The Performance of Bayesian Analysis in Structural Equation Modelling to Construct The Health Behaviour During Pandemic COVID-19

Ferra Yanuar[1]* and Aidinil Zetra[2]

* Corresponding Author

[1]Department of Mathematics, Faculty of Mathematics and Natural Science, Andalas University,
 Kampus Limau Manis, 25163, Padang – Indonesia, ferrayanuar@sci.unand.ac.id
[2]Department of Political Science, Faculty of Social Science and Political Science, Andalas University,
 Kampus Limau Manis, 25163, Padang – Indonesia, aidinil@soc.unand.ac.id

Abstract

Originating from Wuhan, China, COVID-19 is spreading rapidly throughout the world. The epidemiological model is required to provide evidence for public health policymakers to reduce the spread of COVID-19. Health behaviour is assumed could reduce the spread of this virus. This study purposes to construct an acceptable model of health behaviour. To achieve this goal, a Bayesian structural equation modelling (SEM) is implemented. This current study is also purposed to evaluate the performance of Bayesian SEM, including the sensitivity, adequacy, and the acceptability of parameters estimated with the result that the acceptable model is obtained. The sensitivity of the Bayesian SEM estimator is evaluated by choosing several types of prior and the model results are compared. The adequacy of the Bayesian SEM estimate is checked by doing the convergence test of the corresponding model parameters. The acceptability of the Bayesian approach and its associated algorithm in recovering the true parameters are monitored by the Bootstrap simulation study. The Bayesian SEM applies the Gibbs sample approach in estimating model parameters. This method is applied to the primary data gathered from an online survey from March to May 2020 during COVID-19 to individuals living in West Sumatera, Indonesia. It is found that health motivation is significantly related to health behaviour. Whereas socio-demographic and perceived susceptibility has no significant effect on health behaviour.

Key Words: Structural Equation Modelling; Bayesian SEM; Health Behaviour; Health Motivation

Mathematical Subject Classification: 62C10, 62F10, 62F15.

1. Introduction

The first cases of COVID-19 (coronavirus disease 2019) were found in the city of Wuhan, Hubei Province, China in December 2019 (Hou et al., 2020). The pneumonia-causing virus was identified as the new coronavirus named a severe acute respiratory syndrome (SARS-CoV) 2, and the disease was named COVID-19 by World Health Organization (Suksatan et al., 2021; Zhu et al., 2020). Within two months, the infection spread rapidly to other countries and regions. As of May 4, 2021, worldwide, there were 154,159,212 cases of COVID-19 and 3,226,235 death related to it (Worldometers, 2021). In Indonesia, 1,682,004 cases with 45,949 death among them have been confirmed. Several policies have been established to prevent the spread of this pandemic, including in Indonesia. The policies include a strict city-wide lockdown and a ban on foreign tourists from entering the country. These policies have considerably influenced people's lives, including changing their health habits and lifestyle. It needs someone to wear masks, wash their hands with soap or alcohol hand sanitizer, avoid crowded places, etc. (Choompunuch et al., 2021). Health behaviour is needed to protect people from illness. Khoso et al (2016) state that health behaviour is a health-related activity to maintain good health and prevent potential health hazards.

Health behaviour contains several primary concepts that predict why people will take action to prevent, screen for, or control illness conditions; these include susceptibility, seriousness, benefits, and barriers to a behaviour, cues to action, and most recently, self-efficacy (Zvolensky et al., 2020). Health behaviour is achieved

since individuals expect to avoid diseases. They expect that a specific health action may prevent illness besides individuals having the motivation to stay healthy. The expectancy was further delineated in terms of the individual's estimates of personal susceptibility to and perceived susceptibility to an illness and the likelihood of being able to reduce that threat through personal action (Conner & Norman, 2007; Glanz et al., 2008).

Because of the COVID-19 pandemic, it's important to model the health behaviour of individuals to prevent the spread of this disease (Perlman, 2020). Various studies have been carried out to identify the factors of health behaviour. Corner and Norman (Conner & Norman, 2007) identified that individuals have health motivation and perceived susceptibility that affects their health behaviour to avoid illness. Glanz et al. (2008) modeled that perceived susceptibility varies for different ages, gender, and ethnicity. They found that perceived susceptibility could affect the individual's health behaviour.

Health behaviour, socio-demographics, health motivation, and perceived susceptibility cannot be measured directly from the research object unless measured by observable variables. These factors are known as unobservable or latent variables. The structural equation modelling (SEM) approach is particularly suitable for modelling the interrelationship between these latent variables and their corresponding indicators (Kelava et al., 2014; Olsson et al., 2000; Yanuar, 2016; Yanuar et al., 2010). SEM assumes that the observations are independent and identically distributed using the multivariate normal distribution. If this assumption is not met, the estimate of the parameter model could be difficult to obtain in the usual way. Therefore, many researchers, such as Asparouhov & Muthén (2021) and Yanuar et al. (2013), proposed using the Bayesian approach in SEM to overcome these problems. This method uses a Gibbs sampler to obtain samples of arbitrary sizes for summarizing the posterior distribution for describing the parameters of interest (Nam et al., 2018; Y. Thanoon et al., 2016). The user can compute point estimates, standard deviations, and interval estimates from these samples to make an inference. The Bayesian approach is appealing because it allows the user to use prior information to update current information about the parameters of interest (Asparouhov & Muthén, 2021; Cain & Zhang, 2019). Because of this framework, parameter estimation based on Bayesian is more precise than standard SEM (Garnier-Villarreal & Jorgensen, 2020; Yanuar et al., 2013). Accordingly, the main purpose of this study is to evaluate the performance of Bayesian SEM in constructing the health behaviour model during the COVID-19 pandemic of an individual living in West Sumatra. The performance of Bayesian SEM will be evaluated in terms of its adequacy, acceptability, and sensitivity, since the traditional fit indices, such as RMSEA, CFI or TLI are not available when performing Bayesian SEM (Asparouhov & Muthén, 2021; Cain & Zhang, 2019). Not much works have been done on modelling of health behaviour using Bayesian SEM, particularly when information on socio-demographics, perceived susceptibility, health motivation are considered. This study becomes important to be done in the condition of pandemic COVID-19 outbreak.

