Religiosity and Suicide Risk by Sexual Orientation

(BIOS 619)

Nam-Anh Tran

Abstract

Suicide is a significant public health concern, with various risk factors influenced by age, gender, sexual orientation, and socioeconomic context. The continuous rise in suicide rates from 2020 to 2022 suggests that suicide prevention efforts have been ineffective. Research indicates that religion may positively affect suicidal ideation, which may be an essential factor in suicide prevention projects. However, there is a lack of persuasive evidence to support this claim fully. This study aims to evaluate the association between suicidal ideation and religious beliefs and compare it among groups of sexual orientation and age. The result shows that people who are homosexual and have religious beliefs are less likely to think about suicide. This association is weaker in the group of bisexual people and determined to be null in the group of heterosexual people.

1 Introduction

The United States in 2022 witnessed a continual increase in suicide deaths with a comparable rate in 2021.[1] Numerous solutions are proposed with the primary goal of reducing this concerning figure over an extended period. However, these suicide prevention campaigns seem not effective as the suicide rate does not reduce. Investigating and comprehending factors associated with suicide is the cornerstone of these efforts. Religious belief is considered, and its connection with the suicide rate is ostensibly highlighted in studies, although persuasive evidence is lacking.[2–4] Another challenge of studies about the association between religion and suicide is that religious beliefs are not constant among communities. This leads to unpredictable suicide risk in varied subpopulation.

LGBTQ+ individuals, perhaps, have different religious experiences, possibly causing additional mental challenges, which sometimes increase suicidal ideation or attempts. Further, this connection becomes more unpredictable as the religious beliefs can change during lifetime. Some religions provides a supportive community and a sense of spiritual grounding, while other religions holding traditional views often lead to internal conflicts and community rejection. This incorporates challenges to study and comprehend the effect of religiosity on the suicide risk in this community.[5]

Thus, we aim to assess the association between religion and suicidal ideation among groups of sexual identities; particularly, we compare the suicidal ideation of people who have a strong belief in religion to those who do not among three communities: heterosexual, homosexual and bisexual people. This result will reflect disparity in the effect of religion on suicide risk in these communities, which facilitates specialized methods with the tailored effect for each group being proposed in the future. This result illuminates our understanding about the influence of religious belief on suicide rate in LGB community, beneficial to suicide prevention activities.

The association between religiosity and suicidal ideation was considered in the literature. A meta-analysis study published in 2021 showed an inverse association between religion and suicidal ideation.[2] However, this effect is complex and varies among different patients populations.[2–4] Hence, more studies that are undertaken with varied cohorts are necessary to elucidate the relationship between religiosity and suicidal ideation, facilitating a more comprehensive understanding of this relationship.

We address the investigation based on statistical models fitted using the Bayesian paradigm. This approach allows incorporating external information to the observed data, strengthening evidence by informative sources, beside the data, through the prior distribution, reflecting our knowledge about the hypothesis before observing the data. For instance, the data we utilise to fit the model have missing value outcomes, where the missingness seems to be influenced by multiple factors, including outcome itself. The Bayesian framework enables to combine evidence relating to the missingness from other studies to improve the missing mechanism model jointly fitted with the analysis model under the non-ignorability assumption.[6]

The outcome of interest is the proportion of suicidal ideation of people who have religious belief compared to those who do not. The difference is estimated over segments of interest, involving heterosexual, homosexual, and bisexual versus four age groups, 18-25, 26-34, 35-49, and 50+. The models are specified based on the data structure which is discussed in the next section.

2 Data exploration

The cross-sectional data, relating to The National Surveys on Drug Use and Health (NSDUH), are extracted from the site of Substance Abuse and Mental Health Services Administration. NSDUH conducted both face-to-face household interviews and web-based interviews. Only data in 2022 are extracted for this analysis. The samples includes all people living in the United States who are 12 or more years old; not living in institutions, such as prisons or nursing homes; not military personnel on active duty; not experiencing homelessness with no fixed address; able to answer the questionnaire in either English or Spanish.

The survey acquired the answers to the question “at any time in the past 12 months, up to and including today, did you seriously think about trying to kill yourself?”. The exposure was also obtained based on the question “Your religious beliefs influence how you make decisions in your life?”. Other covariates that are also considered to include in the model are shown in Table 1. The refined data have 2129 subjects with five missing values. The subjects are characterised by 13 covariates, divided into two groups. The first group includes three essential covariates: religious beliefs, sexual identity and age group, of which coefficients are the estimands of interest; the second group includes 10 confounding variables, which are selected, based on the variable-selection models, to adjust in the model.

The upper plot in Figure 1 indicates that the proportion of suicidal ideation is lower if people, regardless of sexual identity and age, have religious beliefs. However, the difference is not constant across groups of age and sexual identity, as shown in the lower plot. This suggests the association between religion and suicidal ideation should be examined separately in each group of age and sexual identity. Thus, the model should include the interaction terms of the exposure versus age, and the exposure versus sexual identity.

3 Statistical analysis

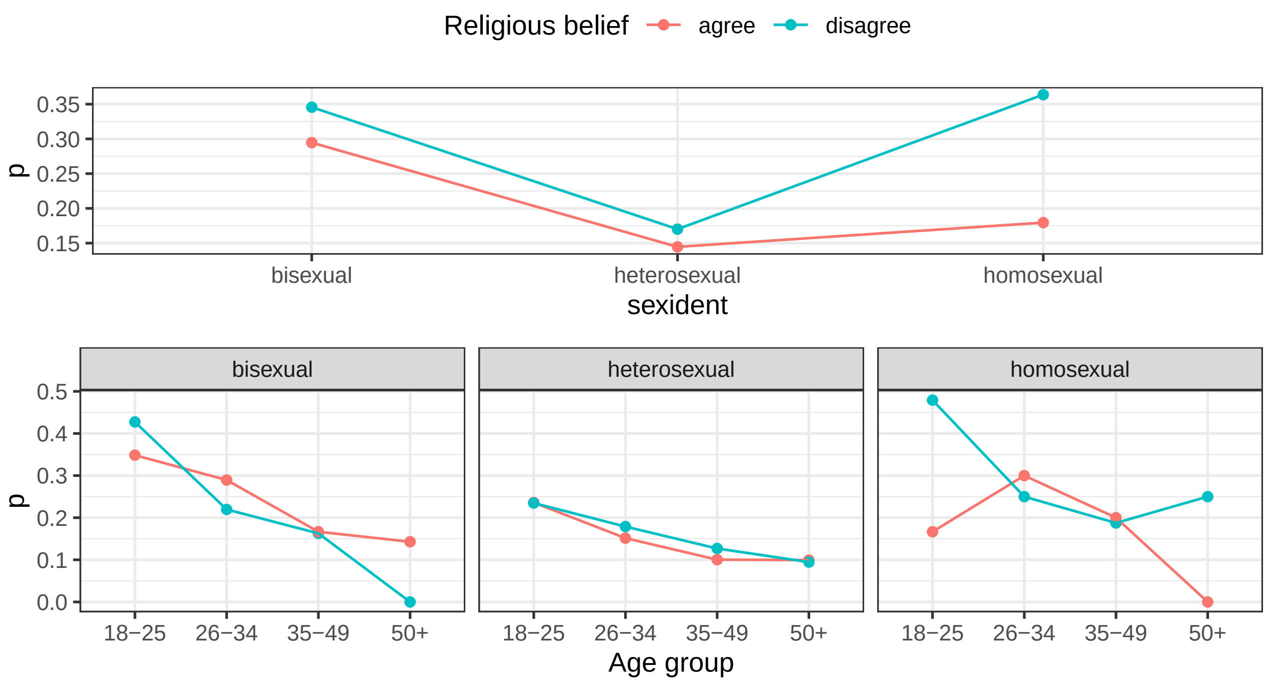
We first discuss models for variable selections. As many confounding variables are available, we aim to simplify the model without losing the model’s objective in this study, which is drawing the inference about suicidal ideation proportion in different groups characterized by sexual identity and

Table 1: The primary outcome, binary variable indicating suicidal plans, and covariates of the data.

