Title
Estimating attitudes toward vaccination: A Bayesian framework

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Madsen, Jens Koed
Pilditch, Toby D
Zagala, Helena
et al.

Publication Date
2023

Peer reviewed

Estimating attitudes toward vaccination: A Bayesian framework

Jens Koed Madsen (j.madsen2@lse.ac.uk)
Department of Psychological and Behavioural Science,
London School of Economics and Political Science, Houghton Street, WC2A 2AE, London, UK

Diane Cordier (D.Cordier@lse.ac.uk)
Department of Psychological and Behavioural Science,
London School of Economics and Political Science, Houghton Street, WC2A 2AE, London, UK

Toby D. Pilditch (t.pilditch@ucl.ac.uk)
Department of Experimental Psychology, UCL
26 Bedford Way, London WC1H 0AP

Helen Zagala (H.Zagala@lse.ac.uk)
Department of Psychological and Behavioural Science,
London School of Economics and Political Science, Houghton Street, WC2A 2AE, London, UK

Abstract

Vaccines are among the best tools to limit the spread of preventable diseases. yet, recent years have seen a rise in anti-vaccination sentiments for vaccines against COVID-19, the MMR, and more. It is critical to understand the factors that influence whether a person will accept or reject a vaccine for a given disease. This paper tests a Bayesian model to predict attitudes toward vaccines. with five factors (subjective beliefs concerning danger of illness, safety of the vaccine, prevalence of the disease, perceived social norm, and governmental recommendation).

To parameterize the model, Study 1 elicits the full conditional probability table while Study 2 tests model predictions by eliciting people's priors for COVID-19 and the common flu. We find a good fit between predictions and observations, accounting for 53% and 44% of the variance. This suggests the usefulness of a formal model to capture people's beliefs about vaccination.

Keywords: Vaccination; Attitudes; Bayesian Modelling; Trust; COVID-19

Introduction

Vaccines are a proven and effective way to protect against many diseases. If enough people are vaccinated, it can help to create herd immunity, which protects those who are unable to receive vaccines. Understanding people's attitudes towards vaccination is key in preventing the outbreak of curable diseases, as some refuse to take needed vaccines.

COVID-19 demonstrated the need to understand people's vaccination attitudes. Since 2020, worldwide protests have erupted in response to COVID-19 vaccination mandates–in several countries. Indeed, the WHO argues that the pandemic exposed the urgent need for a deeper understanding of the causes for the anti-vaccination sentiment and its consequences on global health (*World Health Organization*, 2019). At the onset of COVID, there was understandably confusion on the safety of the vaccine, the causes of COVID, and the consequences of catching COVID. In 2021, this meant that skepticism of the vaccine was present in many countries. For example, data collected in November-December 2020 in the USA showed that 50% would take the vaccine, 40% would wait and learn, and 10% did not want it (Salmon et al., 2021). Eventually, as of January 2023, according to the CDC, 85.4% have accepted one vaccination dose in the USA and 73.2% have accepted two doses. With an adult population (18+) of ca. 209 million, this refusal rate means that 30-32 million Americans have not been vaccinated. This is in line with the UK where 14% of respondents reported unwillingness to vaccinate (Paul et al., 2021) – although the eventual vaccine uptake in the UK, according to government data, was higher with 93.6% of people aged 12+ taking one dose and 88.3% accepting two.

Vaccine hesitancy extends beyond COVID-19 responses as hesitancy for preventable diseases (VPDs) have resurfaced (Benecke & DeYoung, 2019). From 2017 to 2018 measles cases in Europe tripled (IBMS, 2019) and 2019 marked an emergency-level outbreak of measles worldwide (Wamsley, 2019). In the early 2000s, controversies around measles, mumps, and rubella (MMR) linking MMR vaccination with inflammatory bowel disease and autism planted permanent uncertainty around the safety of vaccines. Although research has confirmed the falsity of these claims, MMR vaccination rates have dropped, fueling the re-emergence of these highly infectious diseases in communities with low vaccination rates (Serpell & Green, 2006; Abad & Safdar, 2015).

Vaccine refusal can have several deleterious consequences. First, people who chose to forego the vaccine are left more vulnerable to the disease in question (assuming, of course, that the proposed vaccine is safe and effective). Beyond the impact on the health of individual people, attitudes toward vaccination also carries social judgments. A cross-national study of 21 countries concerning COVID-19 vaccination

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

show that vaccinated people had antipathy for unvaccinated people while unvaccinated people did not show greater antipathy for vaccinated people. This suggests a moral component and a social norm is mixed with feelings of safety when evaluating personal choices and the choices of others (Bor et al., 2022). Concurrent (and perhaps related) with this, hesitancy may put additional strain on societal resources if a person without a vaccine is more likely to be hospitalized. As such, there are individual and social aspects at play when considering attitudes toward vaccination.

