Econometrics and Statistics 27 (2023) 16–35

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

A Multivariate Randomized Response Model for Sensitive Binary Data

Amanda M.Y. Chu[a], Yasuhiro Omori[b,*], Hing-yu So[c], Mike K.P. So[d]

[a] *Department of Social Sciences, The Education University of Hong Kong, Hong Kong*
[b] *Faculty of Economics, The University of Tokyo, Japan*
[c] *Division of Quality & Safety, New Territories East Cluster, Hong Kong Hospital Authority, Hong Kong*
[d] *Department of Information Systems, Business Statistics and Operations Management, The Hong Kong University of Science and Technology, Hong Kong*

ARTICLE INFO

Article history:
Received 22 December 2020
Revised 5 January 2022
Accepted 18 January 2022
Available online 26 January 2022

Keywords:
Bayesian analysis
Data privacy
Multivariate probit models
Patient safety
Sensitive questions

ABSTRACT

A new statistical method is proposed to combine the randomized response technique, probit modeling, and Bayesian analysis to analyze large-scale online surveys of multiple binary randomized responses. The proposed method is illustrated by analyzing sensitive dichotomous randomized responses on different types of drug administration error from nurses in a hospital cluster. A statistical challenge is that nurses' true sensitive responses are unobservable because of a randomization scheme that protects their data privacy to answer the sensitive questions. Four main contributions of the paper are highlighted. The first is the construction of a generic statistical approach in modeling multivariate sensitive binary data collected from the randomized response technique. The second is studying the dependence of multivariate sensitive responses via statistical measures. The third is the calculation of an overall attitude score using sensitive responses. The last one is an illustration of the proposed statistical method for analyzing administration policies that potentially involve sensitive topics which are important to study but are not easily investigated via empirical studies. The particular healthcare example on drug administration policies demonstrated in this paper also presents a scientific way to elicit managerial strategies while protecting data privacy through analytics.

1. Introduction

It is very common in survey and behavioral research to collect binary responses. The probability of observing a particular binary response can be related to other information or independent variables by using regression models (e.g., logit regression or probit regression). In particular, the purpose of a probit regression is to link the occurrence of a response with latent variables, which can be interpreted as random utilities (Manski, 1977; McFadden, 1980). When there are multiple binary responses in a survey, multivariate probit regressions can be set up by using multiple latent variables. By studying the correlation of latent variables, we can understand the relationship between binary responses (e.g., the co-occurrence of two binary outcomes). In the literature, all binary response variables of interest are assumed to be observable. In this paper, we develop a Bayesian approach to analyze binary responses that cannot be directly observed, or that are 'masked' by other

[*] Corresponding author. Tel.: +81 3 5841 5516; fax: +81 3 5841 5521.
E-mail address: omori@e.u-tokyo.ac.jp (Y. Omori).

https://doi.org/10.1016/j.ecosta.2022.01.003
2452-3062/© 2022 EcoSta Econometrics and Statistics. Published by Elsevier B.V. All rights reserved.

data. The motivation of this study is to circumvent the need to ask respondents sensitive questions directly when we need to collect sensitive responses. A common statistical method to elicit trustful responses is the randomized response technique (RRT) pioneered by Warner (1965), Greenberg et al. (1969) and Greenberg et al. (1971). See also Blair et al. (2015) for a recent survey on RRT methods.

A core idea of the RRT is to let respondents answer questions according to a randomization scheme, meaning researchers have no way of knowing respondents' sensitive responses exactly. The unrelated question design (UQD) of Greenberg et al. (1969) illustrates this, where a sensitive question, denoted by x, is paired with an innocuous question y, and then z, a random mixture of x and y is obtained: $z = dx + (1-d)y$, where d is a Bernoulli random variable with a preset probability of success. For example,

x : I had the experience in giving medications to a wrong patient and did not report it to the hospital.

y : I had breakfast at home at least 3 times in the past week.

The question x above is sensitive, and its response is mixed with the response of y randomly in an RRT survey. Because x is not directly observed, the use of the RRT can help to protect respondents' data privacy and enhance data confidentiality. In this paper, we focus on a multivariate probit regression model where x is unobservable but its related information is obtained via z. We develop posterior inference methods for probit regression parameters using z and illustrate how to estimate the relationship of binary sensitive responses for greater understanding of drug administration policies in a hospital cluster in Hong Kong.

1.1. Probit regression

The use of probit regressions has a long history. Probit regression has been popular for analyzing choice models in business research (Doyle, 1977), for analyzing behavioral responses in social sciences studies (Muthén, 1979), and in economic research (Amemiya, 1981). A useful approach for the statistical inference of binary response is to conduct a Bayesian analysis using Markov chain Monte Carlo (MCMC) methods (e.g. Albert and Chib (1993), Holmes and Held (2006), Frühwirth-Schnatter and Frühwirth (2007), Roy and Hobert (2007), Berrett and Calder (2012), Polson et al. (2013), Qin and Hobert (2019), while Durante (2019) and Fasano et al. (2021) use i.i.d. Monte Carlo samples). By augmenting suitable latent variables in the analysis, Albert and Chib (1993) demonstrated that their methods can be generalized to handle multinomial responses. Chib and Greenberg (1998) considered further Bayesian inferences of multivariate probit models, and analyzing the dependence between multiple binary responses became possible. Gibbons and Wilcox-Gök (1998) applied a multivariate probit model to healthcare data where estimations were done using an EM algorithm. O'brien and Dunson (2004) studied a multivariate logistic regression model by using a 'logistic distribution'. Song and Lee (2005) developed a multivariate probit model for binary responses. Lee et al. (2010) worked on nonlinear structural equation models of binary variables with probit and logit links. Imai and Van Dyk (2005) studied the multinomial probit model with marginal data augmentation. Talhouk et al. (2012) proposed Bayesian inference methods of multivariate probit models with a decomposable graphical model for the inverse of the correlation matrix of the latent variables. Laffont et al. (2014) and Barcella et al. (2018) studied multivariate binary responses over time using a probit regression. Our contribution is to develop a full Bayesian scheme for analyzing multiple RRT binary responses. A statistical challenge is that true sensitive responses are unobservable because of a randomization scheme that protects respondents' data privacy. We extend the multivariate probit regression framework of Chib and Greenberg (1998) to study dependence structures of RRT binary responses, and in particular, to examine the relationships between demographic and organizational characteristics, and drug administration error. To the best of our knowledge, this paper is the first to integrate multivariate RRT responses in a probit modeling framework. Although we do not directly observe the sensitive responses, our Bayesian multivariate RRT modeling enables the dependence analysis of sensitive responses. From this, we can gain insights into how to extract opinions from sensitive topics, or in our example, how to improve drug administration policies.

The efficient MCMC estimation method is proposed to estimate model parameters for the multivariate randomized response model for binary data. This is important as it is often the case that sampling the correlation matrix results in highly autocorrelated MCMC samples (e.g., as in Liu and Daniels (2006) and Zhang et al. (2006)). Our estimation method is based on the idea of data augmentation. In order to define the prior distribution of the correlation matrix, we first consider the inverse Wishart distribution for the covariance matrix. We then decompose this into the priors for the variances and the correlation matrix as in Zhang et al. (2006) where the marginally uniform prior density for the correlation matrix is obtained as a special case in Barnard et al. (2000). Then we use the components of the standard deviations to multiply model parameters and sample the transformed parameters using data augmentation as in Talhouk et al. (2012). Through the inverse transformations, we obtain MCMC samples of the model parameters. Furthermore, instead of sampling the correlation matrix, we can construct a factor model for the multivariate probit regression with randomized responses. We also describe the MCMC algorithm, which is shown to capture the factor structure of the correlation matrix in the analysis of drug administration survey data.

1.2. A drug administration survey

It is not surprising that staff's failure to follow working procedures or guidelines poses a serious threat to organizations. To further understand the issue, organizations may conduct a survey study among their staff. However, response distortions may occur, especially if respondents are unwilling to provide honest answers or if their responses may incur legal liability (Locander et al., 1976). The use of the RRT (Warner, 1965) with the UQD (Greenberg et al., 1969) has been recommended as a method to tackle response distortion in research involving sensitive topics (Kwan et al., 2010; Chu et al., 2018; Chung et al., 2018; Chong et al., 2019). A key challenge is that existing RRT modeling methods cannot handle multivariate randomized responses for binary data. In other words, we cannot understand the multivariate dependence in binary sensitive questions collected through the RRT. We present a drug administration survey example motivating this paper and pose four research questions. One of the hospital clusters in Hong Kong manages several public hospitals and serves approximately one-fifth of the population in Hong Kong. Each hospital includes nurses in different ranks and with varying levels of nursing experience. A number of practices and guidelines in the drug administration process are implemented in the hospital cluster to enhance patient safety and reduce drug administration errors. The top management of the hospital cluster would like to learn more about whether nurses have ever made drug administration errors resulting from a lack of adherence to the working practices or guidelines. They would also like to understand the relationships among the errors. Thus, four research questions were raised: 1) How do we construct a generic approach in modeling multivariate sensitive binary data collected from the randomized response technique? 2) Is there any dependence between the different types of drug administration error? 3) How can we calculate an overall attitude score of nurses using sensitive responses? 4) For understanding drug administration policies, do demographic and organizational characteristics explain the drug administration errors? The investigation of drug administration errors, preventive measures related to drug administration errors, and nurses' attitudes toward the drug administration system form an important research agenda in the healthcare sector (Llewellyn et al., 2009; Sheu et al., 2009; Keers et al., 2013; Kim and Bates, 2013; Cross et al., 2017). However, little research has been conducted in these areas because sensitive questions are involved, and thus nurses are less likely to provide truthful responses if questions are asked directly. Therefore, an online drug administration survey that applied the RRT with the UQD was conducted. Both the online and RRT methodologies can encourage respondents to provide truthful responses to sensitive questions (Burkill et al., 2016; Chu et al., 2018). As eight dichotomous sensitive questions using the RRT with the UQD were asked, we had to manage multivariate dependence problems on binary variables with incomplete information resulting from the randomized procedure in the survey design. To the best of our knowledge, this is the largest online RRT survey (in terms of the number of sensitive questions and respondents) in the area of healthcare to investigate issues related to patient safety. We are the first to perform multivariate RRT modeling on such large-scale sensitive healthcare data.

The remainder of this article is structured as follows. In Section 2, we show how a multivariate probit model can be built to manage multivariate randomized responses for binary data. In Section 3, we demonstrate the application of the proposed model in a real dataset concerning drug administration. We also carry out a dependence analysis of the sensitive binary responses in this section. Section 4 describes an alternative factor analytic approach for the prior on the latent correlation matrix in multivariate Probit modeling and summarizes our main findings.

2. The multivariate probit model for multiple RRT data

2.1. The RRT setting

Suppose that there are responses of K pairs of binary questions x_{ik} and y_{ik} from n individuals, $i = 1, ..., n$, $k = 1, ..., K$. Suppose also that the two binary random variables, x_{ik} and y_{ik}, denote the k-th responses to sensitive and innocuous questions for the i-th individual. Then, we can observe the following randomized response,

$$z_{ik} = d_{ik}x_{ik} + (1 - d_{ik})y_{ik}, \quad i = 1, \dots, n, \quad k = 1, \dots, K, \tag{1}$$

where $d_{ik} \sim i.i.d.\ Bernoulli(\pi_i)$ for each i, and $\pi_i's$ $(0 < \pi_i < 1)$ are assumed to be some known constants, and

$$Pr(z_{ik} = j) = \pi_i Pr(x_{ik} = j) + (1 - \pi_i)Pr(y_{ik} = j), \quad j = 0, 1, \quad k = 1, \dots, K. \tag{2}$$

In the construction above, π_i is the probability of the i-th individual answering a sensitive question. Following the UQD of Greenberg et al. (1969), Kwan et al. (2010), Chung et al. (2018), and Chu et al. (2020), we split the sample into two groups according to the different probabilities, p_1 and p_2, of answering sensitive questions. Under this UQD, we may assume for simplicity that

$$\pi_i = \begin{cases} p_1, & i = 1, \dots, n_1, \\ p_2, & i = n_1 + 1, \dots, n. \end{cases}$$

We define $n \times K$ matrices, $\mathbf{Z} = (\boldsymbol{z}_1, \dots, \boldsymbol{z}_n)'$, $\mathbf{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)'$, $\mathbf{Y} = (\boldsymbol{y}_1, \dots, \boldsymbol{y}_n)'$ and $\mathbf{D} = (\boldsymbol{d}_1, \dots, \boldsymbol{d}_n)'$ where

$$\boldsymbol{z}_i = \begin{pmatrix} z_{i1} \\ z_{i2} \\ \vdots \\ z_{iK} \end{pmatrix}, \quad \boldsymbol{x}_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iK} \end{pmatrix}, \quad \boldsymbol{y}_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iK} \end{pmatrix}, \quad \boldsymbol{d}_i = \begin{pmatrix} d_{i1} \\ d_{i2} \\ \vdots \\ d_{iK} \end{pmatrix}.$$

In a typical RRT design, we cannot observe $\boldsymbol{x}_i$, $\boldsymbol{y}_i$, or $\boldsymbol{d}_i$ in order to protect the privacy of the individuals. In this paper, we develop Bayesian methods for inferring $\boldsymbol{x}_i$ from the observed responses $\boldsymbol{z}_i$ under multivariate probit modeling. One primary objective is to study the relationships of sensitive binary responses $\boldsymbol{x}_i$ under the RRT setting.