Variables	Information
Suicidal thought (outcome)	*Survey question:* At any time in the past 12 months, up to and including today, did you seriously think about trying to kill yourself? (yes or no [ref.])
Religious beliefs (exposure)	*Survey question:* Your religious beliefs influence how you make decisions in your life? (agree or disagree [ref.])
Sexual orientation	*Survey question:* Which one of the following do you consider yourself to be? (lesbian/gay, biosexual, heterosexual [ref.])
Age	*Four categories:* 18-25 (ref.), 26-34, 35-49, 50+
Gender	*Two categories:* male [ref.], female
Race	*Three categories:* black, asian, white [ref.]
Total income	*Four categories:* less than 20K [ref.], 30K-49K, 50K-74.9K, 57K+
Employment status	*Three categories:* unemployed (ref.), part-time, full time.
Marital status	*Three categories:* widowed/divorced/separated [ref.], single, married
Overal health*	*Four categories:* 1 (excellent), 2 (very good), 3 (good), 4 (fair/poor)
Emotional distress*	*Four categories:* 1 (mild), 2 (moderate), 3 (severe), 4 (very severe)
Sleep disorder	*Two categories:* No [ref.], Yes
Education	*Four categories:* No [ref.], Yes
Drugs (including heroin, marijuana, cocaine)	*Two categories:* No [ref.], Yes

* treated as continuous variables to reflect a continuous range between 1 and 4.

Figure 1: The porportion of suicidal ideation in two groups characterised by the religion across the age group.



age. Second, we discuss analysis models fitted in this study. Finally, we discuss the approach for handling missing data based on the Bayesian framework.

3.1 Variable selection

Two variable-selection models, Spike-and-Slab and LASSO, are fitted to select predictors that highly contribute to the outcome. The Spike-and-Slab model is formulated as follows:

$$\begin{aligned} y_i &\sim \text{Bin}(1, p_i) \\ \text{logit}(p_i) &= \beta_0 + \boldsymbol{x}_{1i}^\top \boldsymbol{\beta}_1 + \boldsymbol{x}_{2i}^\top \boldsymbol{\beta}_2, \quad i = 1, 2, \ldots, 2129. \end{aligned} \tag{1}$$

where y_i is the binary outcome indicating suicidal ideation, p_i is the corresponding proportion; $\boldsymbol{x}_{1i}$ is a vector of primary variables, including religious beliefs, age, sexual identity, and interaction terms between religious beliefs versus age, and religious beliefs versus sexual identity; $\boldsymbol{x}_{2i}$ is a vector of 10 adjusted covariates. β_0, $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are the intercept, and coefficients corresponding to $\boldsymbol{x}_{1i}$ and $\boldsymbol{x}_{2i}$, respectively. The prior distributions are defined as follows.

$$
\begin{aligned}
&\beta_0 \sim N(0,100); \quad \beta_{1j} \sim N(0,100); \quad \beta_{2j} = \gamma_j \times \delta_j \\
&\gamma_j \sim \text{Bin}(1,0.5); \quad \delta_j \sim N(0,1/\tau) \\
&\tau \sim \text{Ga}(0.1,0.1)
\end{aligned} \tag{2}
$$

where β_{1j} and β_{2j} are elements of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$.

The structure of LASSO model is analogous to Spike-and-Slab except the prior distribution of $\boldsymbol{\beta}_2$ defined in Equation 2. More precisely, the block of the priors in Equation 2 is redefined as follows.

$$
\beta_{2j} \sim \text{dexp}(0,\tau), \quad \tau \sim \text{Ga}(0.1,0.1). \tag{3}
$$

Only the top-ranked covariates in the variable-selection models are adjusted in the reduced model, which will be discussed latter.

3.2 Analysis models

We fit four logistic models to evaluate the proportion of suicidal ideation. First, the full model has the same structure as Equation 1. However, the prior distributions are defined as follows:

$$
\beta_0, \beta_{1j} \text{ and } \beta_{2j} \sim N(0,100). \tag{4}
$$

Second, the crude model is fitted without the term $\boldsymbol{x}_{2i}^{\top}\boldsymbol{\beta}_2$ in Equation 1. Third, the reduced model is also similar to the the full model except $\boldsymbol{x}_{2i}$ only includes predictors suggested from the variable-selection models. Note also that the reduced model is fitted with two different prior distributions of the missingness mechanism, which is discussed in the next subsection, while the analysis model and its corresponding priors do not change, i.e. the vague priors in Equation 4.

3.3 Handling missing data

As the outcome has missing values, we fit the missingness mechanism model jointly with each analysis model, discussed in the previous subsection, under the assumption of non-ignorability to account for the uncertainty of missing values. The missingness mechanism model is formulated as follows:

Table 2: Informative prior distribution of the coefficients in the missing mechanism model.

Variable	odds ratio	95% CI lower	95% CI upper	Mean of log odds ratio	Variance of log odds ratio
intercept	0.02	0.01	0.03	-3.89	56.48
suicidal ideation	0.48	0.37	0.62	-0.73	12.47
health status	0.63	0.51	0.78	-0.46	8.45
emotional distress	0.63	0.42	0.94	-0.46	30.37

* For intercept, it is odds, and odds ratio for other coefficients.

$$R_i \sim \text{Bin}(1, p_i^{missing})$$

$$\text{logit}(p_i^{missing}) = \boldsymbol{z}_i^\top \boldsymbol{\alpha}, \quad i = 1, 2, \ldots, 2129$$

where R_i is a binary variable indicating the missing value, i.e. $R_i = 1$ if the outcome is missing and $R_i = 0$, otherwise; $p_i^{missing}$ is the probability of missing data; $\boldsymbol{z}_i$ is a vector with the length of four representing intercept, suicidal ideation, overall health, and emotional distress. These predictors are selected based on a study examining characteristics associated with non-disclosure of suicidal ideation in adults.[7] Three assumptions are made on how the missing data are associated with those predictors:

1. As people think more about suicide, they do not want to disclose their thought;
2. As people have worse health, they do not want to disclose their thought;
3. As people encounter emotional distress, they do not want to disclose their thought.

However, these assumptions lack validity and were disproved in Merelle et. al.[7] as the odds ratios associated with non-disclosing suicidal ideation, corresponding to those predictors were less than the unity. We, therefore, fit this model with two different prior distributions. First, the model is fitted using the vague priors $\boldsymbol{\alpha} \sim N_4(\mathbf{0}, 10^2\boldsymbol{I}_4)$, implying that the information used to estimate the missing values are from the data. Second, we fit the model using the informative prior distribution specified based on the results in Merelle et. al.[7] Specifically, the mean of the intercept are calculated based on the odds relating to the overall proportion of non-disclosing suicidal ideation; the mean of coefficients are calculated based on the odds ratios relating to health status, suicidal ideation, and psychological distress. Also, the variances are calculated using the corresponding 95% confidence intervals. The informative priors are shown in Table 2.