As described in the following, studies have explored how attitudes toward vaccination are shaped and key components that influence these attitudes. While these studies provide key insights into vaccine hesitancy, a formal mapping of vaccine-related attitudes has yet to be established. That is, a causal model that describes elements that should lead a person to accept or refuse vaccination for a given disease. Indeed, most efforts to investigate anti-vaccination attitudes focus on attitudes towards disease-specific vaccines (e.g., COVID, MMR, or some other disease). In this paper, we provide a formal Bayesian framework to account for people's attitudes toward vaccines for a given disease. While this is initial work, we show that the model can capture attitudes for people who refuse and accept vaccines. This provides a positive start to model attitudes toward vaccination for diseases in general.

Attitudes toward vaccination

Anecdotally, people who hold vaccine-hesitant attitudes are often dismissed as unreasonable, and thereby, excluded from vaccination campaigns–e.g., The World Health Organization (WHO) instructs public health officials to exclude "the vocal vaccine denier" from targeted messaging (*WHO/Europe*, 2017) and research links the widespread anti-vaccination sentiment to excessive mediatization of controversies and relentless misinformation on social media platforms (Benecke & DeYoung, 2019).

While qualitative and descriptive studies into vaccine attitudes are tremendously important to give an impression of the factors that may lead people to accept or reject a vaccine for a specific disease, there is a need for formal modelling of vaccine attitudes in general. To offer this, we suggest a Bayesian Network to capture people's attitudes toward vaccines for a given disease.

To develop such a model, we explore literature on attitudes toward vaccination. Dube et al. (2013) provide a thorough literature review defining potential causes for the increase in vaccine hesitancy in recent years. Their conceptual model describes six psychosocial factors influencing vaccine hesitancy at the individual level and their interactions within the wider historic, political, and socio-cultural context in which vaccination decisions occur. These form the basis for the present study: Knowledge and information about vaccination; Past experiences with vaccination services (e.g., negative encounters with administrators; fear of needles and/or pain post-vaccination); Perceived importance of vaccination for maintaining health (e.g., usage of alternative medicine such as homeopathy is associated with non-vaccination); Risk perception and trust–established as strong predictors for adult vaccination behaviors.; Subjective social norm–identified as a potent driver of vaccine acceptance; and Religious and moral convictions. Recent work related to COVID-19 has evidenced the influence of social norms and perception of risk in adhering to prevention and control guidelines, including vaccination acceptance (Bellato, 2020; Young & Goldstein, 2021, see also Hornsey et al., 2018). Building on Dube et al. (2013), the present study integrates key factors in vaccine hesitancy into a formal Bayesian model of the causal structure for vaccine acceptance.

Formal modelling has been used to gauge people's beliefs about vaccines. Powell et al (2023) uses a Bayesian learning algorithm to construct a model of participants' intuitive theory on vaccines. They find 14 variables and 29 edges (e.g. vaccine danger, disease severity, parental protectiveness, and medical scepticism). In comparison, the current work uses a predictive Bayesian model with an empirically elicited Conditional Probability Table to compare model predictions with observations. We believe these methods complement rather than challenge each other.

Following these studies, the proposed model integrates people's subjective beliefs concerning the likelihood of contracting the disease in question, whether the available vaccine for the disease is safe, their beliefs about the dangers posed if exposed to the disease, their government's advice on vaccination, and whether they believe there is a social norm to take the vaccine.

A Bayesian approach to predict attitudes toward vaccination

In this paper, we adopt a Bayesian perspective on vaccine attitudes. Bayes' theorem gives a normative belief revision model by integrating people's subjective prior degrees of belief (between 0 and 1) with the likelihood ratio to estimate the posterior degree of belief and expresses how a rational agent should revise their belief in a hypothesis, h when faced with new evidence E, captured by Bayes' theorem.

$$p(h|e) = \frac{p(h)p(e|h)}{p(h)p(e[h) + p(\neg h)p(e|\neg h)}$$

where, $p(h|e)$, represents the posterior degree of belief in the hypothesis, h. The posterior degree of belief is a function of what the person believed about the hypothesis before hearing the new evidence called the prior belief, $p(h)$, the conditional probability of observing the evidence, e, if the hypothesis was true $(e|h)$ and the conditional probability of observing the evidence, e, if the hypothesis was false $(e|\neg h)$.