2. Material and Method

2.1. Data

The data in this study was obtained by distributing the online questionnaires and house visits between March to May 2020. Online questionnaires are conducted because due to the physical distancing policy. The questionnaires are disseminated via WhatsApp and email. A small number of respondents are chosen randomly based on the house visit. The house visit is chosen randomly based on the list of available households with the enumerators. The respondents eligible to participate in the survey are those who are 18 years old or older and living in West Sumatra, representing the adult population of West Sumatra. More than 1000 respondents participated in this survey. However, 756 respondents with complete information were involved in the analysis. The information gathered in the survey includes information about socio-demographics, perceived susceptibility, health motivation, and health behaviour of individuals living in West Sumatra, a province of Indonesia. The province has an area of 42,012.89 km^2, with a population of 5,534,472 at the 2021 census. The province is subdivided into twelve regencies and seven cities. Padang is the province's capital and the largest city. Table 1 provides the descriptive statistics of respondents involved in this study.

2.2   Factors of Health Behaviour

The health behaviour can be in the form of preventive behaviour or health facilities. Khoso et al. (2016) said that Health behaviour is influenced by Health belief factors that can form healthy living behaviour during the COVID-19 pandemic to avoid the transmission of this disease outbreak. Thus, the model is used in explaining and predicting preventive health behaviour, as well as sick-role and illness behaviour. This health behaviour model has been applied to all studies of health behaviour. A person's motivation to undertake a healthy behaviour can be

divided into two main categories: individual perceptions and modifying behaviour. Individual perceptions are factors that affect the perception of illness or disease. They deal with the importance of health to the individual, in this study named it health motivation. Modifying factors that are included in this study are socio-demographic and perceived susceptibility. Combining these factors causes a response that often manifests into action, provided it is accompanied by an alternative rational course of action to having healthy behaviour (Arora & Grey, 2020).

Table 1. Descriptive Statistics of Respondents

Variable	Category	Frequency	Percentage (%)
Gender	1. Male	325	42.99
	2. Female	431	57.01
Educational level	1. Junior high school	51	6.75
	2. Senior high school	98	12.96
	3. Associate's degree	102	13.49
	4. Bachelor degree	327	43.25
	5. Masters/PhD	178	23.54
Age	1. Between 18 and 20 years old,	27	3.57
	2. Between 21 and 30 years old,	99	13.10
	3. Between 31 and 40 years old,	268	35.45
	4. Between 41 and 50 years old,	215	28.44
	5. Between 51 and 60 years old,	91	12.04
	6. More than 60 years old,	56	7.40

Socio-demographics, health motivation, perceived susceptibility, and health behaviour are unobservable variables measured through indicator variables. The following briefly explains each unobservable variable considered in this study and its respective indicators. The indicators used for describing the socio-demographic factor are age, gender, and educational level (Suksatan et al., 2021). The age of the respondents is classified into six levels, which are 'between 18 and 20 years old', 'between 21 and 30 years old', 'between 31 and 40 years old', 'between 41 and 50 years old', 'between 51 and 60 years old', and 'more than 60 years old', coded as 1 to 6 respectively. For respect to the gender of the respondents, the responses obtained are coded 1 as 'male' and 2 as 'female'. The responses to educational level are coded as 1 as 'attend elementary school', 2 as 'attend junior high school', 3 as 'senior high school' and 4 as 'attend college/university'. The indicators for health motivation as hypothesized in this study are based on a study by Arora & Grey (2020). The indicators considered in this study are avoiding touching the face, avoid shaking hands, avoid meeting or standing in long queues, avoid touching objects in public areas, avoid taking public transportation (online), avoid going home, avoid worshiping in mosque/church/others, and avoid ordering online food. The responses obtained for each indicator are coded 1 for 'never', 2 for 'seldom', 3 for 'sometimes', 4 for 'often', and 5 for 'always'. Meanwhile, perceived susceptibility as a latent variable is measured by indicator variables based on the study by Conner & Norman (2007). The respondents were asked about their worries regarding personal health, the health of their family members, when they go outside, when they are going to their village, work/school, and using public facilities / public transportation, and food availability. The responses obtained for each indicator are coded as 1 to 5 for denoting 'not worried at all', 'not worried', 'a little worried', 'worried', and 'very worried', respectively.

Many of the health behaviours in response to SARS were similar to those recommended to prevent the spread of COVID-19 including frequent handwashing, social distancing, and self-quarantining (Parekh & Deierlein, 2020; Zvolensky et al., 2020). In this study, the indicators of health behaviours are keeping a distance of 2 m, wear a mask, hand sanitizer, wash hands for 20 seconds, inform others if having symptoms of COVID-19. The responses of each indicator are in 5 Likert scales which are considered to include 'never' and coded as 1, 'seldom' coded as 2, 'sometimes' coded as 3, 'often' coded as 4, and 'always' coded as 5.

Figure 1 provides the conceptual model of health behaviour based on the SEM structure constructed in this study. It is also reasonable to hypothesize that health behaviour is affected by socio-demographics, perceived susceptibility, and health motivation. The hypothesis indicators for each latent are provided in Table 2.
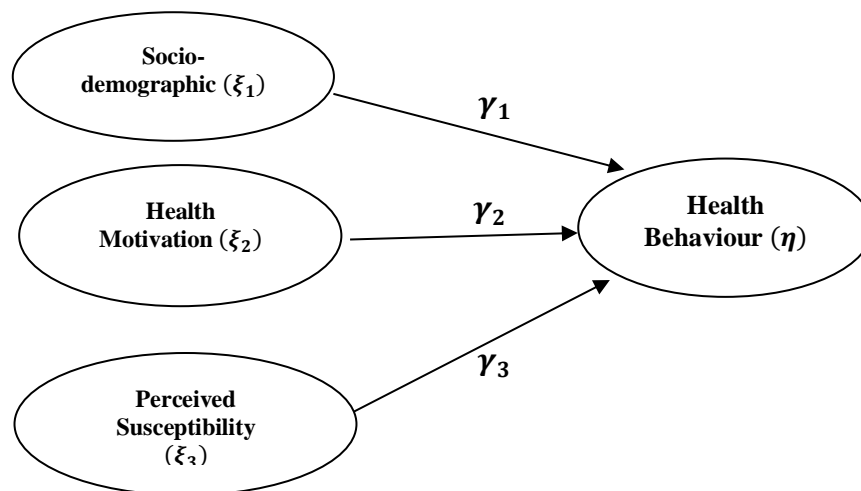
Figure 1. The Conceptual Model of Health Behaviour