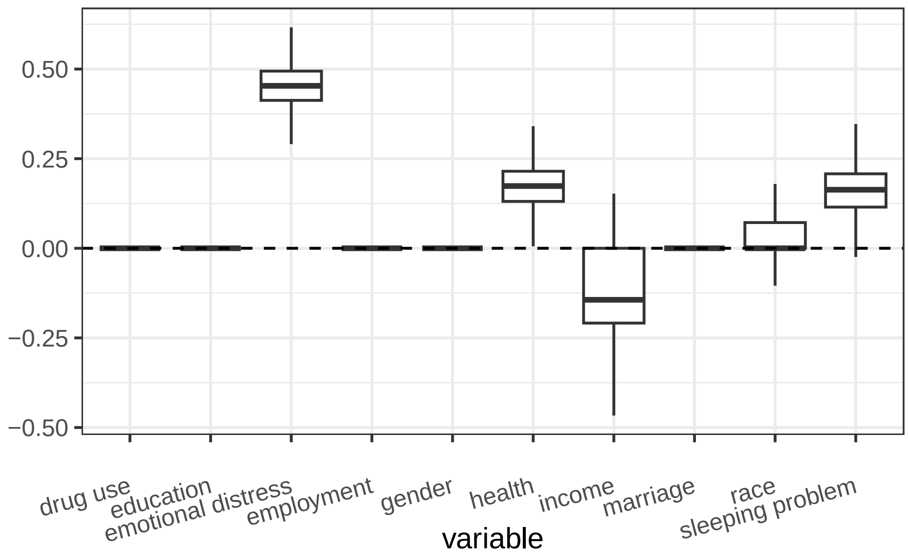
The model convergence is also examined using the the trace plots and Gelman and Rubin's convergence diagnostic.

4 Results

4.1 Variable selection

The posterior distribution of coefficients obtained from the spike-and-slab model is shown in Figure 2. Four covariates of which coeffficients mean differ from zero are emotional distress, health status, income and sleeping problem. This is reasonable as financial crisis and health problem are two the most concerning issues and can contribute to emotional distress, sleep disorder and despairing ideas, including suicidal ideation.

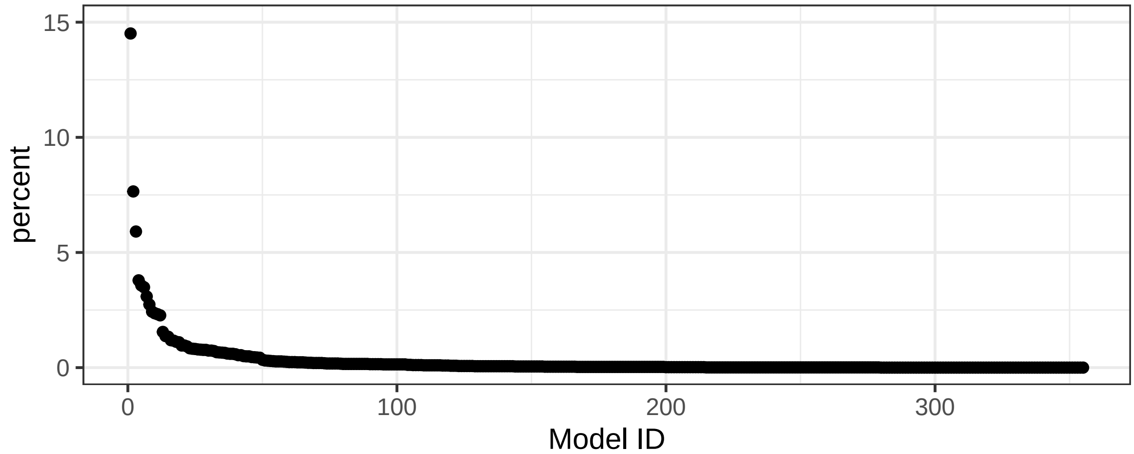
Figure 2: boxplots shows posterior distribution of model coeffcients obtained from the Spike-and-Slab model.

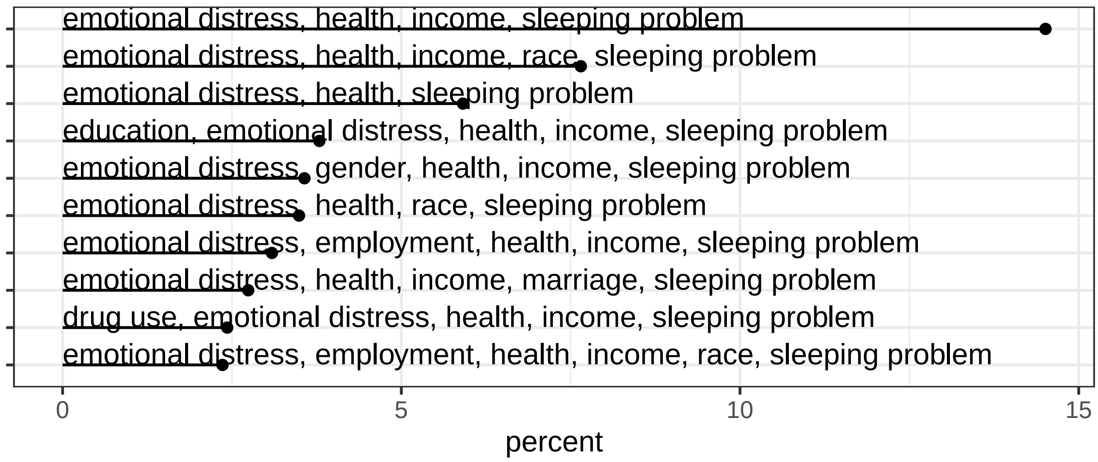


Also, the top ten models ranked by the Spike-and-Slab model are shown in Figure 3. The model with highest probability includes emotional distress, health status, income and sleep disorder. We also compare the Spike-and-Slab model to the LASSO model. Table 3 shows that four covariates with coefficients greater than 0.1 in the absolute scale in the LASSO are in line with those in the Spike-and-Slab model. This implies that the health status, income, emotional distress and sleeping problem play an essential role in estimating the outcome, i.e. proportion of suicidal ideation. Thus, these four variables are included in the reduced analysis model, i.e.

$$x_{2i} = (\text{emotional distress, health status, income and sleep disorder}).$$

Figure 3: The percentage of all model candiates ranked by the Spike-and-Slab model (upper), and the name of covariates included in the top ten models (lower).





4.2 Analysis models

The posterior mean and the corresponding 95% highest posterior density intervals[8] (HDI) (In R, the HDI is computed with ease using the function `hdi` from the package `HDInterval`) of model intercept and coefficients of four analysis models are shown in the appendix (Table 5; Table 6; Table 7; Table 8). The WAICs of the full-, crude-, and two reduced models are equal to 1974.912, 2061.658, 1974.06, and 1968.959, respectively. This indicates that the reduced model with missing mechanism jointly fitted using the informative priors (the last one) outperforms the other three models as its WAIC is the

Table 3: Compare coefficients obtained from Lasso and Spike-and-Slab.

coefficient	Lasso	Spike-and-Slab	$\|\beta_{lasso}\| > 0.1$	$\|\beta_{Spike-and-Slab}\| > 0.1$
gender	0.0448	0.0119	0	0
race	0.0774	0.0359	0	0
income	-0.1609	-0.1329	1	1
employment	0.0382	0.0093	0	0
marriage	0.0043	0.0013	0	0
health	0.1612	0.1686	1	1
emotional distress	0.4473	0.4538	1	1
sleeping problem	0.1507	0.1565	1	1
education	-0.0407	-0.0112	0	0
drug use	-0.0027	-0.0015	0	0

Table 4: Estimands of interest calculated based on the estimates of coefficients of the reduced model where the missing mechanism model is fitted using the informative priors.