Bayesian approaches to belief revision have been applied to reasoning (Oaksford & Chater, 2007), argumentation (Hahn & Oaksford, 2006, 2007), and other areas of cognition (Chater et al., 2010).As vaccine hesitancy has sometimes been described as irrational, Bayesian approaches are particularly useful, as they have been used to explain reasoning biases or errors from a rational perspective, including arguments from ignorance (Hahn et al, 2005), ad hominem (Harris et al., 2012; Oaksford & Hahn, 2012), slippery slope (Corner et al., 2011), and circular arguments (Hahn, Oaksford, & Corner, 2005).

We believe a Bayesian approach is appropriate for three reasons. First, model predictions are generated from people's subjective degrees of belief. If two people disagree on the danger of a particular disease (e.g., where one believes it to be entirely harmless while the other believes it is deadly), they should *reasonably* reach different conclusions regarding the necessity to vaccinate against that disease. Studies have shown that Bayesian updating can lead to outcomes like polarization (Jern et al., 2014) and climate change denialism (Cook & Lewandowsky, 2016; Lewandowsky et al., 2019). This provides a perspective to potentially account for vaccine hesitancy beliefs without reference to irrationality. Second, it provides a falsifiable framework where model predictions can be compared directly with observed attitudes towards vaccines. The proposed model has no free parameters (as we set values for the conditional probability table in Study 1 and the values for the prior beliefs in Study 2). Finally, it provides gradient belief predictions between 0 and 1 (i.e., how useful people believe a vaccine is for a disease) rather than a dichotomous choice of taking the vaccine or now (0 or 1).

We formalize the components identified in the literature review (subjective beliefs for the likelihood of catching the disease, the severity of the disease's symptoms, the safety of the vaccine, social norms regarding vaccination, and advice from government concerning vaccination) within a Bayesian Network (BN) model (Pearl, 1988) to test whether the same rational framework is able to account for vaccine acceptance as well as hesitancy (see Figure 1).

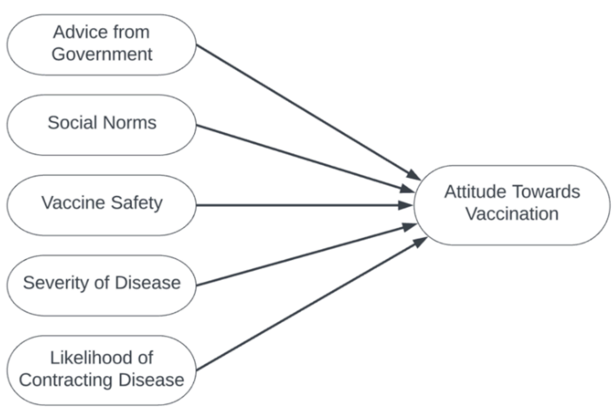

Figure 1: The intended Bayesian Network including the five key factors of the probabilistic model (danger of illness, danger of the vaccine, prevalence of the disease, perceived social norm, and governmental recommendation).

BNs are probabilistic graphical models which represent the relations between items of evidence and possible hypotheses allowing one to draw inferences about specific hypotheses based on observed evidence. The graph consists of a set of nodes representing variables of interest (i.e., hypotheses, evidence, reliability) and a set of directed links representing the probabilistic relations between variables, and in particular, the conditional dependencies. The quantitative component of BNs consists of conditional probability distributions for each variable in the graph. Bayesian networks, therefore, provide the means to test causal models of scenarios – including models of source reliability – and compare intuitive inferences of lay reasoners to a normative standard (Lagnado et al., 2013).

To test the model, Study 1 elicits conditionals probabilities needed to populate a Bayesian model for vaccine acceptance. Following this, Study 2 assesses the model's efficiency at predicting someone's probabilistic disposition to vaccination uptake based on the determined conditionals related to the causal effect. A confirmatory study will test the accuracy of the model by measuring participants' priors and comparing the model's predictions to participants' actual posterior to check if they align. In line with Dube et al.'s (2013) recommendations to understand vaccine hesitancy as a phenomenon occurring in a wider socio-cultural context, the study will compare perceptions related to COVID-19 (high-controversy and low-relevance) and influenza (low-controversy and low-relevance). Specifically, we test the following hypothesis:

H1 (Confirmatory): Predictions from the Bayesian model will correlate positively with attitudes toward vaccination.

