Efficacy

	Do not adopt	Adopt
Not at all	0.04	0.02
-	0.06	0.03
-	0.10	0.06
-	0.21	0.16
-	0.23	0.25
-	0.20	0.27
Very much	0.15	0.22

Societal Efficacy

	Do not adopt	Adopt
Not at all	0.07	0.06
-	0.07	0.07
-	0.12	0.08
-	0.21	0.18
-	0.25	0.25
-	0.18	0.21
Very much	0.10	0.14

Threat Severity

	Do not adopt	Adopt
Strongly disagree	0.03	0.01
-	0.05	0.03
-	0.14	0.10
-	0.35	0.34
Strongly agree	0.42	0.51

Threat Susceptibility

	Do not adopt	Adopt
Not at all worried	0.00	0.00
-	0.02	0.01
-	0.10	0.05
-	0.29	0.21
-	0.36	0.38
-	0.17	0.29
Very worried	0.04	0.06

(a)

Limiting Exposure

Response Efficacy

	Do not adopt	Adopt
Not at all	0.04	0.02
-	0.05	0.02
-	0.09	0.05
-	0.22	0.14
-	0.26	0.23
-	0.21	0.28
Very much	0.13	0.25

Societal Efficacy

	Do not adopt	Adopt
Not at all	0.06	0.06
-	0.07	0.08
-	0.10	0.08
-	0.21	0.17
-	0.26	0.24
-	0.18	0.22
Very much	0.11	0.15

Threat Severity

	Do not adopt	Adopt
Strongly disagree	0.02	0.01
-	0.06	0.02
-	0.18	0.07
-	0.39	0.31
Strongly agree	0.35	0.59

Threat Susceptibility

	Do not adopt	Adopt
Not at all worried	0.00	0.00
-	0.02	0.00
-	0.09	0.04
-	0.29	0.18
-	0.38	0.38
-	0.18	0.32
Very worried	0.03	0.08

(b)

Figure 16: CPTs for PMT nodes in hybrid BBNs for (a) Personal Protection (Figure 12) and (b) Limiting Exposure (Figure 14). Grey boxes indicate higher contribution of response to adopt or do not adopt decisions.

Personal Protection

Response Efficacy

	Do not adopt	Adopt
Not at all	0.05	0.02
-	0.06	0.03
-	0.10	0.06
-	0.21	0.16
-	0.23	0.25
-	0.20	0.27
Very much	0.16	0.22

Societal Efficacy

	Do not adopt	Adopt
Not at all	0.07	0.06
-	0.08	0.07
-	0.12	0.08
-	0.21	0.18
-	0.24	0.25
-	0.18	0.21
Very much	0.10	0.14

Threat Appraisal

	Do not adopt	Adopt
Not at all worried	0.00	0.00
-	0.03	0.01
-	0.10	0.05
-	0.29	0.21
-	0.36	0.38
-	0.17	0.29
Very worried	0.04	0.06

(a)

Limiting Exposure

Response Efficacy

	Do not adopt	Adopt
Not at all	0.04	0.02
-	0.05	0.02
-	0.09	0.05
-	0.22	0.14
-	0.26	0.23
-	0.21	0.28
Very much	0.13	0.25

Societal Efficacy

	Do not adopt	Adopt
Not at all	0.06	0.06
-	0.07	0.07
-	0.10	0.08
-	0.21	0.17
-	0.26	0.24
-	0.18	0.22
Very much	0.11	0.14

Threat Appraisal

	Do not adopt	Adopt
Not at all worried	0.00	0.00
-	0.02	0.00
-	0.09	0.04
-	0.29	0.18
-	0.38	0.38
-	0.18	0.32
Very worried	0.03	0.08

(b)

Figure 17: CPTs for PMT nodes in hybrid BBNs for (a) Personal Protection (Figure 13) and (b) Limiting Exposure (Figure 15). Grey boxes indicate higher contribution of response to adopt or do not adopt decisions.