Age	Bisexual	Heterosexual	Homosexual
18-25	0.77 (0.43, 1.18)	0.91 (0.56, 1.29)	0.47 (0.09, 0.96)
26-34	0.9 (0.39, 1.47)	1.05 (0.57, 1.58)	0.54 (0.11, 1.15)
35-49	0.75 (0.29, 1.3)	0.87 (0.46, 1.34)	0.45 (0.08, 0.96)
50+	0.97 (0.19, 2.07)	1.12 (0.27, 2.26)	0.58 (0.05, 1.42)

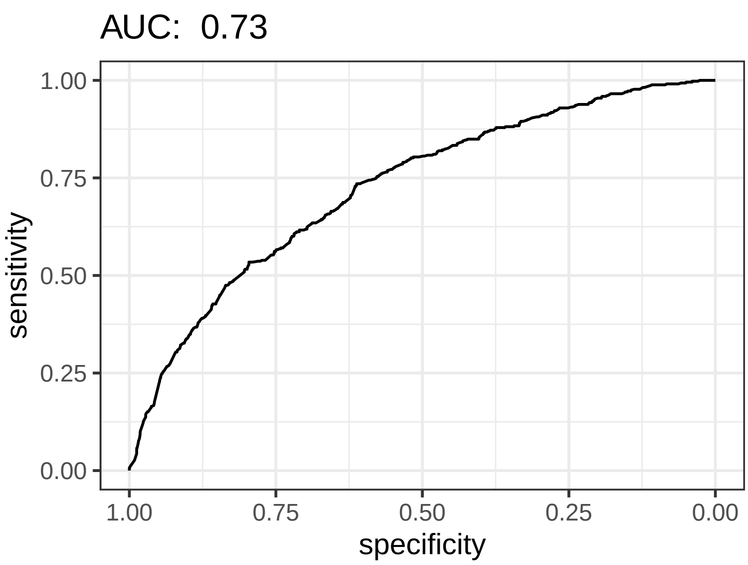
lowest. Thus, this model is used to evaluate the effect of religious beliefs on suicidal ideation.

Table 4 shows the odds ratios of interest and the corresponding 95% HDIs obtained from the reduced model with the informative priors used in the missing mechanism model. After adjusting for health status, emotional distress, income and sleep disorder, the odds ratios associated with suicidal ideation in the groups of interest indicate that people who are homosexual and have religious beliefs are less likely to have suicidal ideation, which is well-supported for those people in the age group 18-25 and 35-49. People who are bisexual, have religious beliefs, and in the age group 18-25 and 35-49 are also less likely to have suicidal ideation, but the effect level is less than that of people who have homosexual identity. In contrast, the effect of religious beliefs on suicidal ideation is not evident for people who have bisexual identity in the age group 26-34 and 50+, and those people who are heterosexual.

The trace plots of primary predictors, religion, sexual identity and age groups are shown in the appendix. For each plot, two chains overlap without the influence of preceding iterations on the current iteration, which indicates the convergence. The convergence is also confirmed based on the Gelman and Rubin's convergence diagnostic with the upper bounds less than 1.1.

Furthermore, Figure 4 shows the receiver operating characteristic (ROC) curve of the reduced model with the informative priors used in the missing mechanism model. The area under the curve of 0.73 indicates that the model has acceptable discrimination.[9]

Figure 4: ROC curve obtained from the reduced model where the missing mechanism model is fitted using the informative priors.



5 Discussion

The study aims to evaluate the association between religious beliefs and suicidal ideation. We found evidence that people self-identifying as homosexual with religious beliefs are less likely to think about suicide. This influence is also found, though less transparent, for those people who self-identify as bisexual, and the effect is least evident for people who are over 50 years of age. Minimal to no evidence of the religious belief effect on suicidal ideation is found for heterosexual people. Although no causality can be inferred, the result suggests that suicide prevention activities should consider religious beliefs as an essential factor for people who are homosexual.

This study has a limitation. The data used for analysis are only in 2022, as the data in 2020 and 2021, when the COVID-19 pandemic occurred, are unavailable. Thus, the result in this study does not account for the effect of the pandemic, which may have a significant impact on suicidal ideation. This should be improved in the future study.

Table 5: The posterior mean and corresponding 95% HDIs of coeffcients obtained from the full model

Coefficient	Mean	95% HDI lower	95% HDI upper
Intercept	0.04	0.01	0.07
religion	0.87	0.52	1.25
Bisexual	1.83	1.25	2.44
Homosexual	2.14	1.11	3.24
Age 26-34	0.64	0.41	0.89
Age 35-49	0.54	0.3	0.8
Age 50+	0.5	0.14	0.93
Religion x Bisexual	0.89	0.44	1.42
Religion x Homosexual	0.56	0.1	1.17
Religion x Age 26-34	1.22	0.57	1.97
Religion x Age 35-49	0.98	0.42	1.63
Religion x Age 50+	1.28	0.29	2.73
WAIC = 1974.912			

Appendix

Figure 5: Trace plots corresponding to six primary covariates included in the full model

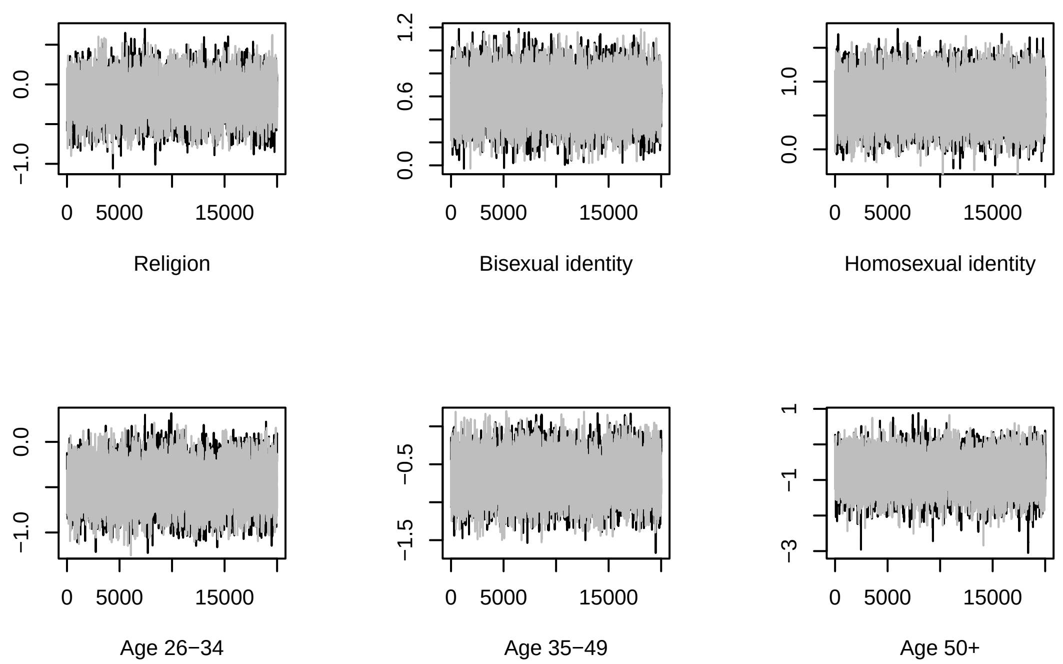


Table 6: The posterior mean and corresponding 95% HDIs of coeffcients obtained from the crude model

[!h]

Coefficient	Mean	95% HDI lower	95% HDI upper
Intercept	0.36	0.27	0.46
religion	0.83	0.52	1.19
Bisexual	1.91	1.34	2.5
Homosexual	2.32	1.29	3.49
Age 26-34	0.53	0.35	0.71
Age 35-49	0.38	0.24	0.53
Age 50+	0.3	0.09	0.54
Religion x Bisexual	1.04	0.51	1.62
Religion x Homosexual	0.58	0.12	1.18
Religion x Age 26-34	1.33	0.64	2.11
Religion x Age 35-49	1.1	0.5	1.84
Religion x Age 50+	1.42	0.32	3
WAIC = 2061.658			

Figure 6: Trace plots corresponding to six primary covariates included in the crude model.