Alongside subjective beliefs concerning the five factors in the proposed model, we suggest that perceived trust in government will be related to vaccine attitudes. In social and political science studies, trust in government has been linked with key observable metrics relevant to vaccination attitudes. Trust is positively correlated with public policy compliance (Ayres & Braithwaite, 1992), high trust in others increases societal cooperation (Fukuyama, 1995), and lack of trust may instigate civic participation through grassroots movements (Levi & Stoker, 2000). We therefore expect participants with high degrees of trust to exhibit higher degree of compliance with government advice. That is:

H2 (Confirmatory): Perceived trust in government will correlate positively with attitudes toward vaccination.

Before discussing the results of the studies, it is important to note the purpose of the current model and paper. We explore whether a Bayesian model can explain the *rationale* for people's subjective attitudes toward vaccines. That is, given their personal beliefs about the dangers of a particular disease and the safety of the vaccine, should they be willing to take the vaccine for that disease? It is possible for a person to act reasonably given their subjective beliefs about the world while simultaneously be entirely mistaken in those beliefs to begin with. For example, by all measurable accounts, it is perfectly safe to consume water. However, if a person earnestly and strongly believes that water is bad for their health (a belief that runs counter to observations), they should *reasonably yet mistakenly* avoid water. In other words, we are not interested in whether people's subjective beliefs are 'correct' (to the best of our ability to assess such questions), but whether people make reasonable inferences and exhibit rational vaccination attitudes *given* their subjective beliefs.

Study 1: Setting the Conditional Probability Table

The proposed model integrates five distinct components to describe people's attitudes toward vaccines for a given disease: their perception of the severity of the disease if contracted, the probability of contracting the disease, the safety of the vaccine, how strongly local governments and health officials encourage or discourage the vaccine, and how strongly local norms encourage or discourage vaccination. To implement the Bayesian network, we require a conditional probability table (CPT). Study 1 elicits the CPT from participants to enable model predictions with no free parameters when testing the mode in Study 2.

Design and materials

A survey design was employed to elicit conditional probabilities to populate the model. The questionnaire was designed and administered using Qualtrics, an online survey tool. As described, the key factors necessary to accept a vaccine were determined through literature review. To fully instantiate the model with the five conditionals we needed to elicit 2^5 i.e., 32 conditionals. The conditionals describe hypothetical scenarios where model parameters are set to 0 or 1 and mixed fully. For example, in considering the likelihood of catching a given disease, 0 is described in the scenarios as 'contracting the disease is certain' while 1 is described as 'contracting the disease is impossible'. Most diseases in the real world fall somewhere between impossible and certain (i.e., between 0 and 1). By eliciting the conditions for the full set of possible scenarios (32 for the proposed model), it enables model predictions for diseases by asking participants what they believe about **that** particular disease (as we do in Study 2). To illustrate CPT scenarios, we described two combinations of conditionals which evoke scenarios that best explain the intended purpose of the survey design.

The scenario where vaccination should be most desirable is the condition where contracting the disease is certain, the vaccine is described as completely safe, the disease has extremely severe symptoms, there is strong a government mandate to take the vaccine, and there is a strong social norm to take the vaccine. This represents the combination of conditionals where individuals should perceive vaccination as the most necessary. Conversely, the scenario, which could be described as the least favorable in terms of warranting vaccination is the conditional where risk of contagion is impossible, the vaccine side-effects are described as severe, the disease has little to no symptoms if caught, where governments discourage vaccination, and where there is a strong social norm to not take the vaccine. We argue this scenario represents a case where people should be unwilling to take the vaccine.

For each hypothetical scenario, participants were asked to rate how necessary they believe taking a vaccine in that given instance on a probabilistic scale from 0 (representing 'It would absolutely NOT be necessary to take a vaccine') to 1 (representing 'It would be absolutely necessary to take a vaccine').

Participants

The study included 92 respondents (44 male, 47 female. 1 preferred not to say) participants, aged 19 to 79 (median = 40) were recruited via Prolific. They had to be aged 18+ and speak fluent English. Twelve were removed for the following reasons: First, response bias: participants who responded similar extremes to all prompts (e.g., participants who answer '1' to all questions) were deemed to be disingenuous (or to have not taken time to read the scenarios). Second, logical incoherence: people who fundamentally misunderstand the direction of the conditionals. For example, if a participant reported less than .15 probability of accepting the vaccine for the most favorable scenario or/and reported more than .85 for the least favorable scenario (the two scenarios described in the designs and materials section), they were deemed to have misunderstood the description of the scenarios.