2.2. A probit model with randomized responses

To estimate the marginal probabilities $Pr(x_{ik}=1)$, $Pr(y_{ik}=1)$ and the joint probability $Pr(x_{ik}=1, x_{il}=1)$ of assessing the degree of dependence of two binary responses, we incorporate latent continuous random variables, x^*_{ik} and y^*_{ik} so that

$$Pr(x_{ik}=1|\boldsymbol{v}_i) = Pr(x^*_{ik}>0|\boldsymbol{v}_i), \quad Pr(y_{ik}=1|\boldsymbol{v}_i) = Pr(y^*_{ik}>0|\boldsymbol{v}_i), \tag{3}$$

where $\boldsymbol{v}_i$ is a $p\times 1$ covariate vector with the first element equal to one. Since x_{ik} and y_{ik} are responses to the sensitive question and the innocuous question (which is supposed to be unrelated to the sensitive question) respectively, it is natural to assume that x^*_{ik} and y^*_{ik} are independent. Thus, we consider the multivariate probit model (see Chib and Greenberg (1998)) given by

$$x^*_{ik}|\boldsymbol{v}_i, \boldsymbol{\beta}_k \sim \mathcal{N}(\boldsymbol{v}_i'\boldsymbol{\beta}_k, 1), \quad y^*_{ik}|\boldsymbol{v}_i, \boldsymbol{\alpha}_k \sim \mathcal{N}(\boldsymbol{v}_i'\boldsymbol{\alpha}_k, 1), \tag{4}$$

and

$$\begin{aligned} \boldsymbol{x}^*_i|\boldsymbol{v}_i, \mathbf{B}, \mathbf{R} &\sim \mathcal{N}(\mathbf{B}'\boldsymbol{v}_i, \mathbf{R}), \quad \mathbf{B} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_K), \\ \boldsymbol{y}^*_i|\boldsymbol{v}_i, \mathbf{A} &\sim \mathcal{N}(\mathbf{A}'\boldsymbol{v}_i, \mathbf{I}_K), \quad \mathbf{A} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_K), \end{aligned} \tag{5}$$

where $\boldsymbol{x}^*_i, \boldsymbol{y}^*_i$ are conditionally independent, $\mathbf{B}$, $\mathbf{A}$ are $p\times K$ matrices, $\boldsymbol{\beta}_k$, $\boldsymbol{\alpha}_k$ are $p\times 1$ coefficient vectors, $\mathbf{R}$ is a correlation matrix and, I_K denotes $K\times K$ identity matrix,

$$\boldsymbol{x}^*_i = \begin{pmatrix} x^*_{i1} \\ x^*_{i2} \\ \vdots \\ x^*_{iK} \end{pmatrix}, \quad \boldsymbol{y}^*_i = \begin{pmatrix} y^*_{i1} \\ y^*_{i2} \\ \vdots \\ y^*_{iK} \end{pmatrix}.$$

Define $n\times K$ matrices, $\mathbf{X}^* = (\boldsymbol{x}^*_1, \ldots, \boldsymbol{x}^*_n)'$, $\mathbf{Y}^* = (\boldsymbol{y}^*_1, \ldots, \boldsymbol{y}^*_n)'$ and a $n\times p$ matrix, $\mathbf{V} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)'$. The probit regression formulation in (5) enables us to compute $q^{\boldsymbol{v}}_k = Pr(x_{ik}=1)$ and $q^{\boldsymbol{v}}_{kl} = Pr(x_{ik}=1, x_{il}=1)$ via multivariate normal distributions. One challenge is that both x_{ik} and y_{ik} are unobservable under the RRT setting. We develop a new Bayesian analysis scheme to estimate the probit model in (5) from which we can conduct the posterior inference of $q^{\boldsymbol{v}}_k$ and $q^{\boldsymbol{v}}_{kl}$ for $k, l = 1, ..., K$.

2.3. Bayesian analysis

Let $\mathrm{vec}(\mathbf{A})$ denote a vectorization of $\mathbf{A}$. Noting that $nK\times 1$ vectors, $\mathrm{vec}\big(\mathbf{X}^{*\prime}\big)$ and $\mathrm{vec}\big(\mathbf{Y}^{*\prime}\big)$, follow the multivariate normal distributions independently,

$$\mathrm{vec}\big(\mathbf{X}^{*\prime}\big)|\mathbf{V}, \mathbf{B}, \mathbf{R} \sim \mathcal{N}(\mathrm{vec}(\mathbf{B}'\mathbf{V}'), \mathbf{I}_n \otimes \mathbf{R}), \quad \mathrm{vec}\big(\mathbf{Y}^{*\prime}\big)|\mathbf{V}, \mathbf{A} \sim \mathcal{N}(\mathrm{vec}(\mathbf{A}'\mathbf{V}'), \mathbf{I}_n \otimes \mathbf{I}_K), \tag{6}$$

where $\mathbf{I}_n$ denotes $n\times n$ identity matrix and $\otimes$ denotes the Kronecker product, we define a matrix variate normal distribution as follows.

Definition 2.1. The random matrix $\mathbf{X}$ $(p\times q)$ is said to have a matrix variate normal distribution with mean matrix $\mathbf{M}$ $(p\times q)$ and covariance matrix $\boldsymbol{\Psi}\otimes\boldsymbol{\Sigma}$ where $\boldsymbol{\Psi}$ $(p\times p)$ and $\boldsymbol{\Sigma}$ $(q\times q)$ are positive definite matrices if $pq\times 1$ vector $\mathrm{vec}(\mathbf{X}') \sim \mathcal{N}(\mathrm{vec}(\mathbf{M}'), \boldsymbol{\Psi}\otimes\boldsymbol{\Sigma})$.

We denote such a matrix as $\mathbf{X} \sim \mathcal{N}_{p,q}(\mathbf{M}, \boldsymbol{\Psi}\otimes\boldsymbol{\Sigma})$.

If $\mathbf{X} \sim \mathcal{N}_{p,q}(\mathbf{M}, \boldsymbol{\Psi}\otimes\boldsymbol{\Sigma})$, the probability density function for $\mathbf{X}$ is given by

$$f(\mathbf{X}) = (2\pi)^{-pq/2}|\boldsymbol{\Psi}|^{-q/2}|\boldsymbol{\Sigma}|^{-p/2} \times \exp\left\{-\frac{1}{2}\mathrm{tr}\Big(\boldsymbol{\Psi}^{-1}(\mathbf{X}-\mathbf{M})\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\mathbf{M})'\Big)\right\},$$

as

$$\begin{aligned} \mathrm{tr}\Big(\boldsymbol{\Psi}^{-1}(\mathbf{X}-\mathbf{M})\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\mathbf{M})'\Big) &= \mathrm{vec}\Big((\mathbf{X}-\mathbf{M})'\Big)'\Big(\boldsymbol{\Psi}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\Big)\mathrm{vec}\Big((\mathbf{X}-\mathbf{M})'\Big) \\ &= \Big\{\mathrm{vec}(\mathbf{X}') - \mathrm{vec}(\mathbf{M}')\Big\}'\Big(\boldsymbol{\Psi}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\Big)\Big\{\mathrm{vec}(\mathbf{X}') - \mathrm{vec}(\mathbf{M}')\Big\}. \end{aligned}$$

Using Definition 2.1, we rewrite (6) as

$$\mathbf{X}^*|\mathbf{V}, \mathbf{B}, \mathbf{R} \sim \mathcal{N}_{n,K}(\mathbf{VB}, \mathbf{I}_n\otimes\mathbf{R}), \quad \mathbf{Y}^*|\mathbf{V}, \mathbf{A} \sim \mathcal{N}_{n,K}(\mathbf{VA}, \mathbf{I}_n\otimes\mathbf{I}_K).$$

The conditional probability density of $(\mathbf{Z}, \mathbf{X}^*, \mathbf{Y}^*, \mathbf{D})$ given $\mathbf{B}, \mathbf{A}, \mathbf{R}$ is then given by

$$\begin{aligned} & f(\mathbf{Z}, \mathbf{X}^*, \mathbf{Y}^*, \mathbf{D}|\mathbf{B}, \mathbf{A}, \mathbf{R}) \\ & \quad \propto |\mathbf{R}|^{-n/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{(\mathbf{X}^* - \mathbf{VB})\mathbf{R}^{-1}(\mathbf{X}^* - \mathbf{VB})' + (\mathbf{Y}^* - \mathbf{VA})(\mathbf{Y}^* - \mathbf{VA})'\right\}\right] \\ & \times \prod_{i=1}^{n}\prod_{k=1}^{K} \pi_i^{d_{ik}}(1-\pi_i)^{1-d_{ik}}\left\{d_{ik}I(x_{ik}^* > 0) + (1-d_{ik})I(y_{ik}^* > 0)\right\}^{z_{ik}}\left\{d_{ik}I(x_{ik}^* \le 0) + (1-d_{ik})I(y_{ik}^* \le 0)\right\}^{1-z_{ik}}, \end{aligned}$$

where $I(A)$ is an indicator function ($I(A) = 1$ if A is true and 0 otherwise). For the prior distributions of $\mathbf{A}$, and $\mathbf{B}$ given $\mathbf{R}$, we assume

$$\mathbf{A} \sim \mathcal{N}_{p,K}(\mathbf{O}_{p,K}, \mathbf{\Phi}_0 \otimes \mathbf{I}_K), \quad \mathbf{B}|\mathbf{R} \sim \mathcal{N}_{p,K}(\mathbf{O}_{p,K}, \mathbf{\Psi}_0 \otimes \mathbf{R}),$$

where $\mathbf{O}_{p,K}$ denotes a $p \times K$ matrix with all elements equal to zero, $\mathbf{\Phi}_0$ and $\mathbf{\Psi}_0$ are $p \times p$ positive definite symmetric matrices. The conditionally conjugate prior distribution is assumed for $\mathbf{B}$ given $\mathbf{R}$. For a prior distribution of $\mathbf{R}$, we assume the prior probability density as

$$\pi(\mathbf{R}) \propto |\mathbf{R}|^{\frac{(n_0-1)(K-1)}{2}-1}\left(\prod_{k=1}^{K}|\mathbf{R}_{kk}|\right)^{-\frac{n_0}{2}}, \tag{7}$$

where n_0 is a hyper-parameter and $\mathbf{R}_{kk}$ denotes a principal submatrix of $\mathbf{R}$ (see Appendix A.1). If we take $n_0 = K + 1$, (7) reduces to the marginally uniform prior distribution in Barnard et al. (2000). Then the posterior probability density given $\mathbf{Z}$ is

$$\begin{aligned} & \pi(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{D}, \mathbf{B}, \mathbf{A}, \mathbf{R}|\mathbf{Z}) \\ & \quad \propto |\mathbf{R}|^{-n/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{(\mathbf{X}^* - \mathbf{VB})\mathbf{R}^{-1}(\mathbf{X}^* - \mathbf{VB})' + (\mathbf{Y}^* - \mathbf{VA})(\mathbf{Y}^* - \mathbf{VA})'\right\}\right] \\ & \quad \times |\mathbf{R}|^{-p/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{\mathbf{\Psi}_0^{-1}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}' + \mathbf{\Phi}_0^{-1}\mathbf{A}\mathbf{A}'\right\}\right] \times \pi(\mathbf{R}) \\ & \quad \times \prod_{i=1}^{n}\prod_{k=1}^{K} \pi_i^{d_{ik}}(1-\pi_i)^{1-d_{ik}}\left\{d_{ik}I(x_{ik}^* > 0) + (1-d_{ik})I(y_{ik}^* > 0)\right\}^{z_{ik}}\left\{d_{ik}I(x_{ik}^* \le 0) + (1-d_{ik})I(y_{ik}^* \le 0)\right\}^{1-z_{ik}}. \end{aligned} \tag{8}$$

In summary, our proposed model is given by

$$\begin{aligned} & \mathbf{Z} = \mathbf{D} \odot \mathbf{X} + (\mathbf{J} - \mathbf{D}) \odot \mathbf{Y}, \\ & \quad d_{ik} \sim i.i.d.\ Bernoulli(\pi_i), \quad k = 1, \ldots, K, \quad i = 1, \ldots, n, \\ & \quad x_{ik} = I(x_{ik}^* > 0), \quad y_{ik} = I(y_{ik}^* > 0), \\ & \quad \mathbf{X}^*|\mathbf{V}, \mathbf{B}, \mathbf{R} \sim \mathcal{N}_{n,K}(\mathbf{VB}, \mathbf{I}_n \otimes \mathbf{R}), \quad \mathbf{Y}^*|\mathbf{V}, \mathbf{A} \sim \mathcal{N}_{n,K}(\mathbf{VA}, \mathbf{I}_n \otimes \mathbf{I}_K), \\ & \quad \mathbf{A} \sim \mathcal{N}_{p,K}(\mathbf{O}_{p,K}, \mathbf{\Phi}_0 \otimes \mathbf{I}_K), \quad \mathbf{B}|\mathbf{R} \sim \mathcal{N}_{p,K}(\mathbf{O}_{p,K}, \mathbf{\Psi}_0 \otimes \mathbf{R}), \quad \mathbf{R} \sim \pi(\mathbf{R}), \end{aligned}$$

where $\odot$ denotes element-wise multiplication, and $\mathbf{J}$ denotes a $n \times K$ matrix with all elements equal to one. The matrices $\mathbf{Z}$ (the response matrix) and $\mathbf{V}$ (the covariate matrix) are observed, while $\mathbf{D}, \mathbf{X}, \mathbf{X}^*, \mathbf{Y}, \mathbf{Y}^*$ are latent variables. The $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{R}$ are parameters, and π_i, $\mathbf{\Phi}_0$, $\mathbf{\Psi}_0$ are the known constant and known matrices.