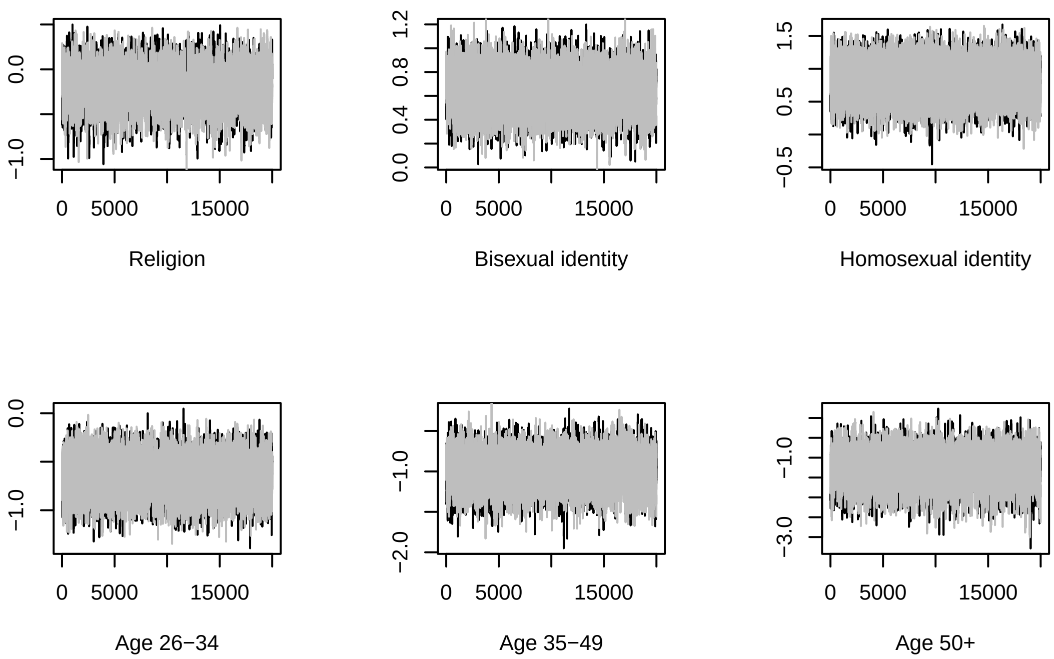


Table 7: Reduced model with the missing mechanism jointly fitted using the vague priors relating to missing values.

[!h]

Coefficient	Mean	95% HDI lower	95% HDI upper
Intercept	0.03	0.01	0.05
religion	0.91	0.56	1.3
Bisexual	1.7	1.18	2.24
Homosexual	2.18	1.16	3.31
Age 26-34	0.58	0.39	0.8
Age 35-49	0.45	0.27	0.64
Age 50+	0.39	0.12	0.7
Religion x Bisexual	0.88	0.45	1.41
Religion x Homosexual	0.52	0.1	1.07
Religion x Age 26-34	1.2	0.58	1.92
Religion x Age 35-49	0.98	0.45	1.65
Religion x Age 50+	1.29	0.28	2.72
WAIC = 1974.06			

Figure 7: Trace plots corresponding to six primary covariates included in the reduced model where the missing mechanism model is fitted with the vague priors.

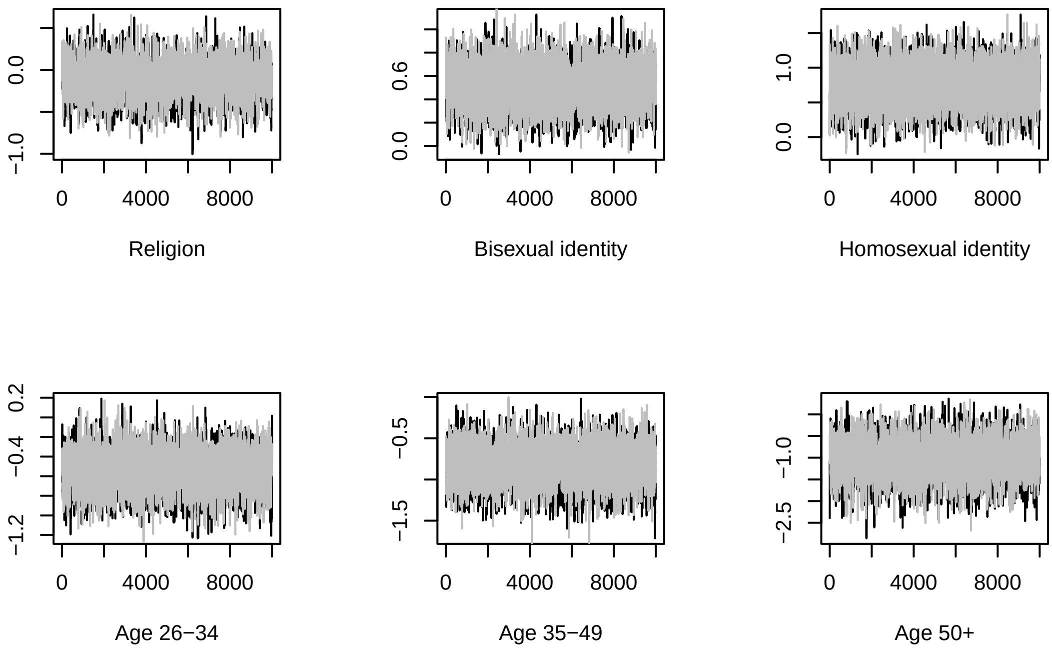


Table 8: Reduced model with the missing mechanism jointly fitted using the informative priors relating to missing values.

[!h]

Coefficient	Mean	95% HDI lower	95% HDI upper
Intercept	0.03	0.01	0.05
religion	0.91	0.56	1.29
Bisexual	1.71	1.18	2.25
Homosexual	2.2	1.18	3.35
Age 26-34	0.59	0.38	0.8
Age 35-49	0.45	0.27	0.65
Age 50+	0.38	0.11	0.7
Religion x Bisexual	0.87	0.44	1.38
Religion x Homosexual	0.52	0.11	1.09
Religion x Age 26-34	1.19	0.57	1.89
Religion x Age 35-49	0.99	0.42	1.61
Religion x Age 50+	1.27	0.28	2.67
WAIC = 1968.959			

Figure 8: Trace plots corresponding to six primary covariates included in the reduced model where the missing mechanism model is fitted with the informative priors.