Results

To set each conditional probability, we take the mean of the participants. For example, in the above scenario (Fig. 2) where participants are told to consider a hypothetical scenario where contracting the disease is certain, the vaccine is safe,

Table 1: Empirically elicited conditional probability table. '0' in the first place refers to 'Contracting the disease is certain' while '1' refers to 'Contracting the disease is impossible'; '0' in the second place refers to 'Vaccine has severe side effects' while '1' refers to 'Vaccine is completely safe'; '0' in the third place refers to 'The disease has extremely severe symptoms' while '1' refers to 'The disease has little to no symptoms'; '0' in the fourth place refers to 'Government STRONGLY discourages vaccination' while '1' refers to 'Government STRONGLY encourages vaccination'; '0' in the fifth place refers to 'There is a STRONG social norm to NOT take the vaccine' while '1' refers to 'There is a STRONG social norm to take the vaccine'.

01011	**10100**	**00000**	**00101**	**00001**	**00010**	**00011**	**00100**
0.88	0.06	0.39	0.23	0.51	0.53	0.66	0.15
00110	**00111**	**01000**	**01001**	**01010**	**01100**	**01101**	**01110**
0.28	0.36	0.74	0.76	0.78	0.38	0.51	0.51
01111	**10000**	**10001**	**10010**	**10011**	**10101**	**10110**	**10111**
0.64	0.13	0.21	0.21	0.29	0.15	0.15	0.21
11000	**11001**	**11010**	**11011**	**11100**	**11101**	**11110**	**11111**
0.27	0.35	0.42	0.44	0.14	0.27	0.29	0.34

the disease has extremely severe symptoms, the government strongly encourages vaccination, and there is a strong social norm to take the vaccine, the average response was 0.88 on a scale from 0 to 1 (case 01011 in Table 1), indicating that people believe that it would be necessary to take the vaccine in this case. By setting the CPT fully, we avoid potential free parameters that can be fitted posteriorly to observations (which is a challenge to Bayesian models raised by Jones & Love, 2011). When asking people for their subjective perception of the priors for each of the five elements, there are no free parameters in the model for subsequent fitting. This means that it is not possible to adjust model predictions a posteriori.

The CPT (Table 1) described the relationship between the five factors and the supposed reason for taking (or not taking) a vaccine for any disease. By asking participants what they believe about the priors for these factors for specific illnesses, the model should be able to project their subjective attitude toward the necessity of taking a vaccine for that disease.

Experiment 2: Testing the model

The paper outlines and tests a computational model for general attitudes toward willingness to be vaccinated for a given disease. As we are proposing a model that should be able to describe (and potentially predict) how people relate to vaccines for diseases in this general manner, we test model predictions against two different illnesses. The two diseases are COVID-19 (chosen due to the importance of this disease in recent years) and the flu (chosen due to the yearly prevalence of this disease).

Design and materials

The priors will be collected by asking participants about the five factors of interest for either COVID-19 or Influenza. The study is a between—subject design so that participants only provide their beliefs about the flu or about COVID. Notable, we ask participants to consider the factors for an imaginary person rather than themselves. Specifically, we ask them to consider 'Imagine a 40-year old American citizen who has not had the COVID-19 vaccine and who has not previously been infected'. This was chosen to normalize the scenario as much as possible between participants (whom were based in the US). However, we also ask if they have personally taken a vaccine for COVID or the flu.

Each factor will be introduced with a neutral description outlining what it signifies, and how it relates to the vaccines. For example, in considering the danger of a disease, the factor was introduced with the following: *'Viruses can be more or less dangerous - some have little to no impact on people (e.g., a mild runny nose) while others can be extremely dangerous (e.g., causing death or extremely severe long-term symptoms like paralysis).'* Following this description,

Participants were then asked to provide their beliefs about the given factor on a 0-100 scale. For example, for the likelihood of contracting the disease, participants will be asked "*How likely do you think this person is to catch Influenza in the next 12 months?"*.

Following this stage, we collect participants' posterior beliefs. These will be collected by asking participants about the vaccination beliefs relating to a given disease (i.e., "*How necessary do you think it is that this person takes the Influenza vaccine*?"). The experiment will include questions on participants' beliefs on governments' trustworthiness and expertise and perceived dependencies between the sources, which will be later used for exploratory study on the source-credibility model (Figure 2 illustrates the survey flow).