Table 2. The Latent and Indicator Variables in the Conceptual Model

Latent Variable	Indicator Variables
Socio-demographic	1. Gender 2. Educational level 3. Age
Health Motivation	1. Avoid touching the face 2. Avoid shaking hands 3. Avoid meeting or standing in long queues 4. Avoid touching objects in public areas 5. Avoid taking public transportation (online) 6. Avoid going home 7. Avoid worshiping in mosque/church/others 8. Avoid ordering online food
Perceived susceptibility	1. Personal health 2. The health of their family members 3. When they go outside 4. When they are going to their village 5. To work/school, public facilities / public transportation 6. Food availability **7.** COVID-19 spread
Health Behaviour	1. Keep a distance of 2m 2. Wear a mask 3. Hand sanitizer 4. Wash hands for 20 seconds 5. Inform others if having symptoms of COVID

2.3 Estimation Method

SEM combines two types of equations simultaneously. These are structural equations and measurement equations. The measurement equation explains the relationship between the indicator variable to its latent variable, which is formulated as follows:

$$x_i = \Lambda\omega_i + \varepsilon_i, \quad i = 1, \dots, n \tag{1}$$

Where x_i represents indicator variable and ω_i is as latent variables, Λ is loading factors and ε_i is measurement errors. Meanwhile, the structural equation is the interrelationship among the latent factors and is formulated by:

$$\eta_i = \Gamma\xi_i + \delta_i , i = 1, ..., n. \tag{2}$$

Let the latent variable ω_i be partitioned into (η_i, ξ_i) where η_i and ξ_i are latent variables respectively, Γ is loading factors, and δ_i is structural error.

In this study, we implemented the Bayesian analysis to estimate the parameter model in order to construct the best model based on the SEM approach in the case of the normality assumption violated. In the Bayesian method, the ordered categorical data a treat as continuous normal distribution by using threshold specification. Here, we consider $X = (x_1, x_2, ... x_n)$ and $Y = (y_1, y_2, ... y_n)$ to be the ordered categorical data matrices and latent continuous variables respectively and $\Omega = (\omega_1, \omega_2, ..., \omega_n)$ be the matrix of latent variables. The observed data X are augmented with the latent data (Y, Ω) in the posterior analysis. In this subsection, we will apply the Bayesian estimation in SEM in posterior analysis to obtain the values of the unknown threshold in $\tau = (\tau_1, \tau_2)$, joint Bayesian estimates of Ω and the structural parameter $\boldsymbol{\theta}$, a vector that includes all the unknown parameters in $\boldsymbol{\Phi}, \boldsymbol{\psi}_\delta, \boldsymbol{\psi}_\varepsilon, \boldsymbol{\Lambda}$ and $\boldsymbol{\Lambda}_\omega$.

The Bayesian method is applied to derive the posterior distribution for $[\tau, \theta, \Omega | X]$. The researchers usually may use the maximum likelihood (ML) estimation method to derive the posterior distribution. However, it is difficult to apply the ML estimation method in this study since a high-dimensional integration is required. Thus, the Markov chain Monte Carlo (MCMC) technique with the Gibbs sampler algorithm is applied to solve this problem. This technique generates a sequence of random observations from the fully conditional posterior distribution of unknown model parameters. The user can create and implement the algorithm easily using WinBUGS. Since ordered categorical variables are used, the latent matrix Y should be augmented in the posterior analysis, resulting in the joint posterior distribution $[\tau, \theta, \Omega, Y | X]$. The Gibbs sampler process starts with the setting of initial starting values $(\tau^{(0)}, \theta^{(0)}, \Omega^{(0)}, Y^{(0)})$, and then conduct the simulation for $(\tau^{(1)}, \theta^{(1)}, \Omega^{(1)}, Y^{(1)})$. At the rth iteration, by making use of the using current values $(\tau^{(r)}, \theta^{(r)}, \Omega^{(r)}, Y^{(r)})$. The Gibbs sampler is carried out as follows (Muharisa et al., 2018):

b. Generate $\Omega^{(r+1)}$ from $p(\Omega | \tau^{(r)}, \theta^{(r)}, Y^{(r)}, X)$
c. Generate $\theta^{(r+1)}$ from $p(\theta | \Omega^{(r+1)}, \tau^{(r)}, Y^{(r)}, X)$
d. Generate $(\tau^{(r+1)}, Y^{(r+1)})$ from $p(\tau, Y | \Omega^{(r+1)}, \theta^{(r+1)}, X)$

The prior distribution for $(\Lambda, \psi_\varepsilon)$ and $\boldsymbol{\Phi}$ are needed in the process of determining the posterior distribution. In this present study, we take the following conjugate prior distribution for those three parameters. Letting $\psi_{\varepsilon k}$ be the kth diagonal element of ψ_ε and Λ_k be the kth row of Λ, we consider:

$$\psi_{\varepsilon k}^{-1} \sim Gamma(\alpha_{0\varepsilon k}, \beta_{0\varepsilon k}) \tag{3}$$

$$(\Lambda_k | \psi_{\varepsilon k}^{-1}) \sim N(\Lambda_{0k}, \psi_{\varepsilon k} \mathbf{H}_{0yk}) \tag{4}$$

$$\boldsymbol{\Phi}^{-1} \sim W_q(R_0, \rho_0) \tag{5}$$

where $Gamma(,)$ is the gamma distribution, $W_q(,)$ is an q dimensional Wishart distribution, parameters $\alpha_{0\varepsilon k}, \beta_{0\varepsilon k}, \Lambda_{0k}, \rho_0$, positive definite matrix ; H_{0yk} and R_0 are collectively hyperparameters that are assumed to be described by an uninformative prior distribution (Yanuar et al., 2013).