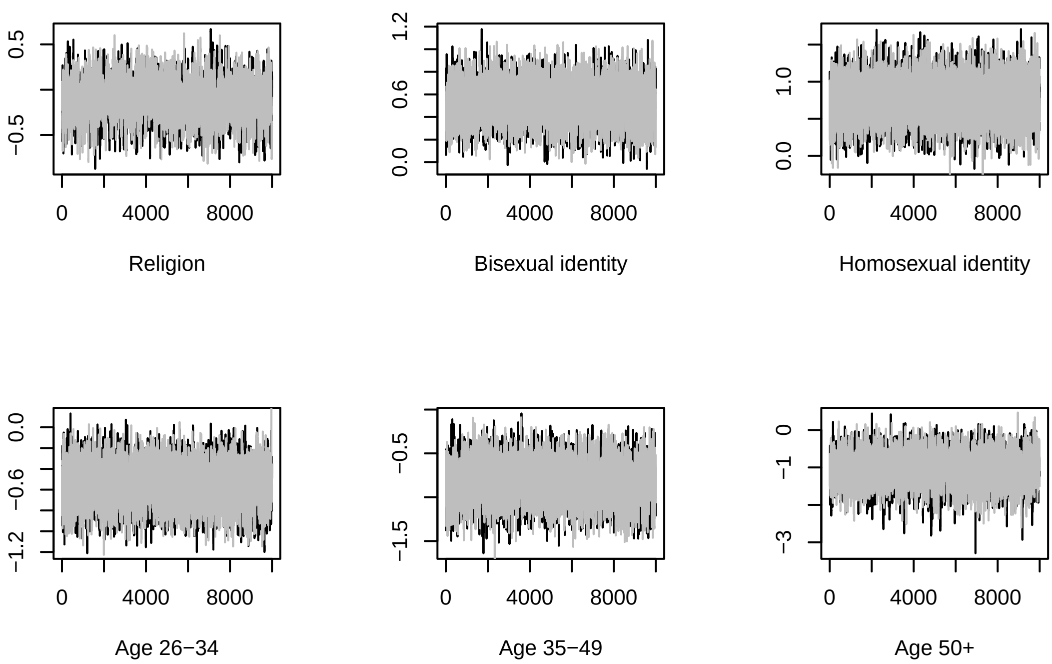


Table 9: Gelman and Rubin's convergence diagnostic of primary coefficients of the Spike-and-Slab model.

coefficient	Point est.	Upper C.I.
Religion	1.000255	1.000314
Bisexual identity	1.000098	1.000503
Homosexual identity	1.000053	1.000053
Age 26-34	1.000190	1.000589
Age 35-49	1.000015	1.000113
Age 50+	1.000831	1.002038

Table 10: Gelman and Rubin's convergence diagnostic of primary coefficients of the LASSO.