Remark 1. We further develop the alternative factor analytic characterization of $\mathbf{R}$ in Section 4.2.

2.4. Markov chain Monte Carlo implementation

For performing the Bayesian inference, the unknown variables $\mathbf{X}^*, \mathbf{Y}^*, \mathbf{D}, \mathbf{B}, \mathbf{A}$ and $\mathbf{R}$ are sampled using the MCMC method. Based on the posterior distribution in (8), we implement the MCMC algorithm in five blocks as follows.

1. Initialize $\mathbf{D}, \mathbf{X}^*, \mathbf{Y}^*, \mathbf{B}, \mathbf{A}$ and $\mathbf{R}$.
2. Generate $\mathbf{D}, \mathbf{X}^*, \mathbf{Y}^*|\mathbf{B}, \mathbf{A}, \mathbf{R}, \mathbf{Z}$.
   (a) Generate $\mathbf{D}|\mathbf{B}, \mathbf{A}, \mathbf{R}, \mathbf{Z}$.
   (b) Generate $\mathbf{X}^*, \mathbf{Y}^*|\mathbf{D}, \mathbf{B}, \mathbf{A}, \mathbf{R}, \mathbf{Z}$.
3. Generate $\mathbf{A}|\mathbf{D}, \mathbf{X}^*, \mathbf{Y}^*, \mathbf{R}, \mathbf{Z}$.
4. Generate $\mathbf{B}, \mathbf{R}|\mathbf{D}, \mathbf{X}^*, \mathbf{Y}^*, \mathbf{A}, \mathbf{Z}$.
5. Go to Step 2.

Step 2: Generation of $\mathbf{D}, \mathbf{X}^*$, and $\mathbf{Y}^*$

We derive the conditional posterior probability density function of $(d_{ik}, x_{ik}^*, y_{ik}^*)$, for $i = 1, \ldots, n$ and $k = 1, \ldots, K$. Let $\mathbf{B}_{-k} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_{k-1}, \boldsymbol{\beta}_{k+1}, \ldots, \boldsymbol{\beta}_K)$. Further, let $\boldsymbol{r}_{k,-k}$ and $\mathbf{R}_{-k}$ denote the row vector of correlations between x_{ik}^* and x_{ij}^* for $j =$

$1, \ldots, k-1, k+1, \ldots, K$, and the matrix that is obtained by excluding the k-th row and k-th column of $\mathbf{R}$. Then, we have the full conditional distribution of $(d_{ik}, x_{ik}^*, y_{ik}^*)$ given by

$$\begin{aligned}
&\pi(d_{ik}, x_{ik}^*, y_{ik}^* | \cdot) \\
&\quad \propto \exp\left\{ -\frac{1}{2\sigma_{x,ik}^2}(x_{ik}^* - \mu_{x,ik})^2 - \frac{1}{2\sigma_{y,ik}^2}(y_{ik}^* - \mu_{y,ik})^2 \right\} \\
&\quad \times \pi_i^{d_{ik}}(1-\pi_i)^{1-d_{ik}} \left\{ d_{ik} I(x_{ik}^* > 0) + (1-d_{ik}) I(y_{ik}^* > 0) \right\}^{z_{ik}} \left\{ d_{ik} I(x_{ik}^* \le 0) + (1-d_{ik}) I(y_{ik}^* \le 0) \right\}^{1-z_{ik}}. \qquad (9)
\end{aligned}$$

where

$$\mu_{x,ik} = \boldsymbol{v}_i' \boldsymbol{\beta}_k + \boldsymbol{r}_{k,-k} \mathbf{R}_{-k}^{-1} (\boldsymbol{x}_{i,-k}^* - \mathbf{B}_{-k}' \boldsymbol{v}_i), \quad \sigma_{x,ik}^2 = 1 - \boldsymbol{r}_{k,-k} \mathbf{R}_{-k}^{-1} \boldsymbol{r}_{k,-k}',$$
$$\mu_{y,ik} = \boldsymbol{v}_i' \boldsymbol{\alpha}_k, \quad \sigma_{y,ik}^2 = 1,$$

and $\boldsymbol{x}_{i,-k}^*$ denotes $\boldsymbol{x}_i^*$ excluding the k-th element x_{ik}^*. There are two main steps for sampling from (9).

Step (i). First we sample d_{ik} from the conditional distribution of d_{ik} given by

$$\begin{aligned}
\pi(d_{ik} | \cdot) &= \int \int \pi(d_{ik}, x_{ik}^*, y_{ik}^* | \cdot) dx_{ik}^* dy_{ik}^* \\
&\propto \pi_i^{d_{ik}} (1-\pi_i)^{1-d_{ik}} \left\{ d_{ik} \Phi\left(\frac{\mu_{x,ik}}{\sigma_{x,ik}} \right) + (1-d_{ik}) \Phi\left(\mu_{y,ik} \right) \right\}^{z_{ik}} \left\{ d_{ik} \Phi\left(\frac{-\mu_{x,ik}}{\sigma_{x,ik}} \right) + (1-d_{ik}) \Phi\left(-\mu_{y,ik} \right) \right\}^{1-z_{ik}}
\end{aligned}$$

by using the following discrete indicator variable samplings.

1. If $z_{ik} = 1$, generate $d_{ik} \sim Bernoulli(\bar{\pi}_{ik})$ where

$$\bar{\pi}_{ik} = \frac{\pi_i \Phi\left(\frac{\mu_{x,ik}}{\sigma_{x,ik}} \right)}{\pi_i \Phi\left(\frac{\mu_{x,ik}}{\sigma_{x,ik}} \right) + (1-\pi_i) \Phi\left(\mu_{y,ik} \right)}.$$

2. If $z_{ik} = 0$, generate $d_{ik} \sim Bernoulli(\bar{\pi}_{ik})$ where

$$\bar{\pi}_{ik} = \frac{\pi_i \Phi\left(\frac{-\mu_{x,ik}}{\sigma_{x,ik}} \right)}{\pi_i \Phi\left(\frac{-\mu_{x,ik}}{\sigma_{x,ik}} \right) + (1-\pi_i) \Phi\left(-\mu_{y,ik} \right)}.$$

Step (ii). Next we generate x_{ik}^*, y_{ik}^* given d_{ik} as follows:

1. If $z_{ik} = 1$ and $d_{ik} = 1$, $x_{ik}^* \sim \mathcal{TN}_{(0,\infty)}(\mu_{x,ik}, \sigma_{x,ik}^2)$ and $y_{ik}^* \sim \mathcal{N}(\mu_{y,ik}, 1)$.
2. If $z_{ik} = 1$ and $d_{ik} = 0$, $x_{ik}^* \sim \mathcal{N}(\mu_{x,ik}, \sigma_{x,ik}^2)$ and $y_{ik}^* \sim \mathcal{TN}_{(0,\infty)}(\mu_{y,ik}, 1)$.
3. If $z_{ik} = 0$ and $d_{ik} = 1$, $x_{ik}^* \sim \mathcal{TN}_{(-\infty,0]}(\mu_{x,ik}, \sigma_{x,ik}^2)$ and $y_{ik}^* \sim \mathcal{N}(\mu_{y,ik}, 1)$.
4. If $z_{ik} = 0$ and $d_{ik} = 0$, $x_{ik}^* \sim \mathcal{N}(\mu_{x,ik}, \sigma_{x,ik}^2)$ and $y_{ik}^* \sim \mathcal{TN}_{(-\infty,0]}(\mu_{y,ik}, 1)$.

where $\mathcal{TN}_{(a,b)}(m, s^2)$ denotes a normal distribution $\mathcal{N}(m, s^2)$ truncated on (a, b).

Step 3: Generation of $\mathbf{A}$

By Gaussian-Gaussian conjugacy, the full conditional of $\mathbf{A}$ is

$$\mathbf{A} | \cdot \sim \mathcal{N}_{p,K}(\mathbf{M}_1, \boldsymbol{\Phi}_1 \otimes \mathbf{I}_K), \qquad (10)$$

where

$$\mathbf{M}_1 = \boldsymbol{\Phi}_1 \mathbf{V}' \mathbf{Y}^*, \quad \boldsymbol{\Phi}_1^{-1} = \boldsymbol{\Phi}_0^{-1} + \mathbf{V}'\mathbf{V},$$

and so, $\mathbf{A}$ can be sampled from the normal distribution in (10).

Step 4: Generation of $\mathbf{B}$ and $\mathbf{R}$

We sample $\mathbf{B}$ and $\mathbf{R}$ using an efficient algorithm below as in Talhouk et al. (2012).

Step (i). Let $\lambda \sim \mathcal{IG}(a, b)$ denote that λ follows inverse gamma distribution whose density is given by

$$f(\lambda) \propto \lambda^{-(a+1)} \exp\left(-\frac{b}{\lambda} \right).$$

Generate

$$\boldsymbol{\Lambda} = \text{diagonal}(\lambda_1, \ldots, \lambda_K), \quad \lambda_k^2 \sim \mathcal{IG}\left(\frac{n_0}{2}, \frac{r^{kk}}{2} \right), \quad k = 1, \ldots, K,$$

where $\lambda_k > 0$, r^{kk} is the (k, k)-th element of $\mathbf{R}^{-1}$, and set $\tilde{\mathbf{X}} = \mathbf{X}^* \boldsymbol{\Lambda}$.

Table 1
The eight sensitive-unrelated question pairs in the drug administration survey.

Variable*	Question Description#	Question Type
Q1	I had the experience in giving medications to a wrong patient and did not report it to the hospital. (*I had breakfast at home at least 3 times in the past week.*)	Wrong patient
Q2	I had the experience in giving a wrong medication and did not report it to the hospital. (*I bought a new mobile phone in the past 12 months.*)	Wrong medication
Q3	I had the experience in giving a medication in a wrong dose and did not report it to the hospital. (*I ate chocolate at least 1 time in the past week.*)	Wrong dose
Q4	I had the experience in giving a medication at a wrong time and did not report it to the hospital. (*I had dinner at home at least 3 times in the past week.*)	Wrong time
Q5	I had the experience in giving a medication via a wrong route and did not report it to the hospital. (*I watched a move in cinema at least 1 time in the past 3 months.*)	Wrong route
Q6	I had an experience NOT to press the button on the scanner to acknowledge the drug administration process has been completed AFTER a patient has taken the medication. (*I have NOT been to Australia.*)	Technology-related administration omitted
Q7	I had an experience NOT to review the pump setting at 15 minutes after the commencement of the heparin infusion. (*I do NOT know how to swim the backstroke.*)	Physical administration omitted
Q8	I had an experience to administer leftover drugs from Patient B to Patient A when the drug prescribed to Patient A is not available. (*I prefer coffee to tea.*)	Physical administration omitted

*Binary response: disagree/agree # Statements in parentheses are the innocuous questions paired with their respective sensitive questions.