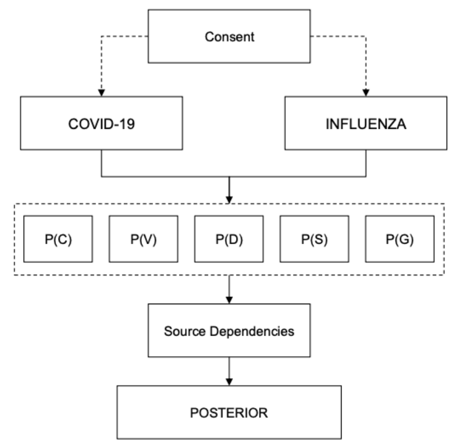

Figure 2: Flow of experiment 2

Participants

We recruited 600 participants from Prolific. Participants had to be aged 18+ with English as their first language. We recruited participants from the US, as vaccine hesitancy at time of collection was higher than in the UK and Australia, increasing the chance of locating participants who had not accepted the COVID-19 vaccine. To ensure a fair wage, we estimated completion time to be around 5 min and paid £0.75 for completion corresponding to £9/hour (actual completion time 4.83 min). We remove participants who completed the survey in less than 2 min, leaving 514 participants. The average age of participants is 38.3 (13.3) and gender is split with 255 female, 253 male, and 6 non-binary participants.

Results

We use participants' responses on whether they accepted a vaccine for COVID or the flu (depending on their condition). For COVID, 210 participants had accepted the vaccine while 47 had not. Using a independent-sample t-test to compare estimates of the five priors for accept and reject participants, we find significant differences for their estimates of contagion ($M_{accept} = 0.64$, $M_{reject} = 0.39$, $t = 7.004$, $p < 0.001$), vaccine safety ($M_{accept} = 0.87$, $M_{reject} = 0.49$, $t = 8.493$, $p <$

0.001), danger (M_{accept} = 0.54, M_{reject} = 0.30, t = 6.191, p < 0.001). We find no difference in perception of government advice, although interestingly participants who reject agree more with the sentiment that the government has encouraged vaccination (M_{accept} = 0.87, M_{reject} = 0.92, t = -1.965, p = 0.051). Similarly, we find no difference in perceived social norm (M_{accept} = 0.72, M_{reject} = 0.69, t = 0.986, p < 0.328).

For the flu, 182 have accepted a vaccine while 75 had not. Similarly, we find significant differences for estimates of contagion, (M_{accept} = 0.51, M_{reject} = 0.37, t = 4.587, p < 0.001), vaccine safety (M_{accept} = 0.87, M_{reject} = 0.71, t = 4.749, p < 0.001), and danger (M_{accept} = 0.44, M_{reject} = 0.34, t = 2.746, p < 0.001). For the flu, we also find significant differences for government advice (M_{accept} = 0.85, M_{reject} = 075, t = 3,764, p < 0.001) and social norm (M_{accept} = 0.71, M_{reject} = 0.60, t = 4.211, p < 0.001). Descriptively, it appears the five factors broadly capture differences in what people who accept and reject vaccines believe about the diseases in question.

To test H1, we use the priors elicited from each participant and the CPT from Study 2 to generate model predictions for each participant, which is compared with that person's attitude toward vaccination. Model predictions enjoy a good fit with attitudes to the COVID-19 vaccine (R^2 = 0.53, p < 0.001) and the flu (R^2 = 0.44, p < 0.001). That is, the model can capture 44 and 53% of the variance for flu and COVID respectively, providing strong confirmatory evidence for H1 (Figure 3 show predictions and observations for the flu).

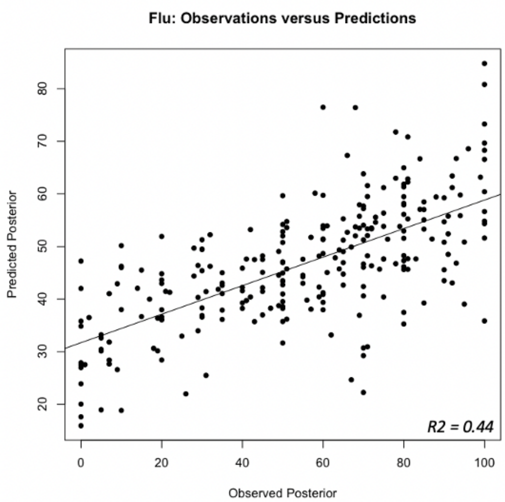

Figure 3: Predictions versus observations (Flu)