coefficient	Point est.	Upper C.I.
Religion	0.9999984	1.0000280
Bisexual identity	1.0000954	1.0003258
Homosexual identity	1.0006919	1.0035574
Age 26-34	0.9999769	0.9999819
Age 35-49	1.0000402	1.0002474
Age 50+	1.0000020	1.0000838

```
# Spike and Slab ------------------------------------------------

dat<- read_csv(file = "01_data/analysis_dat.csv")[,-1]

# define the data for jags
model_dat = list(
  #----------------------------------------------
  n = nrow(dat),
  #-----------------------------------
  suicidal = (dat$suicthink == "yes")*1,

  Xr = cbind(
    religion = (dat$religion=="agree")*1, # exposure
```

```
  bisex = ifelse(dat$sexident == "bisexual", 1, 0), # sex
  hmsex = ifelse(dat$sexident == "homosexual", 1, 0), # sex

  age26.34 = (dat$age == "26-34")*1, # age
  age35.49 = (dat$age == "35-49")*1, # age
  age50 = (dat$age == "50+")*1, # age

  religion.bisex = (dat$religion=="agree")*ifelse(dat$sexident ==
   ↪  "bisexual", 1, 0),
  religion.hmsex =(dat$religion=="agree")* ifelse(dat$sexident ==
   ↪  "homosexual", 1, 0),
  religion.age26.34 = (dat$religion=="agree")*(dat$age == "26-34"),
  religion.age35.49 = (dat$religion=="agree")*(dat$age == "35-49"),
  religion.age50 = (dat$religion=="agree")*(dat$age == "50+")
),

Xc = cbind(
  gender = dat$gender|> factor()|> as.numeric(),
  race = dat$race|> factor(levels = c("white", "black", "asian"))|>
   ↪  as.numeric(),
  income = dat$income|> factor()|> as.numeric(),
  work = dat$work|> factor(levels = c("unemployed", "part-time",
   ↪  "full-time"))|> as.numeric() ,
  married = dat$married|> factor()|> as.numeric(),
  health = dat$health,
  emodistr = dat$emodistr,
  sleeprob = dat$sleeprob|> factor()|> as.numeric(),
  educ = dat$educ|> factor()|> as.numeric(),
  drug = dat$drug|> factor()|> as.numeric()
```

```
)|> apply(2,\(i) (i-mean(i))/sd(i))
)

# model string

model_string <- textConnection("

model{

for (i in 1:n){
suicidal[i] ~ dbin(p.bound[i], 1)
p.bound[i]<- max(0, min(1, p[i]))

logit(p[i])<-
# intercept
b0 +

# required coefficients
inprod(Xr[i,], b[])+

# candidate coefficients
inprod(Xc[i,], beta[])
}

for(j in 1:length(Xc[1,])){
beta[j]<- gamma[j]*delta[j]
gamma[j] ~ dbern(0.5)
delta[j] ~ dnorm(0, tau)
}
```

```
b0 ~ dnorm(0, 0.01)

for(k in 1:length(Xr[1,])){b[k] ~ dnorm(0, 0.01)}
tau ~ dgamma(0.1, 0.1)
 }")


model <- jags.model(model_string,data = model_dat, n.chains=2)
update(model,50000, progress.bar="none")

params_name <- c ("gamma","beta")
samples <- coda.samples(model,
                        variable.names=params_name,
                        n.iter=20000, progress.bar="none");
beepr::beep()
#gelman.diag(samples)
sapply(1:ncol(samples[[1]]), \(i) gelman.diag(list(samples[[1]][,i],
↪  samples[[2]][,i] ))[[1]][,2] )

samp.spike.slab<- samples
# save(samp.spike.slab, file = "01_data/samp_spike_slab.RData")

#----------------------------------------------------------------------
# LASSO

dat<- read_csv(file = "01_data/analysis_dat.csv")[,-1]

# define the data for jags
```

```
model_dat = list(
  #---------------------------------------------
  n = nrow(dat),
  #----------------------------------
  suicidal = (dat$suicthink == "yes")*1,

  Xr = cbind(
    religion = (dat$religion=="agree")*1, # exposure

    bisex = ifelse(dat$sexident == "bisexual", 1, 0), # sex
    hmsex = ifelse(dat$sexident == "homosexual", 1, 0), # sex

    age26.34 = (dat$age == "26-34")*1, # age
    age35.49 = (dat$age == "35-49")*1, # age
    age50 = (dat$age == "50+")*1, # age

    religion.bisex = (dat$religion=="agree")*ifelse(dat$sexident ==
     ↪  "bisexual", 1, 0),
    religion.hmsex =(dat$religion=="agree")* ifelse(dat$sexident ==
     ↪  "homosexual", 1, 0),
    religion.age26.34 = (dat$religion=="agree")*(dat$age == "26-34"),
    religion.age35.49 = (dat$religion=="agree")*(dat$age == "35-49"),
    religion.age50 = (dat$religion=="agree")*(dat$age == "50+")
  ),

  Xc = cbind(
    gender = dat$gender|> factor()|> as.numeric(),
    race = dat$race|> factor(levels = c("white", "black", "asian"))|>
     ↪  as.numeric(),
```

```
    income = dat$income|> factor()|> as.numeric(),
    work = dat$work|> factor(levels = c("unemployed", "part-time",
     ↪  "full-time"))|> as.numeric() ,
    married = dat$married|> factor()|> as.numeric(),
    health = dat$health,
    emodistr = dat$emodistr,
    sleeprob = dat$sleeprob|> factor()|> as.numeric(),
    educ = dat$educ|> factor()|> as.numeric(),
    drug = dat$drug|> factor()|> as.numeric()
  )|> apply(2,\(i) (i-mean(i))/sd(i))

)

# model string

model_string <- textConnection("

model{

for (i in 1:n){
suicidal[i] ~ dbin(p.bound[i], 1)
p.bound[i]<- max(0, min(1, p[i]))

logit(p[i])<-
# intercept
b0 +

# required coefficients
inprod(Xr[i,], b[])+
```

```
# candidate coefficients
inprod(Xc[i,], beta[])
}

for(j in 1:length(Xc[1,])){beta[j] ~  ddexp(0, taubeta)}
taubeta ~ dgamma(0.1, 0.1)

b0 ~ dnorm(0, 0.01)
for(k in 1:length(Xr[1,])){b[k] ~ dnorm(0, 0.01)}
 }")


model <- jags.model(model_string,data = model_dat, n.chains=2)
update(model,50000, progress.bar="none")

params_name <- c ("beta")
samples <- coda.samples(model,
                        variable.names=params_name,
                        n.iter=20000, progress.bar="none");
beepr::beep()
gelman.diag(samples)
summary(samples)

samp.lasso =  samples

# save(samp.lasso, file = "01_data/samp_lasso.RData")

#-------------------------------------------------------------------
```

```
# Full model

dat<- read_csv(file = "01_data/analysis_dat.csv")[,-1]

#-----

#-----

model_dat = list(
  n = nrow(dat),
  r = is.na(dat$suicthink)*1,
  suicthink = (dat$suicthink == "yes")*1, # outcome
  #-----------------------------------
 Xr =  cbind(
   religion = (dat$religion=="agree")*1, # exposure

   bisex = ifelse(dat$sexident == "bisexual", 1, 0), # sex
   hmsex = ifelse(dat$sexident == "homosexual", 1, 0), # sex

   age26.34 = (dat$age == "26-34")*1, # age
   age35.49 = (dat$age == "35-49")*1, # age
   age50 = (dat$age == "50+")*1, # age

   religion.bisex = (dat$religion=="agree")*ifelse(dat$sexident == "bisexual",
   ↪  1, 0),
   religion.hmsex =(dat$religion=="agree")* ifelse(dat$sexident ==
   ↪  "homosexual", 1, 0),
   religion.age26.34 = (dat$religion=="agree")*(dat$age == "26-34"),
   religion.age35.49 = (dat$religion=="agree")*(dat$age == "35-49"),
```

```
  religion.age50 = (dat$religion=="agree")*(dat$age == "50+")
),

#------------------------

Xadj = cbind(
  gender = (dat$gender=="female")*1,

  race.asian = (dat$race == "asian")*1,
  race.black = (dat$race == "black")*1,

  income30.49 = (dat$income %in% c("30-39.9", "50-74.9", "75+"))*1,
  income.50.74 = (dat$income %in% c( "50-74.9", "75+"))*1,
  income.75 = (dat$income %in% c("75+"))*1,

  employ.full = (dat$work == "full-time")*1,
  employ.part = (dat$work == "part-time")*1,

  marry.single = (dat$married == "single")*1,
  marry.married = (dat$married == "married")*1,

  health = dat$health,

  emodistr = dat$emodistr,

  sleeprob = (dat$sleeprob == "yes")*1,

  educ.high = (dat$educ == "highschool")*1,
  educ.somecoll = (dat$educ == "somecoll")*1,
```

```
    educ.collgrad = (dat$educ == "collgrad")*1,


    drug = (dat$drug=="yes")*1
  )
)



model_string <- textConnection("
model{


for (i in 1:n){
suicthink[i] ~ dbin(p.bound[i], 1)
p.bound[i]<- max(0, min(1, p[i]))
logit(p[i])<-


# intercept
b0 +


# required coefficients


inprod(Xr[i,],b[])+




# adjusted covariates
inprod(Xadj[i,], beta[])


# for WAIC
like[i]<- dbin(suicthink[i],p.bound[i], 1)
}
```

```
# missing mechanism
for(i in 1:n){
r[i] ~ dbern(pmissing[i])
logit(pmissing[i]) <- b.miss.intercept + b.miss.suicthink*suicthink[i]+
b.miss.health*Xadj[i,11]+ b.miss.emodistr*Xadj[i,12]
}

# prior - analysis model
b0 ~ dnorm(0, 1/10^2)

for(k in 1:length(Xr[1,])){b[k] ~ dnorm(0,1/10^2)}
for(k in 1:length(Xadj[1,])){beta[k] ~ dnorm(0, 1/10^2)}



# prior - missing mechanism
b.miss.intercept ~ dnorm(0,1/10^2)
b.miss.suicthink ~ dnorm(0, 1/10^2)
b.miss.health ~ dnorm(0,1/10^2)
b.miss.emodistr ~ dnorm(0,1/10^2)

 }")

model <- jags.model(model_string, data = model_dat, n.chains=2)

update(model,50000, progress.bar="none")

params_name <- c ("b0", "b", "beta", "like")
```

```
samples <- coda.samples(model,
                        variable.names=params_name,
                        n.iter=20000, progress.bar="none");
beepr::beep()
# gelman.diag(samples)

gelman.diag(lapply(samples, \(i) i[,str_detect(colnames(i), "b")] ))



samp.full<-  samples

# save(samp.full, file = "01_data/samp_full.RData")

#-----------------------------------------------------------------------
# Reduce model 1 (vague priors)
dat<- read_csv(file = "01_data/analysis_dat.csv")[,-1]
#-----

#-----

model_dat = list(
  n = nrow(dat),
  r = is.na(dat$suicthink)*1,
  suicthink = (dat$suicthink == "yes")*1, # outcome
  #-----------------------------------
  Xr =  cbind(
    religion = (dat$religion=="agree")*1, # exposure

    bisex = ifelse(dat$sexident == "bisexual", 1, 0), # sex
```

```
  hmsex = ifelse(dat$sexident == "homosexual", 1, 0), # sex

  age26.34 = (dat$age == "26-34")*1, # age
  age35.49 = (dat$age == "35-49")*1, # age
  age50 = (dat$age == "50+")*1, # age

  religion.bisex = (dat$religion=="agree")*ifelse(dat$sexident ==
   ↪  "bisexual", 1, 0),
  religion.hmsex =(dat$religion=="agree")* ifelse(dat$sexident ==
   ↪  "homosexual", 1, 0),
  religion.age26.34 = (dat$religion=="agree")*(dat$age == "26-34"),
  religion.age35.49 = (dat$religion=="agree")*(dat$age == "35-49"),
  religion.age50 = (dat$religion=="agree")*(dat$age == "50+")
),

#-------------------------

Xadj = cbind(

  income30.49 = (dat$income %in% c("30-39.9", "50-74.9", "75+"))*1,
  income.50.74 = (dat$income %in% c( "50-74.9", "75+"))*1,
  income.75 = (dat$income %in% c("75+"))*1,

  health = dat$health,

  emodistr = dat$emodistr,

  sleeprob = (dat$sleeprob == "yes")*1
)
```

```
)



model_string <- textConnection("
model{

for (i in 1:n){
suicthink[i] ~ dbin(p.bound[i], 1)
p.bound[i]<- max(0, min(1, p[i]))
logit(p[i])<-

# intercept
b0 +

# required coefficients

inprod(Xr[i,],b[])+



# adjusted covariates
inprod(Xadj[i,], beta[])

# for WAIC
like[i]<- dbin(suicthink[i],p.bound[i], 1)
}

# missing mechanism
for(i in 1:n){
r[i] ~ dbern(pmissing[i])
```

```
logit(pmissing[i]) <- b.miss.intercept + b.miss.suicthink*suicthink[i]+
b.miss.health*Xadj[i,4]+ b.miss.emodistr*Xadj[i,5]
}

# prior - analysis model
b0 ~ dnorm(0, 1/10^2)

for(k in 1:length(Xr[1,])){b[k] ~ dnorm(0,1/10^2)}
for(k in 1:length(Xadj[1,])){beta[k] ~ dnorm(0, 1/10^2)}


# prior - missing mechanism
b.miss.intercept ~ dnorm(0,1/10^2)
b.miss.suicthink ~ dnorm(0, 1/10^2)
b.miss.health ~ dnorm(0,1/10^2)
b.miss.emodistr ~ dnorm(0,1/10^2)

 }")

model <- jags.model(model_string, data = model_dat, n.chains=2)
update(model,50000, progress.bar="none")

params_name <- c ("b0", "b", "beta", "like")

samples <- coda.samples(model,
                        variable.names=params_name,
                        n.iter=10000, progress.bar="none");
beepr::beep()
gelman.diag(lapply(samples, \(i) i[,str_detect(colnames(i), "b")] ))
```

```
summary(samples)


samp.reduced<- samples
# save(samp.reduced, file = "01_data/samp_reduced.RData")


#-----------------------------------------------------------------------
# reduced model 2 (informative)


dat<- read_csv(file = "01_data/analysis_dat.csv")[,-1]
#-----


missing.data = tibble(vars = c("intercept", "suicide", "health", "depressed"),
                      odds = c(0.02/(1-0.02), 0.48, 0.63, 0.63),
                      lower = c(0.01, 0.37, 0.51, 0.42),
                      upper = c(0.03, 0.62, 0.78, 0.94))|>
  mutate(ln.odds = log(odds), variance =
   ↪  (log(upper/lower)/2/qnorm(0.975))^2*719)
miss.prior <- select(missing.data, ln.odds, variance)|> as.matrix()|>
↪  `rownames<-`(missing.data$vars)


#-----


model_dat = list(
  n = nrow(dat),
  r = is.na(dat$suicthink)*1,
  suicthink = (dat$suicthink == "yes")*1, # outcome
  #-----------------------------------
  Xr =  cbind(
    religion = (dat$religion=="agree")*1, # exposure
```

```r
  bisex = ifelse(dat$sexident == "bisexual", 1, 0), # sex
  hmsex = ifelse(dat$sexident == "homosexual", 1, 0), # sex

  age26.34 = (dat$age == "26-34")*1, # age
  age35.49 = (dat$age == "35-49")*1, # age
  age50 = (dat$age == "50+")*1, # age

  religion.bisex = (dat$religion=="agree")*ifelse(dat$sexident ==
   ↪  "bisexual", 1, 0),
  religion.hmsex =(dat$religion=="agree")* ifelse(dat$sexident ==
   ↪  "homosexual", 1, 0),
  religion.age26.34 = (dat$religion=="agree")*(dat$age == "26-34"),
  religion.age35.49 = (dat$religion=="agree")*(dat$age == "35-49"),
  religion.age50 = (dat$religion=="agree")*(dat$age == "50+")
),

#------------------------

Xadj = cbind(

  income30.49 = (dat$income %in% c("30-39.9", "50-74.9", "75+"))*1,
  income.50.74 = (dat$income %in% c( "50-74.9", "75+"))*1,
  income.75 = (dat$income %in% c("75+"))*1,

  health = dat$health,

  emodistr = dat$emodistr,
```

```
    sleeprob = (dat$sleeprob == "yes")*1
  ),
  miss.prior = miss.prior
)



model_string <- textConnection("
model{

for (i in 1:n){
suicthink[i] ~ dbin(p.bound[i], 1)
p.bound[i]<- max(0, min(1, p[i]))
logit(p[i])<-

# intercept
b0 +

# required coefficients

inprod(Xr[i,],b[])+



# adjusted covariates
inprod(Xadj[i,], beta[])

# for WAIC
like[i]<- dbin(suicthink[i],p.bound[i], 1)
}
```

```
# missing mechanism
for(i in 1:n){
r[i] ~ dbern(pmissing[i])
logit(pmissing[i]) <- b.miss.intercept + b.miss.suicthink*suicthink[i]+
b.miss.health*Xadj[i,4]+ b.miss.emodistr*Xadj[i,5]
}

# prior - analysis model
b0 ~ dnorm(0, 1/10^2)

for(k in 1:length(Xr[1,])){b[k] ~ dnorm(0,1/10^2)}
for(k in 1:length(Xadj[1,])){beta[k] ~ dnorm(0, 1/10^2)}



# prior - missing mechanism
b.miss.intercept ~ dnorm(miss.prior[1,1], miss.prior[1,2])
b.miss.suicthink ~ dnorm(miss.prior[2,1], miss.prior[2,2])
b.miss.health ~ dnorm(miss.prior[3,1], miss.prior[3,2])
b.miss.emodistr ~ dnorm(miss.prior[4,1], miss.prior[4,2])

 }")

model <- jags.model(model_string, data = model_dat, n.chains=2)
update(model,50000, progress.bar="none")

params_name <- c ("b0", "b", "beta", "like")

samples <- coda.samples(model,
                        variable.names=params_name,
```

```
                              n.iter=10000, progress.bar="none");
beepr::beep()
gelman.diag(lapply(samples, \(i) i[,str_detect(colnames(i), "b")] ))


samp.reduced.info<-  samples
# save(samp.reduced.info, file = "01_data/samp_reduced_info.RData")
```