This present study suggests evaluating the performance of Bayesian analysis in SEM through three aspects, the sensitivity, adequacy, and acceptability of the Bayesian SEM estimate. It is important to test the sensitivity of the Bayesian analysis concerning the choice of the prior (Rahmadita et al., 2018). Several types of prior inputs can be constructed as long as they can be compared to the model results. The adequacy of the Bayesian SEM estimate is checked by doing the convergence test of the corresponding model parameters. The convergence is assessed by plotting the time series to assess the quality of the individual parameters with different starting values graphically, providing a diagnosis based on the trace plots and density plots (Lee, 2007; Yanuar et al., 2022). This study also uses the Brooks-Gelman-Rubin (BGR) convergence statistics which compares the variation between and within multiple chains, denoted by $\boldsymbol{R}$. The estimated parameters converge if the $\boldsymbol{R}$-value is close to 1. The accuracy of the posterior estimates is inspected by assuring that the Monte Carlo error (an estimate of the difference between the mean of the sampled values and the true posterior mean) for all the parameters to be less than 5% of the sample standard deviation (Depaoli and Clifton 2015). At the same time, the acceptability of the Bayesian approach and its associated algorithm in recovering the true parameters are monitored by doing a simulation study to construct the Bootstrap confidence interval (Zhang & Savalei, 2016). In this Bootstrap method, new data set is generated by sampling with replacement from the original data set, and then estimating the statistics for each new data set. Several model fits are determined to calculate the 95% confidence interval of all parameters model.

3.   Results and Discussion

In the first stage, we include all variables into the model and then fit the hypotheses model to the data, but an acceptable model is not achieved. We then exclude socio-demographics from the hypotheses model since this latent variable is not significant to give effect to the response, health behaviour. At this second stage, we estimate the goodness of fit of the proposed model and we could obtain the acceptable model. The results of the model fitting obtained based on Bayesian SEM are presented in Figure 2.
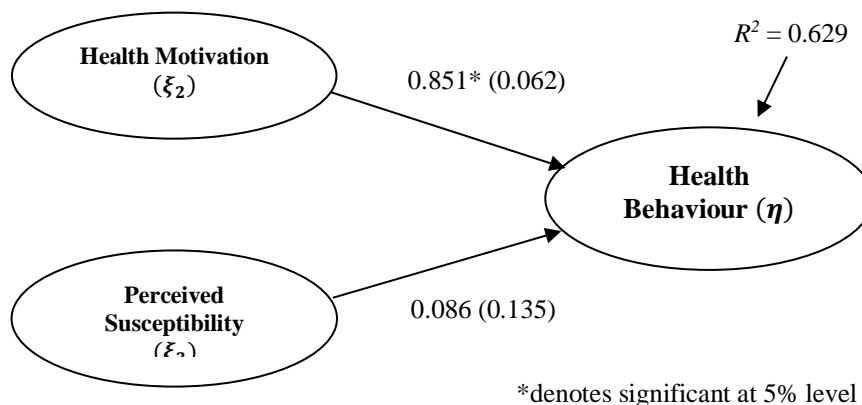

Figure 2. The Proposed Model of Health Behaviour

To test the sensitivity of the Bayesian analysis, we consider three types of prior inputs. The model resulted based on three different priors. In this present study, to have good prior information, we take a small variance for each parameter in assigning the values for the hyperparameter where all parameters are fixed based on a study by Yanuar et al. (2013). Therefore, all prior inputs are summarized as follows:

a.   Type I prior: the unknown loadings in Λ are all taken to be 0.5, those values corresponding to $\gamma_2 = 0.80$ and $\gamma_3 = 0.10$.
b.   Type II prior: the hyperparameters are equal to 75% of the values given in a.
c.   Type III prior: the hyperparameters are equal to 150% of the values given in a.

The results for Bayesian estimation of the structural and measurement equations obtained based on the three types of prior inputs are provided in Table 3.

Based on Table 3, the parameter estimates and standard errors resulting under three different priors are reasonably close. It could be concluded here that the statistics obtained based on the Bayesian SEM are robust to the different prior inputs. Or it can be said that the Bayesian SEM implemented here is not sensitive to these three different prior inputs. Accordingly, the results obtained using Type I prior are used as the proposed model based on the Bayesian SEM method in this present study. The values provided in Table 3 were estimated using WinBUGS 14 (Ntzoufras, 2009). Code in WinBugs is provided in Appendix.

The adequacy of Bayesian SEM estimated is evaluated using the convergence statistics test for all parameters model obtained. Plots of sequences of observations corresponding to $\hat{\gamma}_2$ and $\hat{\lambda}_{34}$ as a selected parameter for illustrative purposes are provided in Figures 3 (a) and (b). Both plots show that data lie within two parallel horizontal lines, which indicates that the algorithm converged in less than 15.000 iterations for two initial values. Density plot in Figure 3 (c) for $\hat{\gamma}_2$ and Figure 3 (d) for $\hat{\lambda}_{34}$ indicate that iteration samples have a normal distribution which informs that the parameter values resulting from using the algorithm have converged. The BGR plot for (a) $\hat{\gamma}_3$, (b) $\hat{\lambda}_{24}$ as a selected parameter is provided in Figure 4. We found that the R-value is close to 1, meaning the estimated parameters have converged. The Monte Carlo errors for all parameters yielded are less than 5% of the sample standard deviation. To assess the plausibility of the proposed model is checked by plotting the estimated residual versus the case number. The distribution of data is centered at zero and no trends are detected. It could be concluded here that the estimated parameter model resulting in this present study is considered adequate and could be accepted.