To test H2, we correlate perceived government trust with attitude toward vaccination. In line with predictions, we see positive correlations for COVID-19 (R^2 = 0.46, p < 0.001) and the flu (R^2 = 0.20, p < 0.001). This suggests people's attitudes toward vaccination is correlated with their attitudes toward trust in government. However, when running a multiple linear regression with model predictions and trust perceptions, we see only little improvement compared with model predictions by themselves (for COVID, R^2 goes from 0.53 for the model fit to 0.55 when including trust and from 0.44 to 0.46 for the flu). This suggests perceived trust does not add much to model fit compared with the five-factor Bayesian in isolation.

Concluding remarks

The paper tests a Bayesian model for predicting attitudes toward vaccines for a given disease. The model integrates five factors (subjective beliefs concerning danger of illness, safety of the vaccine, prevalence of the disease, perceived social norm, and governmental recommendation). The model can capture a significant amount of the variance for people who accept **and** for people who reject the vaccines for COVID-19 (53%) and the flu (44%) respectively. This suggests that people who refuse the vaccine may do so from a reasonable perspective given their subjective beliefs. That is, if a person earnestly believes a disease to be mild with low probability of contagion, and is doubtful of the safety of the vaccine, they may reasonably forego it. As we stress in the literature review, this does not mean we support this view (as we believe the COVID-19 vaccine to be safe and effective against a harmful illness), but it suggests that people who refuse may be rational in doing so from their perspective.

This is important and good news, as it suggests that people would agree to take the vaccine if these beliefs were changed (how to run an effective community outreach program is beyond the scope of this paper). In line with the idea that information campaigns are critical, conspiracy theories and misinformation has been shown to be in line with COVID-19 vaccine hesitancy (Enders et al., 2022).

We further find support for a positive correlation between perceived trust in government and vaccine attitudes. However, when including this in a model with the Bayesian predictions, it fails to increase the explanatory potential significantly.

While the paper provides initial evidence for developing a general model for vaccine attitudes, we note that it is early work with multiple open questions. To name some, the model should be tested against other diseases to gauge whether it is a general model. Further, the data in the paper is from the USA. Cross-cultural validation is also required to explore whether cultures may differ in the conditional probability table (i.e., more fundamentally, how factors are treated) and whether the model enjoys a similar fit across cultures. For example, it is plausible that the social norm and governmental advice factors may function differently across cultures. Further, the model should be tested against psychometric differences within a population (e.g., exploring whether the government and social norm factors differ depending on their score on the Right-Wing Authoritarianism scale).

In sum, the paper provides positive evidence for the development and testing of a general Bayesian model with the aim of predicting attitudes toward vaccination. This is an important step toward understanding why people choose to accept or reject the vaccine. The model has application potential, as it provides a causal model to describe the factors that may be involved with changing overall attitudes. We hope the work will continue in the future by applying the model to other diseases, cross-culturally, and to explore individual differences in more detail.

References

Abad, C. L., & Safdar, N. (2015). The Reemergence of Measles. Current infectious disease reports, 17(12), 51

Ayres I & Braithwaite J. (1992) *Responsive Regulation*, Oxford University Press

Bellato, A. (2020). Psychological factors underlying adherence to COVID-19 regulations: A commentary on how to promote compliance through mass media and limit the risk of a second wave. *Social Sciences & Humanities Open*, 2(1), 100062

Benecke, O., & DeYoung, S. E. (2019). Anti-Vaccine Decision-Making and Measles Resurgence in the United States. Global pediatric health. https://doi.org/10.1177/2333794X19862949

Bor, A., Jørgensen, F. & Petersen, M. B. (2022) Prejudice Against the Vaccinated and the Unvaccinated During the COVID-19 Pandemic: A Global Conjoint Experiment*, Working paper*

Chater, N., Oaksford, M., Hahn, U. & Heit, E. (2010) Bayesian models of cognition*,* Wiley Interdisciplinary Reviews: Cognitive Science 1 (6), 811-823

Cook, J., & Lewandowsky, S. (2016). Rational irrationality: Modeling cli- mate change belief polarization using Bayesian networks. Topics in Cognitive Science, 8(1), 160–179.

Corner, A., Hahn, U. & Oaksford, M. (2011) The psychological mechanisms of the slippery slope argument, *Journal of Memory and Language 64,* 133-152.