Step (ii). Let $\boldsymbol{\Sigma} \sim \mathcal{IW}(\nu, \mathbf{S})$ denote that $\boldsymbol{\Sigma}$ follows the inverse Wishart distribution whose probability density function is given by

$$f(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{\nu+K+1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{S}\right)\right\}.$$

Generate $\boldsymbol{\Sigma}$, $\tilde{\mathbf{B}}$, and $\tilde{\mathbf{X}}$ using

$$\boldsymbol{\Sigma}|\tilde{\mathbf{X}} \sim \mathcal{IW}(n+n_0, \mathbf{S}), \quad \mathbf{S} = \tilde{\mathbf{X}}'\tilde{\mathbf{X}} - \mathbf{M}_2'\boldsymbol{\Psi}_1^{-1}\mathbf{M}_2 + \mathbf{I}_K,$$

$$\tilde{\mathbf{B}}|\boldsymbol{\Sigma}, \tilde{\mathbf{X}} \sim \mathcal{N}_{p,K}(\mathbf{M}_2, \boldsymbol{\Psi}_1 \otimes \boldsymbol{\Sigma}),$$

and generate $\tilde{\mathbf{X}}|\tilde{\mathbf{B}}, \boldsymbol{\Sigma}, \boldsymbol{y}^*, \boldsymbol{d}, \mathbf{A}, \boldsymbol{z}$ from $\mathcal{N}_{n,K}(\mathbf{M}_2, \boldsymbol{\Psi}_1 \otimes \mathbf{I}_K)$ as in (10) where

$$\mathbf{M}_2 = \boldsymbol{\Psi}_1\mathbf{V}'\tilde{\mathbf{X}}, \quad \boldsymbol{\Psi}_1^{-1} = \boldsymbol{\Psi}_0^{-1} + \mathbf{V}'\mathbf{V}.$$

Step (iii). Transform $\boldsymbol{\Sigma}$, $\tilde{\mathbf{B}}$, and $\tilde{\mathbf{X}}$ back to $\mathbf{R}$, $\mathbf{B}$ and $\mathbf{X}^*$ as follows:

$$\mathbf{R} = \boldsymbol{\Lambda}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Lambda}^{-1}, \quad \mathbf{B} = \tilde{\mathbf{B}}\boldsymbol{\Lambda}^{-1}, \quad \mathbf{X}^* = \tilde{\mathbf{X}}\boldsymbol{\Lambda}^{-1}, \quad \boldsymbol{\Lambda} = \mathrm{diag}\left(\sigma_{11}^{1/2}, \sigma_{22}^{1/2}, \ldots, \sigma_{KK}^{1/2}\right).$$

where σ_{ii} denotes the (i, i)-th element of $\boldsymbol{\Sigma}$. See Appendix A.2 for the derivation of the conditional posterior distribution. Note that we also updated $\mathbf{X}^*$ as by-product in addition to Step 2. In this step, we use the Gibbs sampling approach, but other approaches using the Metropolis-Hastings algorithm are also discussed in Liu (2008), Liu and Daniels (2006) and Zhang et al. (2006).

Finally, the numerical example of the MCMC simulation using simulated data is given in Appendix Appendix C.

3. Analysis of the drug administration survey data

3.1. Survey design and sample

In this section, we apply the proposed methods to the drug administration survey, which consisted of eight dichotomous questions on different types of drug administration error (see Table 1). Our target population was the 2,514 full-time nursing staff who are using an electronic system for drug administration in three hospitals, namely hospital A, hospital B, and hospital C under the cluster. We randomly ordered the emails of the target nurses in a list. The first half of the target respondents in the list received version 1 of the online survey (sample 1, $p_1 = 1/3$), and the second half of them received version 2 (sample 2, $p_2 = 2/3$). In other words, respondents in sample 1 and sample 2 have respectively one-third and two-thirds of a chance to answer the sensitive question in each sensitive-unrelated pair under the UQD. There are two main considerations in choosing p_1 and p_2. First, the choice of p has to balance the data privacy and the statistical accuracy (Kwan et al., 2010). We cannot make p, the probability of answering sensitive questions, very large. Otherwise, respondents may think that their data privacy cannot be adequately protected. Very small values of p_1 or p_2 will affect the statistical accuracy. Experience from our previous studies (Kwan et al., 2010; Chung et al., 2018; Chu et al., 2020) shows that p_1 and p_2 lying from 1/3 to 2/3 are promising choices. The second consideration is to facilitate the implementation of the RRT mechanism in an online platform (Chu et al., 2018). Operationally, p has to be either $1/m$ or $1 - 1/m$, where m is a positive integer. To follow the online implementation scheme in Kwan et al. (2010) and Chu et al. (2018), we asked respondents to pick a number from 1 to m, and then a random number was generated to see if there is a match, or not a match, between the respondents'

Table 2
A list of the independent variables in the 10×1 vector $\boldsymbol{v}_i$.

Variable	Description
Const	Constant (Registered nurse at hospital A with more than 15 years of nursing experience)
HospitalB	1 if hospital B, 0 otherwise.
HospitalC	1 if hospital C, 0 otherwise.
Rank2	1 if the nurse is in rank 2.
Rank3	1 if the nurse is in rank 3.
Year1	1 if 0-1 year of nursing experience, 0 otherwise.
Year2	1 if >1-3 years of nursing experience, 0 otherwise.
Year3	1 if >3-5 years of nursing experience, 0 otherwise.
Year4	1 if >5-10 years of nursing experience, 0 otherwise.
Year5	1 if >10-15 years of nursing experience, 0 otherwise.

number and the random number. Therefore, it is helpful for us to fix p_1 and p_2 as $1/m$ (a match) or $1 - 1/m$ (not a match) in online surveys. Again, experience indicates that taking $m = 3$ can help respondents to follow the online randomization procedure effectively while balancing the statistical accuracy and data privacy. We used an anonymous approach to conduct the survey. Ultimately, 401 and 585 usable responses were obtained in sample 1 and sample 2, respectively. In addition to knowing which hospital the nurses are working for, we know which rank (out of ranks 1 to 3) they are at and their years of nursing experience. More than 40% of the respondents have 15 years of nursing experience or more, and around 25% have five or fewer years of nursing experience.

To understand some regular practices of nurses in administrating medications to patients, we included eight pairs of questions ($K = 8$) in the survey, as per Table 1. The RRT survey in this paper was approved by the Ethics Committee of the hospital cluster. In Table 1, we can see that the eight questions of interest are related to medication procedures. Five questions are related to possible actions that may not align with the hospitals' guidelines. For example, they describe giving medications to the 'wrong patient', with the 'wrong dose', and at the 'wrong time'. Three of the sensitive questions are related to the omission of required steps by the hospitals' standard practices. All eight questions are regarded as sensitive as nurses are likely to be hesitant to answer or reluctant to say 'Yes' if the questions are asked directly. Each sensitive question is paired with an innocuous question that is nonsensitive and easy to answer. All respondents were given clear instructions on how the RRT is executed and informed that the main purpose of the RRT is to protect their data privacy. All eight questions were pilot-tested and pretested to remove all ambiguities and any confusing wording. Through the RRT data collected for the questions in Table 1, we perform the proposed multivariate probit analysis in Section 2 to understand how frequently nurses may use a wrong procedure or bypass required steps as stipulated by the hospitals' guidelines for drug administration.

3.2. Probit analysis

We adopt the multivariate probit model in (3) to (5) with $K \times 1$ vectors, $\boldsymbol{x}_i^* | \boldsymbol{v}_i, \mathbf{B}, \mathbf{R} \sim \mathcal{N}(\mathbf{B}'\boldsymbol{v}_i, \mathbf{R})$, $\boldsymbol{y}_i^* | \boldsymbol{v}_i, \mathbf{A} \sim \mathcal{N}(\mathbf{A}'\boldsymbol{v}_i, \mathbf{I}_K)$, and the RRT data to study the nurses' general practice in drug administration. To estimate the effect of hospital, rank, and years of nursing experience on drug administration error, we list all the $p = 10$ covariates $\boldsymbol{v}_i$ in Table 2. We define the covariates as indicator variables such as HospitalB= I(respondent is from hospital B) and HospitalC= I(respondent is from hospital C). With regard to years of nursing experience, the notation '>1-3' indicates more than one year and at most three years of experience. The two other indicators for the rank and other indicator variables for the years of experience are defined similarly.

To perform the Bayesian analysis of the $p \times K$ (10×8) unknown parameter matrices $\mathbf{A}$ and $\mathbf{B}$ and the $K \times K$ (8×8) correlation matrix $\mathbf{R}$, we assume $\mathbf{A} \sim \mathcal{N}_{10,8}(\mathbf{O}_{10,8}, 4\mathbf{I}_{10} \otimes \mathbf{I}_8)$, $\mathbf{B}|\mathbf{R} \sim \mathcal{N}_{10,8}(\mathbf{O}_{10,8}, 4\mathbf{I}_{10} \otimes \mathbf{R})$ and the marginal uniform distribution for $\mathbf{R}$ given in (7) with $n_0 = K + 1$. The prior covariance matrices are large enough to account that all independent variables are indicator variables and that the variance of the latent variables is one. We iterate our MCMC algorithm 200,000 times after discarding 20,000 samples as burn-in periods. The effective sample sizes, which are defined as the sample sizes divided by the corresponding inefficiency factor (IF) in Table 3, vary about from 336 to 670. The MCMC sample paths are found to be stable and the chains mix well. However, the figures of the MCMC sample paths are omitted to save the space.

3.3. Results and implications

Table 3 presents the posterior mean and standard deviations of $\mathbf{B}$ from the MCMC sampling. There are altogether $K = 8$ columns summarizing the results of the eight sensitive questions. Each column gives the posterior estimates of 10×1 vector $\boldsymbol{\beta}_k$, the elements of which represent the effect of the covariates $\boldsymbol{v}_i$ in Table 2, for $k = 1, \ldots, K$. From (4), $x_{ik}^* | \boldsymbol{v}_i, \boldsymbol{\beta}_k \sim \mathcal{N}(\boldsymbol{v}_i' \boldsymbol{\beta}_k, 1)$, and so $Pr(x_{ik} = 1 | \boldsymbol{v}_i) = Pr(x_{ik}^* > 0 | \boldsymbol{v}_i) = \Phi(\boldsymbol{v}_i' \boldsymbol{\beta}_k)$, where $\Phi(\cdot)$ is the distribution function of the standard normal distribution. As the elements of $\boldsymbol{v}_i$ are either 0 or 1, representing respondents' demographic attributes, more positive elements in $\boldsymbol{\beta}_k$ of

Table 3
The posterior means and standard deviations (Std) of the 10×8 matrix **B** in the probit regression.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Const	**-0.6831**	**-0.7156**	**-0.7502**	**-0.8011**	**-0.9401**	**-1.0236**	**-0.8550**	**-0.7744**
(Std)	(0.1979)	(0.1703)	(0.1623)	(0.2001)	(0.2110)	(0.2346)	(0.2535)	(0.2836)
IF	392	381	363	466	394	405	427	475
HospitalB	0.1207	0.0003	-0.2039	-0.1159	-0.0852	-0.4421	-0.1334	-0.2702
(Std)	(0.1879)	(0.1760)	(0.1915)	(0.1961)	(0.2073)	(0.2374)	(0.2906)	(0.2561)
IF	420	456	480	495	427	442	525	476
HospitalC	**0.6920**	**1.0267**	**0.8931**	**0.8843**	0.5731	0.6826	0.5478	**0.9184**
(Std)	(0.2997)	(0.3053)	(0.2818)	(0.3047)	(0.3232)	(0.3737)	(0.4347)	(0.4033)
IF	396	459	444	468	432	463	485	449
Rank2	0.2941	0.2228	0.3127	0.3406	0.3214	**0.8013**	**0.7166**	0.4514
(Std)	(0.2930)	(0.2679)	(0.3297)	(0.2867)	(0.3097)	(0.3053)	(0.3200)	(0.3688)
IF	300	315	441	374	345	298	304	370
Rank3	0.1668	0.1176	0.0182	0.0543	-0.2614	0.0046	0.1885	0.1339
(Std)	(0.2079)	(0.2084)	(0.1872)	(0.1942)	(0.2560)	(0.3084)	(0.3340)	(0.3366)
IF	394	447	418	429	438	487	512	504
Year1	-0.1012	-0.3340	-0.2678	-0.3417	-0.2220	-0.4381	-0.4101	-0.3852
(Std)	(0.4768)	(0.3714)	(0.4038)	(0.3727)	(0.4314)	(0.4703)	(0.5114)	(0.4653)
IF	466	418	463	425	395	414	433	405
Year2	0.0106	0.3693	0.2093	0.1946	0.3817	-0.8504	-1.1615	-1.032
(Std)	(0.3248)	(0.2612)	(0.2546)	(0.2721)	(0.3328)	(0.7049)	(0.7004)	(0.8535)
IF	420	379	368	411	437	584	570	595
Year3	0.0253	-0.2314	-0.0116	-0.0030	-0.1301	0.4915	0.2927	0.3050
(Std)	(0.3988)	(0.3093)	(0.3500)	(0.3229)	(0.4226)	(0.3234)	(0.3325)	(0.3636)
IF	464	426	480	465	487	348	357	390
Year4	-0.2623	0.1367	0.2089	0.1161	0.2738	0.4398	0.5617	0.3792
(Std)	(0.2365)	(0.2772)	(0.2705)	(0.2461)	(0.3058)	(0.2772)	(0.3063)	(0.3244)
IF	341	494	493	461	469	391	422	448
Year5	0.4003	0.3958	0.3178	0.1945	0.1036	-0.2818	-0.4399	-0.1988
(Std)	(0.2743)	(0.2201)	(0.2560)	(0.3159)	(0.3938)	(0.4602)	(0.4948)	(0.5055)
IF	428	373	456	520	523	525	544	539

(The posterior mean with its absolute value greater than two standard deviation is in bold letters.) To measure how well the chain mixes, we calculate the inefficiency factors (IF). The inefficiency factor (equivalently the autocorrelation time) is defined as $1 + 2\sum_{s=1}^{\infty} \rho_s$ where ρ_s is the sample autocorrelation at lag s calculated from the sampled values. It is also the ratio of the numerical variance of the posterior sample mean to the variance of the posterior sample mean from the hypothetical uncorrelated draws. It suggests the relative number of correlated draws necessary to attain the same variance of the posterior sample mean from the uncorrelated draws. The effective sample sizes, which are defined as the sample sizes divided by the corresponding inefficiency factor, vary about from 336 to 670.