The next analysis is to determine the 95% bootstrap confidence interval to test the acceptability of the Bayesian approach and its associated algorithm in recovering the true parameters. Table 4 presents the results from the Bootstrap simulation study. It is clear from Table 4 that all parameter estimates fall within the 95% bootstrap confidence intervals taken from this simulation study. These results inform us that the Bootstrap method

seems to work well in this study. Thus, the estimated posterior mean obtained based on Bayesian SEM are acceptable since we believe that the power of our Bayesian SEM could result in the best fit for the proposed model. The relationship between health behaviour with its endogenous latent variables in the structural equation can be modeled as follow:

$$\hat{\eta}_{BaySEM} = 0.851\,\xi_2 + 0.086\,\xi_3, \quad R^2 = 0.629 \tag{6}$$

Table 3. The Results of Health Behaviour Model Using Bayesian SEM

| No	Parameter	Type I Prior Estimate (SE)	Type II Prior Estimate (SE)	Type III Prior Estimate (SE)
A	**Structural Equation**			
	Health Motivation → Health Behaviour (γ_2)	0.851* (0.062)	0.821* (0.076)	0.854* (0.072)
	Perceived susceptibility → Health Behaviour (γ_3)	0.086 (0.135)	0.077 (0.112)	0.076 (0.145)
B	**Measurement equation**			
	Health Motivation (ξ_2)			
1	Avoid touching the face (λ_{21})	0.529* (0.032)	0.505* (0.064)	0.516* (0.052)
2	Avoid shaking hands (λ_{22})	0.457* (0.064)	0.412* (0.067)	0.465* (0.061)
3	Avoid meeting or standing in long queues (λ_{23})	0.460* (0.065)	0.423* (0.075)	0.430* (0.072)
4	Avoid touching objects in public areas (λ_{24})	0.504* (0.063)	0.475* (0.076)	0.514* (0.071)
5	Avoid taking public transportation (online) (λ_{25})	0.534* (0.076)	0.530* (0.076)	0.501* (0.086)
6	Avoid going home (λ_{26})	0.512* (0.074)	0.499* (0.065)	0.452* (0.070)
7	Avoid worshiping in mosques/churches/others (λ_{27})	0.571* (0.073)	0.565* (0.067)	0.522* (0.063)
8	Avoid ordering online food (λ_{28})	0.420* (0.062)	0.390* (0.076)	0.409* (0.059)
	Perceived susceptibility (ξ_3)			
9	Personal health (λ_{31})	0.546* (0.055)	0.550* (0.051)	0.556* (0.059)
10	The health of their family members (λ_{32})	0.532* (0.057)	0.520* (0.045)	0.512* (0.065)
11	When they go outside (λ_{33})	0.425* (0.058)	0.430* (0.052)	0.403* (0.066)
12	When they are going to their village (λ_{34})	0.449* (0.058)	0.429* (0.063)	0.442* (0.059)
13	To work/school, public facilities / public transportation (λ_{35})	0.431* (0.058)	0.427* (0.054)	0.422* (0.065)
14	Food availability (λ_{36})	0.554* (0.060)	0.530* (0.063)	0.554* (0.060)
15	COVID-19 spread(λ_{37})	0.444* (0.058)	0.430* (0.054)	0.414* (0.062)
	Health Behaviour (η)			
16	Keep a distance of 2 m (η_1)	0.665* (0.043)	0.702* (0.042)	0.662* (0.047)
17	Wear a mask(η_2)	0.564* (0.058)	0.521* (0.053)	0.559* (0.060)
18	Hand sanitizer (η_3)	0.776* (0.032)	0.711* (0.041)	0.771* (0.038)
19	Wash hands for 20 seconds (η_4)	0.754* (0.034)	0.755* (0.036)	0.751* (0.037)
20	Inform others if having symptoms of COVID (η_5)	0.552* (0.064)	0.567* (0.059)	0.558* (0.058)

*Significant at level 0.05

This estimated structural equation modelling indicates that health motivation (ξ_2) has a greater effect on the health behaviour (η) than perceived susceptibility (ξ_3). We can say here that health motivation is significantly correlated to health behaviour, which implies that people with high motivation to be healthy tend to have better health behaviour. It is also obtained that perceived susceptibility has no significant effect on health behaviour. It is possible to incorporate some other indicator variable into the model which is assumed as an indicator of perceived susceptibility. Meanwhile, the values of estimated measurement equations or the coefficient of factor loading and the associated standard errors for each indicator variable are provided in Table 4 (third column). We