Dubé, E., Laberge, C., Guay, M., Bramadat, P., Roy, R., & Bettinger, J. A. (2013). Vaccine hesitancy: An overview. Human Vaccines & Immunotherapeutics, 9(8), 1763-1773.

Enders, A., Uscinski, J., Klofstad, C. & Stoler, J. (2022) On the relationship between conspiracy theory beliefs, misinformation, and vaccine hesitancy, PLoS One 17(10): e0276082

Fukuyama F. (1995) *Trust*, Basic Books

Hahn, U., & Oaksford, M. (2006) A normative theory of argument strength, *Informal Logic* 26*,* 1-24

Hahn, U. & Oaksford, M. (2007b) The burden of proof and its role in argumentation, *Argumentation 21*, 39-61

Hahn, U., Oaksford, M. & Bayindir, H. (2005) How convinced should we be by negative evidence? In Bara, B., Barsalou, L. & Bucciarelli, M. (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society,* Mahwah, NJ: Erlbaum, 887-892

Hahn, U., Oaksford, M., & Corner, A. (2005) Circular arguments, begging the question and the formalization of argument strength, in A. Russell, T. Honkela, K. Lagus & M. Polla (Eds.), *Proceedings of AMKLC '05 International Symposium on Adaptive Models of Knowledge, Language and Cognition*, Helsinki, 34-40

Harris, A. J. L., Hahn, U., Madsen, J. K., & Hsu, A. S. (2015). The Appeal to Expert Opinion: Quantitative support for a Bayesian Network Approach. *Cognitive Science* 40*,* 1496-1533

Harris, A., Hsu, A. & Madsen, J. K. (2012) Because Hitler did it! Quantitative tests of Bayesian argumentation using *Ad Hominem*, *Thinking & Reasoning* 18 (3), 311-343

Hornsey, M. J., Harris, E. A., & Fielding, K. S. (2018). The psychological roots of anti-vaccination attitudes: A 24-nation investigation, *Health Psychology* 37 (4), 307–315.

IBMS. (2019, April 26). Vaccine-preventable diseases on the rise. Institute of Biomedical Science. Retrieved March 17, 2022

Jern, A., Chang, K. K., & Kemp, C. (2014). Belief polarization is not always irrational. Psychological Review, 121(2), 206–224.

Lagnado, D., Fenton, N. & Neil, M. (2013) Legal Idioms: A framework for evidential reasoning, *Argumentation & Computation* 1, 1-18

Levi M & Stoker L. (2000) Political Trust and Trustworthiness, *Annual Review of Political Science* 3, 475-507

Lewandowsky, S., Pilditch, T. D., Madsen, J. K., Oreskes, N. & Risbey, J. S. (2019) Seepage and influence: An evidence-resistant minority can affect scientific belief formation and public opinion, *Cognition* 188, 124-139

Oaksford, M. & Chater, N. (2007) *Bayesian Rationality: the probabilistic approach to human reasoning*, Oxford University Press, Oxford: UK

Oaksford, M., & Hahn, U. (2012). Why are we convinced by the ad hominem argument? Source reliability or pragma-dialectictics. *Bayesian Argumentation*, 39-58.

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 Heath – Europe* 1, 100012

Pearl, J. (1988) *Probabilistic reasoning in intelligent systems*, Morgan Kaufmann

Powell, D., Weisman, K. & Markman, E. M. (2023) Modeling and leveraging intuitive theories to improve vaccine attitudes, *Journal of Experimental Psychology: General*

Salmon, D. A., Dudlet, M. Z., Brewer, J., Kan, L., Gerber, J. E., Budigan, H., Proveaux, T. M., Bernier, R., Rimal, R. & schwartz, B. (2021) COVID-19 vaccination attitudes, values and intentions among United States adults prior to emergency use authorization, *Vaccine* 39, 2698-2711

Serpell, L., & Green, J. (2006). Parental decision-making in childhood vaccination. *Vaccine*, 24(19), 4041-4046

Wamsley, L. (2019, March 12). Measles Is Spiking Around The Globe. How Worried Should We Be? NPR.org. Measles Is Spiking Around The Globe. How Worried Should We Be?

WHO/Europe. (2017) How to respond to vocal vaccine deniers in public. Retrieved March 14, 2022

Young, S. D., & Goldstein, N. J. (2021). Applying social norms interventions to increase adherence to COVID-19 prevention and control guidelines. *Preventative Medicine*, 145, 106424