Question k imply that the i-th respondent with attributes given by $\boldsymbol{v}_i$ is more likely to say ‘Yes’ in the k-th sensitive question, or to have a higher tendency to bypass hospital guidelines related to Question k.

To analyze the effect of the demographic variables (i.e., which hospital they are working in, which rank they are in, and how much nursing experience they have), we study the results in Table 3 in detail. Regarding the two rows of coefficients for HospitalB and HospitalC in Table 3, all coefficients for HospitalC are positive and higher than the corresponding coefficients for HospitalB. This indicates that nurses working in hospital C have a higher tendency to say ‘Yes’ in answer to the eight sensitive questions. Many of the coefficients of HospitalB are negative, except two small coefficients in Q1 and Q2. In general, nurses working in hospital B seem to follow the hospital drug administration guidelines better than those working in hospital A. Regarding the rank of the nurses, all coefficients of Rank2 are positive and higher than the corresponding coefficients of Rank3. This finding suggests that nurses in rank 2 have a higher tendency to say ‘Yes’ in answer to the eight sensitive questions or are more likely to bypass the hospital guidelines than nurses in rank 1 and nurses in rank 3. In the row of Rank3, we observe that most coefficients are positive except those in Q5. Therefore, nurses in rank 1 are likely to follow the hospital guidelines on drug administration better than nurses in rank 3. Concerning the attribute of the years of experience, we set ‘more than 15 years’, or Year0, as the reference group. In the row of Year4, the coefficients for Q7 and Q8 are the highest (among Year1 to Year5) and positive. Nurses with >5-10 years of experience have a higher tendency to omit some steps as seen in Q7 and Q8 than the other five experience groups. In the row of Year5, the coefficients for Q1 to Q4 are the highest (among Year1 to Year5) and positive. Nurses with >10-15 years of experience have a higher tendency than the other five experience groups to have made drug administration errors in the past and to have not reported it to the hospital. For Year1, all coefficients are negative, indicating that nurses with 0-1 year of experience have a lower tendency than nurses in Year0 to not follow the hospital guidelines. In the rows of Year1, Year2, and Year5, all coefficients for Q6, Q7, and Q8 are negative. Therefore, nurses in these three experience groups have a lower tendency than Year0 to omit steps.

Figure 1 gives the estimates of the conditional probability of saying ‘Yes’ given $\boldsymbol{v}_i$, that is, $\mathrm{P}(x_{ik} = 1|\boldsymbol{v}_i)$. We divide the whole sample into nine groups based on the hospitals they serve and their ranks. Different colors indicate the different likelihoods to say ‘Yes’ in answer to the sensitive questions. The four groups formed by the combinations of (hospital A, hospital B) and (rank 1, rank 3) display very coherent probability patterns. Fig. 1 tells us these four groups are quite ho-

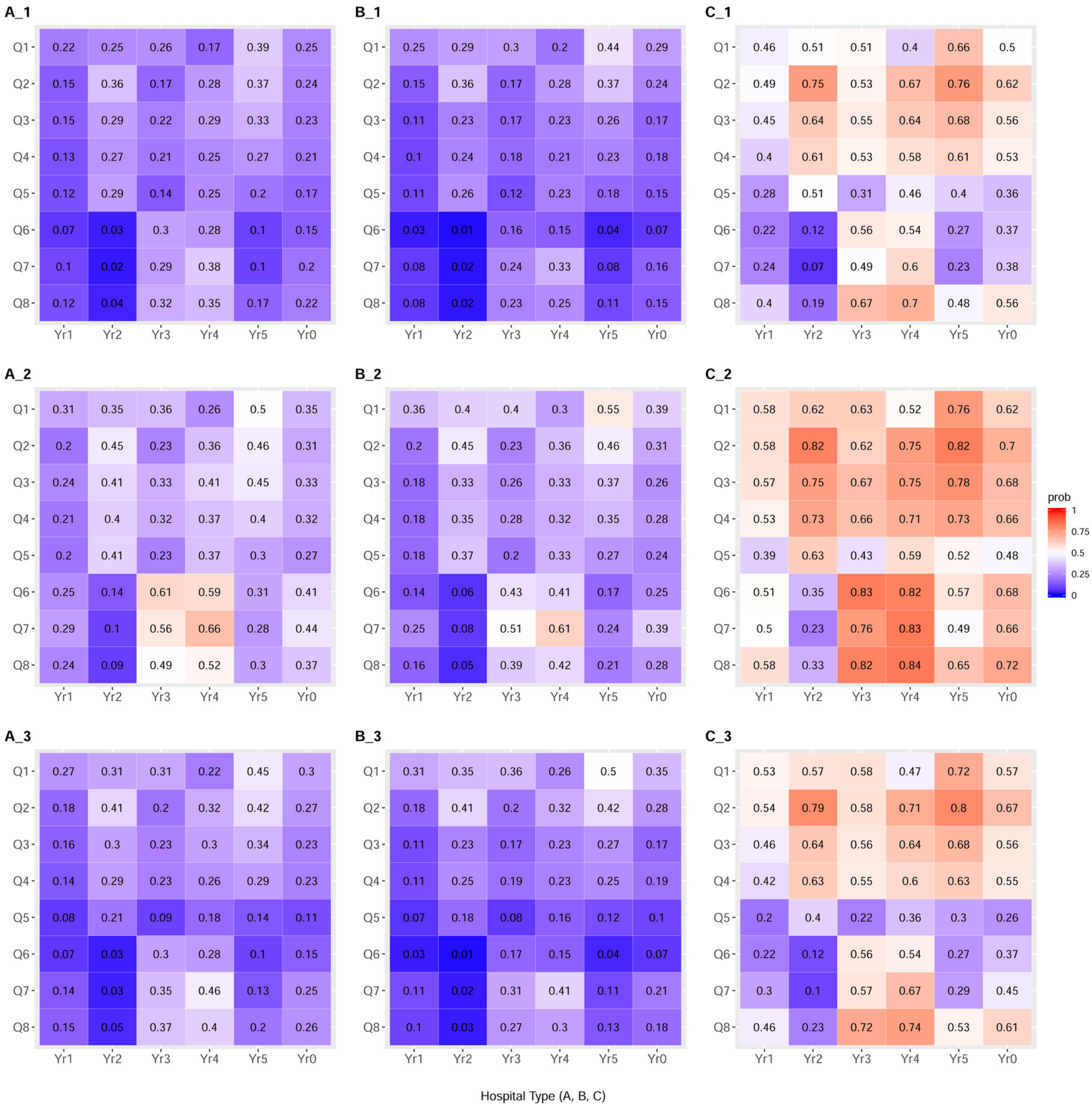


Fig. 1. The Bayesian estimate of $Pr(x_{ik}=1|\boldsymbol{v}_i)$. The notation Hospital_Rank is used for the label (e.g. A_2 implies Hospital A and Rank 2).

mogeneous in terms of their approach to reporting issues and omitting steps in the hospital guidelines. In general, the probability pattern differences of rank 1 (row 1) and rank 3 (row 3) are not practically significant. Among the nine groups, nurses in rank 2 working in hospital C are most likely to say ‘Yes’ in answer to the eight sensitive questions. Senior management can consider devoting more resources to promoting the importance of the hospital guidelines to this group of nurses.

3.4. Dependence analysis

To examine the dependence of responses to different questions of the i-th respondent, we calculate the conditional correlation of the indicator variables $I(x_{ik}=1)$ and $I(x_{ik'}=1)$ given $\boldsymbol{v}_i$ for $k,k'=1,\ldots,K$ as

$$\rho_{k,k'|\boldsymbol{v}_i}=\frac{Pr(x_{ik}=1,x_{ik'}=1|\boldsymbol{v}_i)-Pr(x_{ik}=1|\boldsymbol{v}_i)Pr(x_{ik'}=1|\boldsymbol{v}_i)}{\sqrt{Pr(x_{ik}=1|\boldsymbol{v}_i)Pr(x_{ik}=0|\boldsymbol{v}_i)}\sqrt{Pr(x_{ik'}=1|\boldsymbol{v}_i)Pr(x_{ik'}=0|\boldsymbol{v}_i)}}. \tag{11}$$

The probability $Pr(x_{ik'}=1|\boldsymbol{v}_i)$ is obtained in Section 3.3 and

$$Pr(x_{ik}=1,x_{ik'}=1|\boldsymbol{v}_i)=Pr(x^*_{ik}>0,x^*_{ik'}>0|\boldsymbol{v}_i),$$

Table 4
Estimation results of the correlation matrix **R** in the probit model in (5). Top: posterior mean, middle: (posterior standard deviation), bottom: inefficiency factor.

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	
1	**0.9621** (0.0266) 486	**0.9639** (0.0329) 538	**0.9692** (0.0229) 475	**0.9447** (0.0360) 463	0.7728 (0.1191) 566	0.7131 (0.1290) 568	0.7063 (0.1258) 546	Q1
	1	**0.9761** (0.0202) 507	**0.9775** (0.0194) 500	**0.9556** (0.0321) 483	0.7523 (0.1053) 542	0.6878 (0.1091) 538	0.6817 (0.1126) 531	Q2
		1	**0.9834** (0.0155) 485	**0.9638** (0.0269) 465	0.8257 (0.0782) 516	0.7679 (0.0812) 502	0.7628 (0.0878) 505	Q3
			1	**0.9652** (0.0257) 449	0.8075 (0.0890) 539	0.7463 (0.1002) 545	0.7402 (0.0986) 523	Q4
				1	0.8036 (0.0919) 515	0.7425 (0.1042) 517	0.7280 (0.1230) 535	Q5
					1	**0.9597** (0.0286) 431	**0.9489** (0.0376) 443	Q6
						1	**0.9600** (0.0258) 390	Q7
							1	Q8

(The posterior mean greater than 0.9 in bold letters)

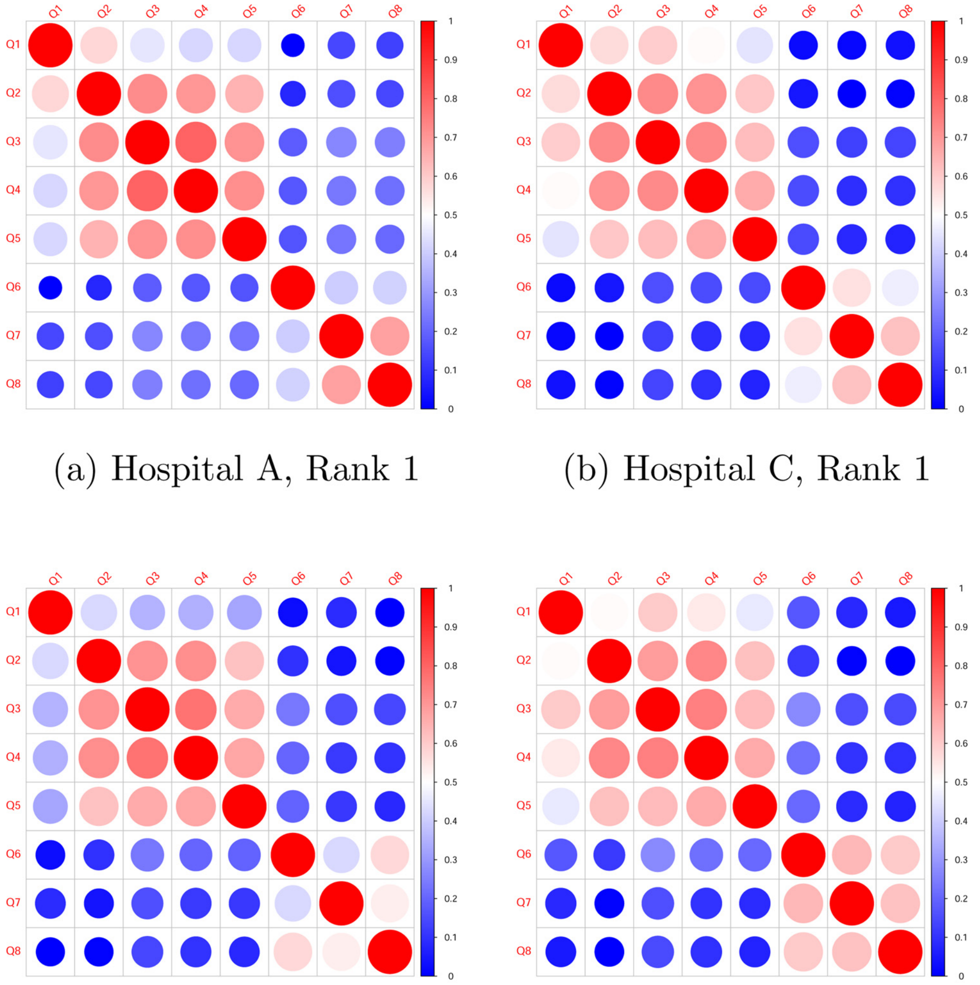
Table 5
Varimax rotated factor loadings for the posterior mean of **R**.