References

1. Garnett MF, Curtin SC. Suicide mortality in the united states, 2002–2022. 2024;

2. Poorolajal J, Goudarzi M, Gohari-Ensaf F, Darvishi N. Relationship of religion with suicidal ideation, suicide plan, suicide attempt, and suicide death: A meta-analysis. Journal of Research in Health Sciences. 2022;22(1):e00537.

3. Lawrence RE, Brent D, Mann JJ, Burke AK, Grunebaum MF, Galfalvy HC, et al. Religion as a risk factor for suicide attempt and suicide ideation among depressed patients. The Journal of nervous and mental disease. 2016;204(11):845–50.

4. Park K, Hsieh N. A national study on religiosity and suicide risk by sexual orientation. American journal of preventive medicine. 2023;64(2):235–43.

5. Gibbs JJ, Goldbach J. Religious conflict, sexual identity, and suicidal behaviors among LGBT young adults. Archives of suicide research. 2015;19(4):472–88.

6. Daniels MJ, Hogan JW. Missing data in longitudinal studies: Strategies for bayesian modeling and sensitivity analysis. chapman; hall/CRC; 2008.

7. Mérelle S, Foppen E, Gilissen R, Mokkenstorm J, Cluitmans R, Van Ballegooijen W. Characteristics associated with non-disclosure of suicidal ideation in adults. International journal of environmental research and public health. 2018;15(5):943.

8. Turkkan N, Pham-Gia T. Computation of the highest posterior density interval in bayesian analysis. Journal of statistical computation and simulation. 1993;44(3-4):243–50.

9. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. Journal of Thoracic Oncology. 2010;5(9):1315–6.