can see from Table 4 that all indicator variables are significantly given effects on their respective latent variables. The goodness of fit of this proposed model is good enough, which is 62.9%, indicated by the value of R^2.
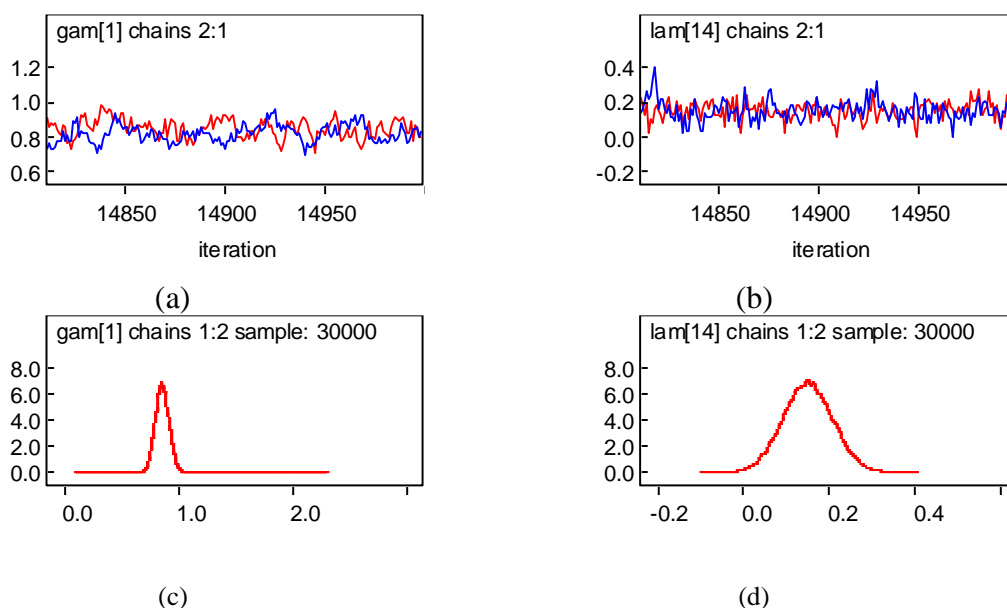
Figure 3. Selected trace plot for (a) $\hat{\gamma}_2$, (b) $\hat{\lambda}_{34}$, selected density plot for (c) $\hat{\gamma}_2$, (d) $\hat{\lambda}_{34}$.
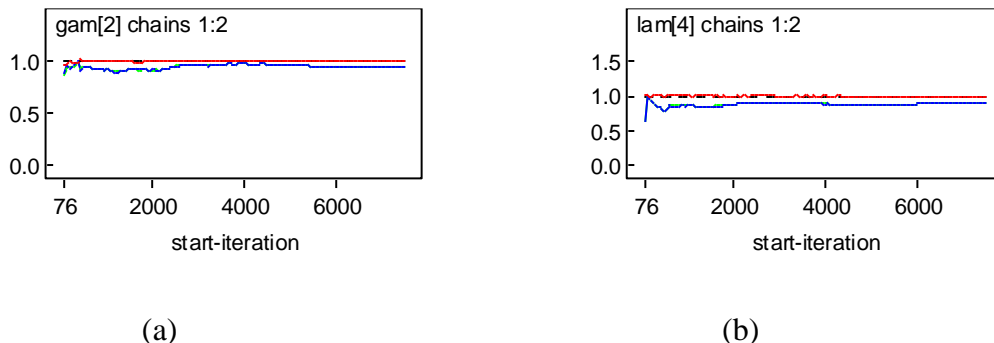
Figure 4. BGR plot for (a) $\hat{\gamma}_3$, (b) $\hat{\lambda}_{24}$,

4.   Conclusions

Our enthusiasm for SEM is based on its ability to enhance our understanding of health behaviour modelling, as both latent and observed variables are considered. Through the use of SEM, one can understand better how direct and indirect factors affect the health behaviour model of the population in West Sumatra during the COVID-19 outbreak. SEM method is quite flexible as it allows one to use data that is violated the normality assumption. In this study, the commonly used approach of maximum likelihood or robust weighted least square (RWLS) is not applied. Yanuar et al. (2013) have proved that Bayesian SEM results in a better model than classical SEM. Therefore, the Bayesian method is used for parameter estimation here. The estimated posterior mean obtained based on Bayesian SEM are acceptable since we believe that the power of our Bayesian SEM could result in the best fit for the proposed model. Besides, Bayesian SEM is found to be pertinent to be used for constructing the health behaviour of an individual living in West Sumatra. This study found that the significant factor that affects health behaviour is health motivation.

<p style="text-align:center">Table 4. The Results of the Bootstrap Simulation Study</p>

No	Parameter	Estimated Posterior Mean (SE)	Bootstrap Simulation Study	
			Estimated Bootstrap Mean (SE)	95% Bootstrap Confidence interval
A	**Structural Equation**			
	Health Motivation → Health Behaviour (γ_2)	0.851* (0.062)	0.842* (0.066)	(0.713 ; 0.971)
	Perceived susceptibility → Health Behaviour (γ_3)	0.086 (0.135)	0.082 (0.122)	(-0.157 ; 0.321)
B	**Measurement equation**			
	Health Motivation (ξ_2)			
1	Avoid touching the face (λ_{21})	0.529* (0.032)	0.535* (0.044)	(0.449 ; 0.621)
2	Avoid shaking hands (λ_{22})	0.457* (0.064)	0.443* (0.066)	(0.314 ; 0.572)
3	Avoid meeting or standing in long queues (λ_{23})	0.460* (0.065)	0.457* (0.067)	(0.326 ; 0.588)
4	Avoid touching objects in public areas (λ_{24})	0.504* (0.063)	0.495* (0.066)	(0.366 ; 0.624)
5	Avoid taking public transportation (online) (λ_{25})	0.534* (0.076)	0.536* (0.065)	(0.409 ; 0.663)
6	Avoid going home (λ_{26})	0.512* (0.074)	0.494* (0.077)	(0.343 ; 0.645)
7	Avoid worshiping in mosque/church/others (λ_{27})	0.571* (0.073)	0.568* (0.066)	(0.439 ; 0.697)
8	Avoid ordering online food (λ_{28})	0.420* (0.062)	0.387* (0.066)	(0.258 ; 0.516)
	Perceived susceptibility (ξ_3)			
9	Personal health (λ_{31})	0.546* (0.055)	0.548* (0.059)	(0.432 ; 0.664)
10	The health of their family members (λ_{32})	0.532* (0.057)	0.526* (0.061)	(0.406 ; 0.646)
11	When they go outside (λ_{33})	0.425* (0.058)	0.433* (0.049)	(0.337 ; 0.529)
12	When they are going to their village (λ_{34})	0.449* (0.058)	0.431* (0.060)	(0.313 ; 0.549)
13	To work/school, public facilities / public transportation (λ_{35})	0.431* (0.058)	0.437* (0.056)	(0.327 ; 0.529)
14	Food availability (λ_{36})	0.554* (0.060)	0.545* (0.066)	(0.416 ; 0.674)
15	COVID-19 spread(λ_{37})	0.444* (0.058)	0.439* (0.057)	(0.327 ; 0.551)
	Health Behaviour (η)			
16	Keep a distance of 2 m (η_1)	0.665* (0.043)	0.642* (0.054)	(0.536 ; 0.748)
17	Wear a mask(η_2)	0.564* (0.058)	0.533* (0.061)	(0.413 ; 0.653)
18	Hand sanitizer (η_3)	0.776* (0.032)	0.762* (0.034)	(0.695 ; 0.829)
19	Wash hands for 20 seconds (η_4)	0.754* (0.034)	0.751* (0.039)	(0.675 ; 0.827)
20	Inform others if having symptoms of COVID (η_5)	0.552* (0.064)	0.563* (0.061)	(0.443 ; 0.683)

*Significant at level 0.05

The proposed model obtained in this study provides the health behaviour to combat COVID-19, such as keeping physical distance, wearing a mask, washing hands with soap and running water or using hand sanitizer, avoiding crowded places, and limiting the mobilization and interaction. One other important thing to protect yourself from COVID-19 is always keeping the motivation to be healthy. In practical terms, this model is useful for an epidemiologist or a health manager. It will be easier for the epidemiologist or a health manager to control public health, reduce the rate of COVID-19's spreading effectively, and streamline government programs in the health sector.