	Factor 1	Factor 2
Q1	**0.9006**	0.3949
Q2	**0.9210**	0.3597
Q3	**0.8726**	0.4713
Q4	**0.8884**	0.4398
Q5	**0.8647**	0.4429
Q6	0.4865	**0.8515**
Q7	0.3956	**0.8995**
Q8	0.3890	**0.8971**

which is based on the bivariate normal distribution of $(x^*_{ik}, x^*_{ik'})'$ with mean $(\boldsymbol{v}_i'\boldsymbol{\beta}_k, \boldsymbol{v}_i'\boldsymbol{\beta}_{k'})'$ and covariance matrix $\begin{pmatrix} 1 & \rho_{kk'} \\ \rho_{kk'} & 1 \end{pmatrix}$, where $\rho_{kk'}$ is the (k, k')-th element of **R** as given in Table 4: where we can see high correlations among $Q1 - Q5$ and among $Q6 - Q8$. The measure $\rho_{k,k'|\boldsymbol{v}_i}$ in (11) can be interpreted as a kind of 'risk correlation' as it measures how strongly the incident in question k is related to the incident in question k'. Fig. 2 shows the correlation matrix plot of four specific groups of nurses having 0-1 year of experience. We can observe two blocks of highly correlated questions. They are Q1-Q5 and Q6-Q8. For Q1-Q5, Q1 is relatively less correlated with others in hospital A, and Q5 is relatively less correlated with others in hospital C. Generally, Q1-Q5 are highly correlated, indicating that the five 'wrongs', namely 'wrong patient', 'wrong medication', 'wrong dose', 'wrong time' and 'wrong route' as per Table 1, tend to occur together. The correlation plot of nurses in rank 1 and nurses in rank 2 are not that different. (Q6, Q7, Q8) is a relatively less correlated block than the one in Q1-Q5 but omissions of the three steps also tend to occur together. We also observe weak correlations between the two blocks Q1-Q5 and Q6-Q8, indicating that the occurrence of the five 'wrongs' is not much related to the omission of steps in Q6-Q8.

To understand the structure of **R** in Table 4, we perform a factor analysis on the correlation matrix. Table 5 displays the varimax rotated factor loadings. Two factors are extracted together, accounting for 95.23% of the total variance. Variables Q1, Q2, Q3, Q4, and Q5 have high respective loadings of 0.9006, 0.9210, 0.8726, 0.8884, and 0.8647 on factor 1. This suggests that the response variables for the first five questions define the first factor. Variables Q6, Q7, and Q8 have high respective loadings of 0.8515, 0.8995, and 0.8971 on factor 2, and thus the response variables for the last three questions define the second factor.

To further analyze the correlation structure, we produce posterior mean estimates for $\mathbf{R}^{-1}$ using the MCMC draws of **R**. Under the multivariate normal distributional assumption, the (k, k')-th element of $\mathbf{R}^{-1}$ is the conditional correlation of x^*_{ik} and $x^*_{ik'}$ given other x^*_{ij}, $j \neq k$ or k'. Fig. 3 displays a heatmap of the conditional correlation pattern. We can observe that there are two clusters highlighted by blue regions. One cluster is defined by Q1-Q5 and the second cluster is defined by



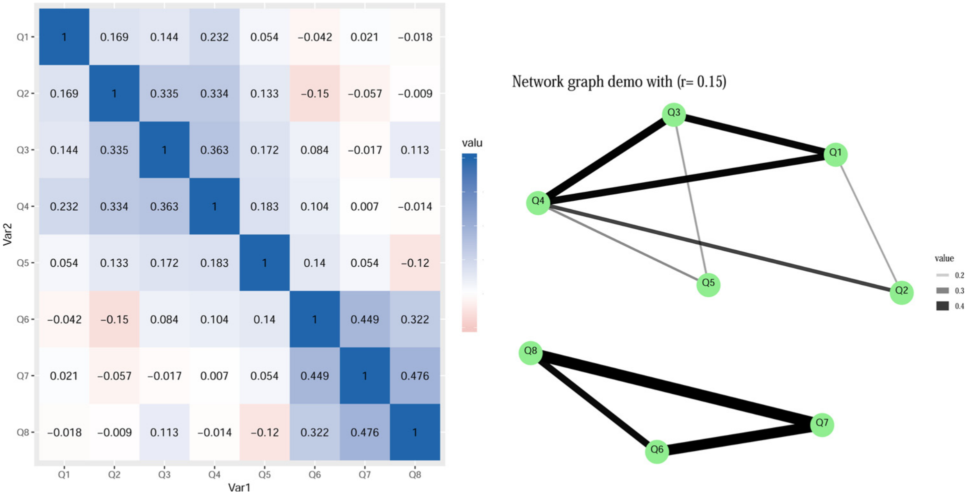
(a) Hospital A, Rank 1

(b) Hospital C, Rank 1

(c) Hospital A, Rank 2

(d) Hospital C, Rank 2

Fig. 2. Bayesian estimate of the conditional correlation $\rho_{k,k'|\boldsymbol{v}_i}$ for $k, k' = 1, \ldots, K$ of nurses having 0-1 year of experience.



Var2 \ Var1	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Q1	1	0.169	0.144	0.232	0.054	−0.042	0.021	−0.018
Q2	0.169	1	0.335	0.334	0.133	−0.15	−0.057	−0.009
Q3	0.144	0.335	1	0.363	0.172	0.084	−0.017	0.113
Q4	0.232	0.334	0.363	1	0.183	0.104	0.007	−0.014
Q5	0.054	0.133	0.172	0.183	1	0.14	0.054	−0.12
Q6	−0.042	−0.15	0.084	0.104	0.14	1	0.449	0.322
Q7	0.021	−0.057	−0.017	0.007	0.054	0.449	1	0.476
Q8	−0.018	−0.009	0.113	−0.014	−0.12	0.322	0.476	1

Fig. 3. The Bayesian estimates of the partial correlation matrix $\mathbf{R}^{-1}$ on the left and a network graph based on the partial correlation matrix on the right.

Table 6
The probability distribution of the attitude score, S_i in the three hospital groups, $Pr(S_i = j|i$ is a hospital group) and their standard errors (S.E.)

j	Hospital A	Hospital B	Hospital C
0	0.532 (0.025)	0.540 (0.025)	0.215 (0.021)
1	0.090 (0.014)	0.105 (0.015)	0.088 (0.014)
2	0.064 (0.012)	0.075 (0.013)	0.067 (0.012)
3	0.058 (0.012)	0.052 (0.011)	0.062 (0.012)
4	0.049 (0.011)	0.048 (0.011)	0.075 (0.013)
5	0.058 (0.012)	0.061 (0.011)	0.088 (0.014)
6	0.036 (0.009)	0.033 (0.009)	0.084 (0.014)
7	0.040 (0.010)	0.033 (0.009)	0.107 (0.015)
8	0.074 (0.013)	0.054 (0.011)	0.214 (0.021)
$E[S_i\|\mathbf{Z}]$	1.966	1.769	**4.113**

Q6, Q7, and Q8. From the graphical network perspective, if we treat the eight sensitive questions as nodes and link the nodes whenever their correlations are greater than 0.15, we see the network graph on the right-hand side of Fig. 3. Again, it is not difficult to identify two clusters of nodes. Both the factor analysis and the conditional correlation analysis based on $\mathbf{R}^{-1}$ suggest that a two-factor model structure can explain the correlation pattern in $\mathbf{R}$. In other words, on top of the demographic attributes of nurses, there are potentially two hidden factors (possibly related to their attitudes), which may help to further explain the dependence of the responses to the eight sensitive questions. We discuss the posterior analysis of a two-factor multivariate probit model for the RRT data in Section 4.2.

4. Discussion

4.1. Overall attitude scoring

From the administration perspective, we are also interested in learning about overall attitude of nurses as measured by the sensitive questions. A score for respondent i can be compiled as

$$S_i = \sum_{k=1}^{K} w_k I(x_{ik} = 1), \tag{12}$$

where w_k is the weight associated with question k to account for the relative importance of the response to question k. The conditional expectation of S_i given $\boldsymbol{v}_i$ is a byproduct of the results in Section 3 as $E[S_i|\boldsymbol{v}_i] = \sum_{k=1}^{K} w_k Pr(x_{ik} = 1|\boldsymbol{v}_i)$. We can also determine the conditional distribution of S_i given $\boldsymbol{v}_i$ by calculating $Pr(x_{ik_1} = \cdots = x_{ik_g} = 1)$ for distinct $k_1, ..., k_g$ and $g = 1, ..., K$ using the multivariate normal distribution of $\boldsymbol{x}_i^*$ in (5) with the posterior mean estimates of $\mathbf{B}$ and $\mathbf{R}$ in Table 3 and 4. Alternatively, we can conduct a full Bayesian analysis on S_i. Suppose $x_{ik}^{*[m]}$, $m = 1, ..., M$ are the MCMC draws using the sampling scheme in Section 2.4, where M is the number of MCMC iterations. For each m from 1 to M, we calculate $S_i^{[m]}$ using (12). Then, we can estimate the conditional distribution of S_i given i in A, the subpopulation of interest, by obtaining the posterior distribution based on $S_i^{[m]}$ for all $i \in A$. For example, if w_k's are equal to one such that $S_i = \sum_{k=1}^{K} I(x_{ik} = 1)$, or the number of questions with a 'Yes' answer, S_i has a discrete distribution with possible values, 0,..., K. Then, we can form posterior estimates of the distribution of S_i by estimating

$$Pr(S_i = j|i \in A) \approx \frac{\sum_i I(i \in A) \sum_{m=1}^{M} I(S_i^{[m]} = j)}{M \sum_i I(i \in A)}. \tag{13}$$

As an illustration, we compute $Pr(S_i = j|i \in A)$ using (13), where A defines a subpopulation of a hospital group. Table 6 presents the probability distribution of the attitude score, S_i, in Hospital A, B, and C. Hospitals A and B have similar distributions in S_i, and both hospitals have more than a 50% chance of having $S_i = 0$ or more than half of the nurses with no experience of the five 'wrongs' in Q1-Q5 and no experience of omitting steps in Q6-Q8. Their average attitude scores are 1.966 and 1.769. For hospital C, the average attitude score is higher. The attitude score result is consistent with the findings in Fig. 1.

4.2. Factor structure for the correlation matrix

Instead of sampling the correlation matrix $\mathbf{R}$, we can consider ways of simplifying $\mathbf{R}$ or building structures similar to those in Lee et al. (2010) and Talhouk et al. (2012). In Section 3, we have employed a factor model to understand the structure of the posterior mean of $\mathbf{R}$ from the MCMC posterior draws under the prior (7). In the discussion below, we give a brief account on an alternative approach that uses a factor analytic approach directly as a prior on $\mathbf{R}$ instead, where we sample latent factors and estimate factor loadings. Let $\boldsymbol{f}_i = (f_{i1}, \ldots, f_{iL})'$ denote the $L \times 1$ vector with L factors for the i-th individual and $\mathbf{C} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_K)$ denote the corresponding $L \times K$ factor loading matrix where $\boldsymbol{c}_k = (c_{k1}, \ldots, c_{kL})'$. For identification,

Table 7
Factor loadings in a factor model (Factor 1).

Variable	Mean	Std	95% Interval	IF
c_{11}	**0.9714**	0.0241	(0.9093, 0.9981)	377
c_{21}	**0.9702**	0.0255	(0.9011, 0.9976)	499
c_{31}	**0.9568**	0.0382	(0.8531, 0.9977)	622
c_{41}	**0.9686**	0.0290	(0.8940, 0.9982)	601
c_{51}	**0.9190**	0.0571	(0.7827, 0.9892)	524
c_{61}	0.6233	0.1300	(0.3409, 0.8280)	629
c_{71}	0.5831	0.1185	(0.3119, 0.7720)	631
c_{81}	0.5679	0.1305	(0.2865, 0.8030)	610

Table 8
Factor loadings in a factor model (Factor 2).

Variable	Mean	Std	95% Interval	IF
c_{22}	0.1181	0.0970	(0.0041, 0.3461)	562
c_{32}	0.2025	0.1397	(-0.0577, 0.5135)	620
c_{42}	0.1314	0.1472	(-0.1184, 0.4342)	611
c_{52}	0.1957	0.1796	(-0.1301, 0.5387)	562
c_{62}	**0.7303**	0.1129	(0.5091, 0.9258)	626
c_{72}	**0.7754**	0.0899	(0.5885, 0.9373)	617
c_{82}	**0.7701**	0.0984	(0.5471, 0.9315)	594

The MCMC algorithm is iterated 200,000 times after discarding 20,000 samples as burn-in periods.

we assume

$$\begin{aligned} &c_{ll} > 0, \quad l = 1, \dots, L, \\ &c_{l,l+1} = \dots = c_{lL} = 0, \quad l = 1, \dots, L-1. \end{aligned}$$

Then the factor model for the $K \times 1$ latent vector $\boldsymbol{x}_i^*$ is given by

$$\begin{aligned} \boldsymbol{x}_i^* &= \mathbf{B}'\boldsymbol{v}_i + \mathbf{C}'\boldsymbol{f}_i + \boldsymbol{\Omega}^{1/2}\boldsymbol{u}_i, \quad \boldsymbol{f}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_L), \quad \boldsymbol{u}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_K), \\ &\boldsymbol{\Omega} = \mathrm{diag}\big(1 - \mathbf{c}_1'\mathbf{c}_1, \dots, 1 - \mathbf{c}_K'\mathbf{c}_K\big), \\ &\mathbf{c}_k'\mathbf{c}_k < 1, \quad k = 1, \dots, K, \end{aligned}$$

where $\boldsymbol{u}_i = (u_{i1}, \dots, u_{iK})'$ and

$$\mathbf{C}' = \begin{pmatrix} \mathbf{c}_1' \\ \vdots \\ \mathbf{c}_K' \end{pmatrix} = \begin{pmatrix} c_{11} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ c_{L-1,1} & \cdots & c_{L-1,L-1} & 0 \\ c_{L1} & \cdots & c_{L,L-1} & c_{LL} \\ \vdots & & \vdots & \vdots \\ c_{K1} & \cdots & c_{K,L-1} & c_{KL} \end{pmatrix}.$$

Note that the marginal model over $\boldsymbol{f}_i$ is given by

$$\boldsymbol{x}_i^* = \mathbf{B}'\boldsymbol{v}_i + \boldsymbol{\epsilon}_i, \quad \boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \quad \mathbf{R} = \mathbf{C}'\mathbf{C} + \boldsymbol{\Omega},$$

and

$$|\mathbf{R}| = \prod_{k=1}^{K}(1 - \mathbf{c}_k'\mathbf{c}_k) \times \left| \mathbf{I} + \sum_{k=1}^{K} \frac{1}{1 - \mathbf{c}_k'\mathbf{c}_k} \mathbf{c}_k\mathbf{c}_k' \right|.$$

The MCMC algorithm for the factor model is described in detail in Appendix B. For the drug administration survey data, we fit the $L = 2$ factor model and the estimates of the factor loadings are given in Tables 7 and 8. The factor loadings for the first five questions are large for the first factor, while those for the last three questions are large for the second factor. These results are consistent with those obtained for the varimax rotated factor loadings in Section 3.4 and support the evidence of dependence among the eight sensitive questions through the factor structure.

The results in Sections 3.3 and 3.4 are mostly the same using the factor model above and the estimates in Tables 7 and 8. While the factor structure is yet to be explained from the behavioral point of view, the proposed multivariate randomized response model and its Bayesian inference offer a generic approach to analyze responses collected using the RRT.

4.3. Summary of findings with policy implications

In this paper, we propose a multivariate probit regression model to analyze multiple randomized responses for binary data. We employed a drug administration survey with eight sensitive questions and close to 1000 respondents in a hospital cluster in Hong Kong.

There are numerous potential applications of the proposed multivariate probit model for risk management based on surveys. These include research in cybersecurity, educational research in handling plagiarism and school bullying, and studies for understanding behaviors in sensitive social and environmental issues. In the application we presented here, we have illustrated that drug administration error is affected by demographic and organizational factors. Drug administration errors are highly correlated, with the five ‘wrongs’ (wrong patient, wrong medication, wrong dose, wrong time, and wrong route) defining one dimension and the three omission of administration process steps (Q6, Q7, and Q8) defining a second dimension.

In terms of policy implications, the identification of gaps in hospitalization services is a useful starting point for improvements and such signals should not be ignored. However, it is equally important to realize that unsatisfactory performance is a manifestation, rather than a cause of the problem. An organization must investigate and look into the underlying reasons for unsatisfactory performance before effective resolutions of the issues and real improvements can be implemented. Staff performance is often affected by many contextual factors such as training, staffing, workflow, structures, regulation, and leadership. It could be useful for an organization to include the RRT as a method for the identification of gaps and then it can investigate into the underlying issues to enable improvements. This paper presents a comprehensive drug administration survey example that researchers and organizations can follow in modeling multiple sensitive responses.

Acknowledgments

The work described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (UGC/FDS14/P05/16). This work was also partially supported by JSPS KAKENHI Grant Number JP19H00588.

Appendix A. Posterior inference for R and B

A1. Prior distribution of $\mathbf{R}$

Suppose a $K \times K$ matrix $\boldsymbol{\Sigma} \sim \mathcal{IW}(\nu, \mathbf{I}_K)$ as in Barnard et al. (2000).

$$f(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{\nu+K+1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\right)\right\}. \tag{14}$$

Consider the transformation from $\boldsymbol{\Sigma}$ to $(\mathbf{R}, \boldsymbol{\Lambda})$ where $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\mathbf{R}\boldsymbol{\Lambda}$. The Jacobian is $2^K \prod_{k=1}^K \lambda_k^K$ and

$$f(\mathbf{R}, \boldsymbol{\Lambda}) \propto |\mathbf{R}|^{-\frac{\nu+K+1}{2}} \prod_{k=1}^{K} \lambda_k^{-(\nu+1)} \exp\left(-\frac{r^{kk}}{2\lambda_k^2}\right), \tag{15}$$

where r^{kk} is the (k,k)-th element of $\mathbf{R}^{-1}$. Given $\mathbf{R}$, λ_k^2’s are conditionally independent and the probability density function of λ_k^2 is given by

$$f(\lambda_k^2|\mathbf{R}) \propto \left(\lambda_k^2\right)^{-\left(\frac{\nu}{2}+1\right)} \exp\left(-\frac{r^{kk}}{2\lambda_k^2}\right),$$

and

$$\lambda_k^2|\mathbf{R} \sim \mathcal{IG}\left(\frac{\nu}{2}, \frac{r^{kk}}{2}\right), \quad k = 1, \dots, K.$$

The marginal probability density function of $\mathbf{R}$ is

$$f(\mathbf{R}) \propto |\mathbf{R}|^{-\frac{\nu+K+1}{2}} \prod_{k=1}^{K} \left(r^{kk}\right)^{-\frac{\nu}{2}} = |\mathbf{R}|^{\frac{(\nu-1)(K-1)}{2}-1} \prod_{k=1}^{K} |\mathbf{R}_{kk}|^{-\frac{\nu}{2}},$$

where $r^{kk} = |\mathbf{R}_{kk}|/|\mathbf{R}|$ and $\mathbf{R}_{kk}$ denotes a principal submatrix of $\mathbf{R}$ as shown also in Zhang et al. (2006). If we take $\nu = K + 1$, we obtain the marginally uniform prior density

$$\pi(\mathbf{R}) \propto |\mathbf{R}|^{\frac{K(K-1)}{2}-1} \left(\prod_{k=1}^{K} |\mathbf{R}_{kk}|\right)^{-\frac{K+1}{2}}.$$

A2. Efficient generation of **R** *and* **B** *in the multivariate probit model*

To sample $\mathbf{R}$, we introduce the latent variables

$$\lambda_k^2|\mathbf{R} \sim \mathcal{IG}\left(\frac{n_0}{2}, \frac{r^{kk}}{2}\right), \quad k = 1, \ldots, K,$$

and the joint prior probability density of $\mathbf{R}$ and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_K)$ is

$$\pi(\mathbf{R}, \boldsymbol{\Lambda}) = \pi(\boldsymbol{\Lambda}|\mathbf{R})\pi(\mathbf{R}) \propto |\mathbf{R}|^{-\frac{n_0+K+1}{2}} \prod_{k=1}^{K} \lambda_k^{-(n_0+1)} \exp\left(-\frac{r^{kk}}{2\lambda_k^2}\right),$$

as in (15). We note that

$$\pi(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n_0+K+1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\right)\right\},$$

where $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\mathbf{R}\boldsymbol{\Lambda}$ as in (14).

As the conditional joint posterior probability density of $\mathbf{B}$, $\mathbf{R}$, $\boldsymbol{\Lambda}$ and $\mathbf{X}^*$ is

$$\begin{aligned}
&\pi(\mathbf{B}, \mathbf{R}, \boldsymbol{\Lambda}, \mathbf{X}^*|\boldsymbol{y}^*, \boldsymbol{d}, \mathbf{A}, \boldsymbol{z}) \\
&\quad \propto |\mathbf{R}|^{-\frac{n+p}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{x}_i^* - \mathbf{B}'\boldsymbol{v}_i)'\mathbf{R}^{-1}(\boldsymbol{x}_i^* - \mathbf{B}'\boldsymbol{v}_i) - \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Psi}_0^{-1}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\right)\right\} \times \pi(\mathbf{R}, \boldsymbol{\Lambda}) \\
&\quad \propto |\mathbf{R}|^{-\frac{n+p}{2}} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{\Psi}_0^{-1} + \mathbf{V}'\mathbf{V})\mathbf{B}\mathbf{R}^{-1}\mathbf{B}' - 2\mathbf{B}\mathbf{R}^{-1}(\mathbf{X}^{*\prime}\mathbf{V}) + \mathbf{R}^{-1}(\mathbf{X}^{*\prime}\mathbf{X}^*)\right\}\right] \times \pi(\mathbf{R}, \boldsymbol{\Lambda}),
\end{aligned}$$

we obtain

$$\pi(\tilde{\mathbf{B}}, \boldsymbol{\Sigma}, \tilde{\mathbf{X}}|\boldsymbol{y}^*, \boldsymbol{d}, \mathbf{A}, \boldsymbol{z}) \propto |\boldsymbol{\Sigma}|^{-\frac{p+n+n_0+K+1}{2}} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{\boldsymbol{\Psi}_1^{-1}(\tilde{\mathbf{B}} - \mathbf{M}_2)\boldsymbol{\Sigma}^{-1}(\tilde{\mathbf{B}} - \mathbf{M}_2)' + \boldsymbol{\Sigma}^{-1}\mathbf{S}\right\}\right],$$

where

$$\begin{aligned}
&\tilde{\mathbf{B}} = \mathbf{B}\boldsymbol{\Lambda}, \quad \boldsymbol{\Sigma} = \boldsymbol{\Lambda}\mathbf{R}\boldsymbol{\Lambda}, \quad \tilde{\mathbf{X}} = \mathbf{X}^*\boldsymbol{\Lambda}, \\
&\mathbf{M}_2 = \boldsymbol{\Psi}_1\mathbf{V}'\tilde{\mathbf{X}}, \quad \boldsymbol{\Psi}_1^{-1} = \boldsymbol{\Psi}_0^{-1} + \mathbf{V}'\mathbf{V}, \quad \mathbf{S} = \tilde{\mathbf{X}}'\tilde{\mathbf{X}} - \mathbf{M}_2'\boldsymbol{\Psi}_1^{-1}\mathbf{M}_2 + \mathbf{I}_K.
\end{aligned}$$

Thus, the conditional posterior distribution of $\tilde{\mathbf{B}}$ given $\boldsymbol{\Sigma}, \tilde{\mathbf{X}}$ is

$$\tilde{\mathbf{B}}|\boldsymbol{\Sigma}, \tilde{\mathbf{X}} \sim \mathcal{N}_{p,K}(\mathbf{M}_2, \boldsymbol{\Psi}_1 \otimes \boldsymbol{\Sigma}),$$

and the marginal posterior probability density of $\boldsymbol{\Sigma}$ given $\tilde{\mathbf{X}}$ is

$$\pi(\boldsymbol{\Sigma}|\tilde{\mathbf{X}}, \boldsymbol{y}^*, \boldsymbol{d}, \mathbf{A}, \boldsymbol{z}) \propto |\boldsymbol{\Sigma}|^{-\frac{n+n_0+K+1}{2}} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{\boldsymbol{\Sigma}^{-1}\mathbf{S}\right\}\right].$$

Thus, we generate $\mathbf{B}$ and $\mathbf{R}$ as follows.