Reference

1. Arora, T., & Grey, I. (2020). Health behaviour changes during COVID-19 and the potential consequences: A mini-review. *Journal of Health Psychology*, *25*(9), 1155–1163. https://doi.org/10.1177/1359105320937053

2. Asparouhov, T., & Muthén, B. (2021). Advances in Bayesian Model Fit Evaluation for Structural Equation Models. *Structural Equation Modelling: A Multidisciplinary Journal*, *28*(1), 1–14. https://doi.org/10.1080/10705511.2020.1764360

3. Cain, M. K., & Zhang, Z. (2019). Fit for a Bayesian: An Evaluation of PPP and DIC for Structural Equation Modelling. *Structural Equation Modelling: A Multidisciplinary Journal*, *26*(1), 39–50. https://doi.org/10.1080/10705511.2018.1490648

4. Choompunuch, B., Suksatan, W., Sonsroem, J., Kutawan, S., & In-udom, A. (2021). Stress, adversity quotient, and health behaviours of undergraduate students in a Thai university during COVID-19 outbreak. *Belitung Nursing Journal*, *7*(1), 1–7. https://doi.org/10.33546/bnj.1276

5. Conner, M., & Norman, P. (Eds.). (2007). *Predicting health behaviour: Research and practice with social cognition models* (2. ed., repr). Open Univ. Press.

6. Garnier-Villarreal, M., & Jorgensen, T. D. (2020). Adapting fit indices for Bayesian structural equation modelling: Comparison to maximum likelihood. *Psychological Methods*, *25*(1), 46–70. https://doi.org/10.1037/met0000224

7. Glanz, K., Rimer, B. K., & Viswanath, K. (Eds.). (2008). *Health behaviour and health education: Theory, research, and practice* (4th ed). Jossey-Bass.

8. Hou, W., Zhang, W., Jin, R., Liang, L., Xu, B., & Hu, Z. (2020). Risk factors for disease progression in hospitalized patients with COVID-19: A retrospective cohort study. *Infectious Diseases*, *52*(7), 498–505. https://doi.org/10.1080/23744235.2020.1759817

9. Kelava, A., Nagengast, B., & Brandt, H. (2014). A Nonlinear Structural Equation Mixture Modelling Approach for Nonnormally Distributed Latent Predictor Variables. *Structural Equation Modelling: A Multidisciplinary Journal*, *21*(3), 468–481. https://doi.org/10.1080/10705511.2014.915379

10. Khoso, P. A., Yew, V. W. C., & Mutalib, M. H. A. (2016). *Comparing and Contrasting Health Behaviour With Illness Behaviour. 11*(2), 12.

11. Lee, S.-Y. (2007). *Structural Equation Modelling: A Bayesian Approach*. John Wiley & Sons, Ltd.

12. Muharisa, C., Yanuar, F., & Devianto, D. (2018). Simulation Study of the Using of Bayesian Quantile Regression in Non- normal Error. *Cauchy - Jurnal Matematika Murni Dan Aplikasi*, *5*(November), 121–126.

13. Nam, E. J., Lee, E. K., & Oh, M.-S. (2018). Bayesian quantile regression analysis of Korean Jeonse deposit. *Communications for Statistical Applications and Methods*, *25*(5), 489–499. https://doi.org/10.29220/CSAM.2018.25.5.489

14. Ntzoufras, I. (2009). *Bayesian modelling using WinBUGS*. Wiley.

15. Olsson, U. H., Foss, T., Troye, S. V., & Howell, R. D. (2000). The Performance of ML, GLS, and WLS Estimation in Structural Equation Modelling Under Conditions of Misspecification and Nonnormality. *Structural Equation Modelling: A Multidisciplinary Journal*, *7*(4), 557–595. https://doi.org/10.1207/S15328007SEM0704_3

16. Parekh, N., & Deierlein, A. L. (2020). Health behaviours during the coronavirus disease 2019 pandemic: Implications for obesity. *Public Health Nutrition*, *23*(17), 3121–3125. https://doi.org/10.1017/S1368980020003031

17. Perlman, S. (2020). Another Decade, Another Coronavirus. *The New England Journal of Medicine*, 3.

18. Rahmadita, A., Yanuar, F., & Devianto, D. (2018). The Construction of Patient Loyalty Model Using Bayesian Structural Equation Modelling Approach. *CAUCHY*, *5*(2), 73. https://doi.org/10.18860/ca.v5i2.5039

19. Suksatan, W., Choompunuch, B., Koontalay, A., Posai, V., & Abusafia, A. H. (2021). Predictors of Health Behaviours Among Undergraduate Students During the COVID-19 Pandemic: A Cross-Sectional Predictive Study. *Journal of Multidisciplinary Healthcare*, *Volume 14*, 727–734. https://doi.org/10.2147/JMDH.S306718

20. Worldometers. (2021). *COVID-19 Coronavirus pandemic*. Https://Www.Worldometers.Info/Coronavirus/.

21. Y. Thanoon, T., Adnan, R., & Saffari, S. E. (2016). Bayesian Analysis of Linear and Nonlinear Latent Variable Models with Fixed Covariate and Ordered Categorical Data. *Pakistan Journal of Statistics and Operation Research*, *12*(1), 125–140. https://doi.org/10.18187/pjsor.v12i1.952

22. Yanuar, F. (2016). The Health Status Model in Urban and Rural Society in West Sumatera, Indonesia: An Approach of Structural Equation Modelling. *Indian Journal of Science and Technology*, *9*(4). https://doi.org/10.17485/ijst/2016/v9i4/72601

23. Yanuar, F., Ibrahim, K., & Aziz Jemain, A. (2013). Bayesian structural equation modelling for the health index. *Journal of Applied Statistics*, *40*(6), 1254–1269.

24. Yanuar, F., Ibrahim, K., & Jemain, A. A. (2010). On the application of structural equation modelling for the construction of a health index. *Environmental Health and Preventive Medicine*, *15*(5), 285–291. https://doi.org/10.1007/s12199-010-0140-7

25. Yanuar, F., Yozza, H., & Zetra, A. (2022). Modified Quantile Regression for Modelling the Low Birth Weight. *Frontiers in Applied Mathematics and Statistics*, *8*, 890028. https://doi.org/10.3389/fams.2022.890028

26. Zhang, X., & Savalei, V. (2016). Bootstrapping Confidence Intervals for Fit Indexes in Structural Equation Modelling. *Structural Equation Modelling: A Multidisciplinary Journal*, *23*(3), 392–408. https://doi.org/10.1080/10705511.2015.1118692

27. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., & Tan, W. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, *382*(8), 727–733. https://doi.org/10.1056/NEJMoa2001017

28. Zvolensky, M. J., Garey, L., Rogers, A. H., Schmidt, N. B., Vujanovic, A. A., Storch, E. A., Buckner, J. D., Paulus, D. J., Alfano, C., Smits, J. A. J., & O'Cleirigh, C. (2020). Psychological, addictive, and health behaviour implications of the COVID-19 pandemic. *Behaviour Research and Therapy*, *134*, 1–16. https://doi.org/10.1016/j.brat.2020.103715

Appendix

Code in Winbugs

model{

```
for(i in 1:N){
# Measurement equation model
for(j in 1:P){
y[i,j]~dnorm(mu[i,j],psi[j])I(thd[j,z[i,j]],thd[j,z[i,j]+1])
ephat[i,j]<-y[i,j]-mu[i,j]
}

mu[i,1]<-xi[i,1]
mu[i,2]<-lam[1]*xi[i,1]
mu[i,3]<-lam[2]*xi[i,1]
mu[i,4]<-lam[3]*xi[i,1]
mu[i,5]<-lam[4]*xi[i,1]
mu[i,6]<-lam[5]*xi[i,1]
mu[i,7]<-lam[6]*xi[i,1]
mu[i,8]<-lam[7]*xi[i,1]
mu[i,9]<-xi[i,2]
mu[i,10]<-lam[8]*xi[i,2]
mu[i,11]<-lam[9]*xi[i,2]
mu[i,12]<-lam[10]*xi[i,2]
mu[i,13]<-lam[11]*xi[i,2]
mu[i,14]<-lam[12]*xi[i,2]
mu[i,15]<-lam[13]*xi[i,2]
mu[i,16]<-eta[i]
mu[i,17]<-lam[14]*eta[i]
mu[i,18]<-lam[15]*eta[i]
mu[i,19]<-lam[16]*eta[i]
mu[i,20]<-lam[17]*eta[i]


# Structural equation model
xi[i,1:2]~dmnorm(u[1:2],phi[1:2,1:2])
eta[i]~dnorm(nu[i],psd)
nu[i]<-gam[1]*x[i]+gam[2]*xi[i,1]+gam[3]*xi[i,2]
dthat[i]<-eta[i]-nu[i]
}# end of i

for(i in 1:2){u[i]<-0.0}
# Priors on loadings and coefficients
var.lam[1]<-10.0*psi[2]
var.lam[2]<-10.0*psi[3]
var.lam[3]<-10.0*psi[4]
var.lam[4]<-10.0*psi[5]
var.lam[5]<-10.0*psi[6]
var.lam[6]<-10.0*psi[7]
var.lam[7]<-10.0*psi[8]
var.lam[8]<-10.0*psi[10]
var.lam[9]<-10.0*psi[11]
var.lam[10]<-10.0*psi[12]
var.lam[11]<-10.0*psi[13]
var.lam[12]<-10.0*psi[14]
var.lam[13]<-10.0*psi[15]
var.lam[14]<-10.0*psi[17]
var.lam[15]<-10.0*psi[18]
var.lam[16]<-10.0*psi[19]
var.lam[17]<-10.0*psi[20]

for(i in 1:17){lam[i]~dnorm(0.5,var.lam[i])}
var.gam<-10.0*psd
gam[1]~dnorm(0.8,var.gam)
gam[2]~dnorm(0.2,var.gam)
```

gam[3]~dnorm(0.2,var.gam)

#priors on precisions
for(j in 1:P){
psi[j]~dgamma(10,8)
sgm[j]<-1/psi[j]
}
psd~dgamma(10,8)
sgd<-1/psd
phi[1:2,1:2]~dwish(R[1:2,1:2], 30)
phx[1:2,1:2]<-inverse(phi[1:2,1:2])
} #end of model

 # L O A D   D A T A
list(N=756, P=20,
R=structure(
     .Data=c(8.0, 0.0,
             0.0, 8.0),
     .Dim=c(2,2)),
thd=structure(
 .Data=c(-200.000, -1.748, -1.090, -0.082, 0.779, 200.000,
             ……
             -200.000, -1.667, -1.329, -0.810, 0.009, 200.000),
         .Dim=c(20,6)),
z=structure(
.Data=c(……..),
.Dim=c(756, 20))) ,

 # Two different Initial Values
list(
    lam=c(0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0,0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0),
    psi=c(0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5),
        psd=0.6,
        gam=c(0.0, 0.0, 0.0),
        phi=structure(
        .Data=c(0.5, 0.0,
                 0.0, 0.5),
        .Dim=c(2,2)),
xi=structure(
.Data=c(…….),
.Dim=c(756,2)))

list(
        lam=c(0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5),
        psi=c(1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0),
        psd=1.0,
        gam=c(1.0, 1.0, 1.0),
        phi=structure(
        .Data=c(1.0, 0.0,
                0.0, 1.0),
        .Dim=c(2,2)),
xi=structure(
.Data=c(…….……),
.Dim=c(756,2)))