1. Generate $\lambda_k^2|\mathbf{R} \sim \mathcal{IG}\left(\frac{n_0}{2}, \frac{r^{kk}}{2}\right)$ for $k = 1, \ldots, K$, and set $\boldsymbol{\Lambda} = \mathrm{diagonal}(\lambda_1, \ldots, \lambda_K)$.
2. Compute $\tilde{\mathbf{X}}$, $\mathbf{S}$, and $\mathbf{M}_2$.
3. Generate $\boldsymbol{\Sigma}|\tilde{\mathbf{X}}, \boldsymbol{y}^*, \boldsymbol{d}, \mathbf{A}, \boldsymbol{z} \sim \mathcal{IW}(n + n_0, \mathbf{S})$ and set $\boldsymbol{\Lambda} = \mathrm{diagonal}(\sigma_{11}^{1/2}, \ldots, \sigma_{KK}^{1/2})$ where σ_{ii} is the (i, i)-th element of $\boldsymbol{\Sigma}$.
4. Generate $\tilde{\mathbf{B}}|\boldsymbol{\Sigma}, \tilde{\mathbf{X}}, \boldsymbol{y}^*, \boldsymbol{d}, \mathbf{A}, \boldsymbol{z} \sim \mathcal{N}_{p,K}(\mathbf{M}_2, \boldsymbol{\Psi}_1 \otimes \boldsymbol{\Sigma})$.
5. Generate $\tilde{\mathbf{X}}|\tilde{\mathbf{B}}, \boldsymbol{\Sigma}, \boldsymbol{y}^*, \boldsymbol{d}, \mathbf{A}, \boldsymbol{z}$ as in Section 2.4.
6. Compute $\mathbf{B} = \tilde{\mathbf{B}}\boldsymbol{\Lambda}^{-1}$, $\mathbf{R} = \boldsymbol{\Lambda}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Lambda}^{-1}$ and $\mathbf{X}^* = \tilde{\mathbf{X}}\boldsymbol{\Lambda}^{-1}$.

Appendix B. The Bayesian sampling scheme for the factor structure

Assuming the same normal prior distributions for **A** and **B** as in (7) and the uniform prior distribution for **C** over the region defined in Section 4.2, the joint posterior density is given by

$$\begin{aligned}
&\pi(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{D}, \mathbf{A}, \mathbf{B}, \mathbf{C}|\mathbf{Z}) \\
&\quad \propto |\mathbf{R}|^{-n/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{(\mathbf{X}^* - \mathbf{V}\mathbf{B})\mathbf{R}^{-1}(\mathbf{X}^* - \mathbf{V}\mathbf{B})' + (\mathbf{Y}^* - \mathbf{V}\mathbf{A})(\mathbf{Y}^* - \mathbf{V}\mathbf{A})'\right\}\right] \\
&\qquad \times |\mathbf{R}|^{-p/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{\boldsymbol{\Psi}_0^{-1}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}' + \boldsymbol{\Phi}_0^{-1}\mathbf{A}\mathbf{A}'\right\}\right] \\
&\qquad \times \prod_{i=1}^{n}\prod_{k=1}^{K} \pi_i^{d_{ik}}(1 - \pi_i)^{1-d_{ik}}\left\{d_{ik}I(x_{ik}^* > 0) + (1 - d_{ik})I(y_{ik}^* > 0)\right\}^{z_{ik}}\left\{d_{ik}I(x_{ik}^* \le 0) + (1 - d_{ik})I(y_{ik}^* \le 0)\right\}^{1-z_{ik}}, \qquad (16)
\end{aligned}$$

where $\mathbf{R} = \mathbf{C}'\mathbf{C} + \mathbf{\Omega}$. The conditional posterior density of $(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{D})$ in (16) is the same as in Section 2.4 and we sample $(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{D})$ using $\mathbf{R} = \mathbf{C}'\mathbf{C} + \mathbf{\Omega}$. Similarly, we sample $\mathbf{A}$ as in Section 2.4. As the conditional posterior probability density of $\mathbf{B}$ is

$$\pi(\mathbf{B}|\mathbf{X}^*, \mathbf{Y}^*, \mathbf{D}, \mathbf{A}, \mathbf{C}, \mathbf{Z}) \propto \exp\left[-\frac{1}{2}\mathrm{tr}\left\{(\mathbf{X}^* - \mathbf{VB})\mathbf{R}^{-1}(\mathbf{X}^* - \mathbf{VB})' + \mathbf{\Psi}_0^{-1}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\right\}\right],$$

we generate

$$\mathbf{B}|\cdot \sim \mathcal{N}_{p,K}(\mathbf{M}_2, \mathbf{\Psi}_1 \otimes \mathbf{R}), \quad \mathbf{M}_2 = \mathbf{\Psi}_1\mathbf{V}'\mathbf{X}, \quad \mathbf{\Psi}_1^{-1} = \mathbf{\Psi}_0^{-1} + \mathbf{V}'\mathbf{V}, \quad \mathbf{R} = \mathbf{C}'\mathbf{C} + \mathbf{\Omega}.$$

Using the posterior density marginalized over $\mathbf{F}$ and $\mathbf{G}$, (16), we generate c_{kl} one at a time given the other element of $\mathbf{C}$, denoted by $\mathbf{C}_{-kl}$, and other parameters and latent variables. As

$$\begin{aligned} &0 < c_{ll} < \sqrt{1 - \sum_{m \neq l} c_{lm}^2}, \quad l = 1, \ldots, L, \\ &|c_{kl}| < \sqrt{1 - \sum_{m \neq l} c_{km}^2}, \quad l = 1, \ldots, \min(k-1, L), \quad k = 2, \ldots, K, \end{aligned}$$

we consider the transformation

$$\begin{aligned} w_{ll} &= \log \frac{c_{ll}}{\sqrt{1 - \sum_{m \neq l} c_{lm}^2} - c_{ll}}, \quad \left(c_{ll} = \sqrt{1 - \sum_{m \neq l} c_{lm}^2} \times \frac{\exp(w_{ll})}{\exp(w_{ll}) + 1}\right), \quad l = 1, \ldots, L, \\ w_{kl} &= \log \frac{\sqrt{1 - \sum_{m \neq l} c_{km}^2} + c_{kl}}{\sqrt{1 - \sum_{m \neq l} c_{km}^2} - c_{kl}}, \quad \left(c_{kl} = \sqrt{1 - \sum_{m \neq l} c_{km}^2} \times \frac{\exp(w_{kl}) - 1}{\exp(w_{kl}) + 1}\right), \\ &\qquad l = 1, \ldots, \min(k-1, L), \quad k = 2, \ldots, K, \end{aligned}$$

with Jacobian

$$\begin{aligned} \left|\frac{dc_{ll}}{dw_{ll}}\right| &= \sqrt{1 - \sum_{m \neq l} c_{lm}^2} \times \frac{\exp(w_{ll})}{\{1 + \exp(w_{ll})\}^2}, \quad l = 1, \ldots, L, \\ \left|\frac{dc_{kl}}{dw_{kl}}\right| &= 2\sqrt{1 - \sum_{m \neq l} c_{km}^2} \times \frac{\exp(w_{kl})}{\{1 + \exp(w_{kl})\}^2}, \quad l = 1, \ldots, \min(k-1, L), \quad k = 2, \ldots, K, \end{aligned}$$

and use the independent acceptance-rejection MH algorithm with the normal density proposal based on the second order Taylor expansion of the logarithm of the conditional posterior density around the mode. At the end of this step, we also compute and save $\mathbf{R}$.

Appendix C. Illustrative example using simulated data

This section illustrates our proposed algorithm using the simulated data. To simulate the data, we set $n = 2000$, $K = 5$, $p = 3$, and $\pi_i = 1/3$ for $i = 1, \ldots, 800$ and $2/3$ for $801, \ldots, 2000$. The latent variables $\boldsymbol{x}_i^*$ and $\boldsymbol{y}_i^*$ are generated using $\boldsymbol{\beta}_1' = \boldsymbol{\beta}_2' = \cdots = \boldsymbol{\beta}_5' = (1, 1, -1)$, $\boldsymbol{\alpha}_1' = \boldsymbol{\alpha}_2' = \cdots = \boldsymbol{\alpha}_5' = (1, -1, 1)$, and $\boldsymbol{v}_i = (1, v_{i1}, v_{i2})'$ where $v_{ij} \sim$ i.i.d. $\mathcal{N}(0, 1)$ for $i = 1, \ldots, 2000$ and $j = 1, 2$. For the correlation matrix of $\boldsymbol{x}_i^*$, we use the equi-correlation matrix $\mathbf{R} = \{\rho_{ij}\} = (1 - \rho)\mathbf{I}_5 + \rho\mathbf{1}_5\mathbf{1}_5'$ where $\mathbf{1}_5 = (1, 1, 1, 1, 1)'$. All correlations are assumed to be equal to ρ and, in our illustrative example, we set $\rho = 0.4$. The prior distributions for parameters are

$$\mathbf{A} \sim \mathcal{N}_{3,5}(\mathbf{O}_{3,5}, 4\mathbf{I}_{15}), \quad \mathbf{B}|\mathbf{R} \sim \mathcal{N}_{3,5}(\mathbf{O}_{3,5}, 4\mathbf{I}_3 \otimes \mathbf{R}), \quad \pi(\mathbf{R}) \propto |\mathbf{R}|^9 \left(\prod_{k=1}^{5} |\mathbf{R}_{kk}|\right)^{-3},$$

where we assume the marginally uniform prior distribution for $\mathbf{R}$. We iterated the MCMC sampling of the posterior distribution 55,000 times and discarded the initial 5000 variates as the burn-in period. The subsequent 50,000 values are retained for the posterior inference. Below we report estimation results for $\boldsymbol{\alpha}_1$, $\boldsymbol{\beta}_1$ and ρ_{i1} $(i = 2, 3, 4)$ since those for $\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}_i$ $(i = 2, \ldots, 5)$ and ρ_{ij}'s $(j = 2, 3, 4, 5)$ are quite similar.

Figure 4 shows the sample paths for $\boldsymbol{\alpha}_1$, $\boldsymbol{\beta}_1$ and ρ_{i1} $(i = 2, 3, 4)$. They look stable indicating that the chains mix very well. In Table 9, the summary statistics are given for $\boldsymbol{\alpha}_1$, $\boldsymbol{\beta}_1$ and ρ_{i1} $(i = 2, 3, 4)$. The posterior means are overall close to the true values taking account of the posterior standard deviations, and all true values are contained in the 95% credible intervals.

Finally, the estimated posterior densities are shown in Fig. 5, indicating that they are successful to capture the true values of the parameters.

Remark 2. When we have many high correlation coefficients in $\mathbf{R}$, we would need more MCMC iterations to obtain the convergence for the posterior inference.

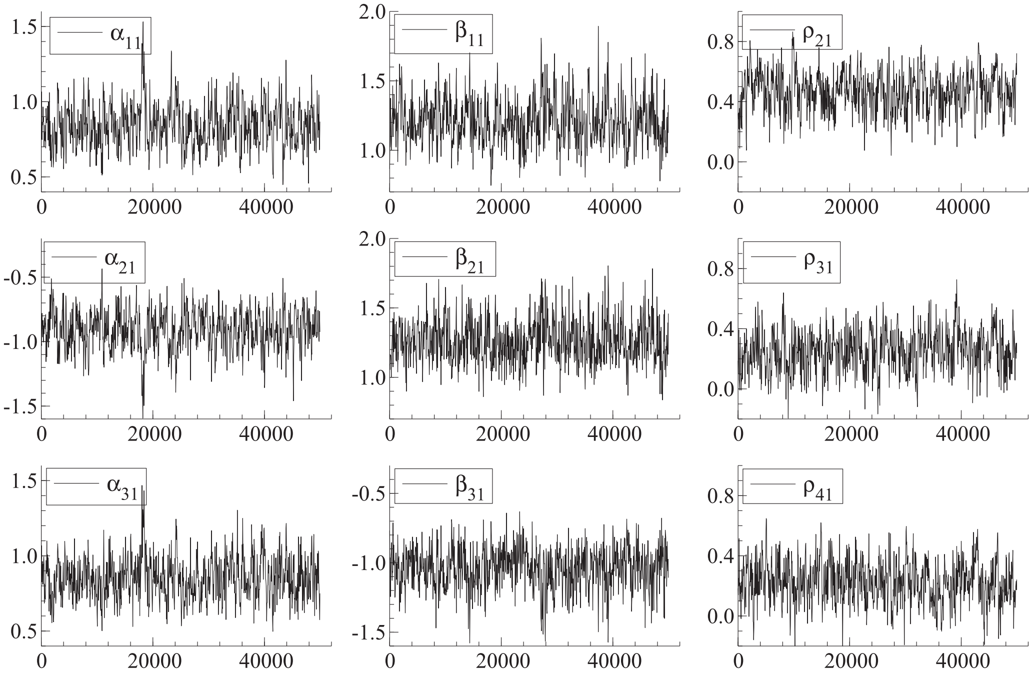


Fig. 4. Sample paths for $\boldsymbol{\alpha}_1$, $\boldsymbol{\beta}_1$ and ρ_{i1}'s. True values: $\boldsymbol{\alpha}_1 = (1, -1, 1)'$, $\boldsymbol{\beta}_1 = (1, 1, -1)'$ and $\rho_{i1} = 0.4$ ($i = 2, 3, 4$).

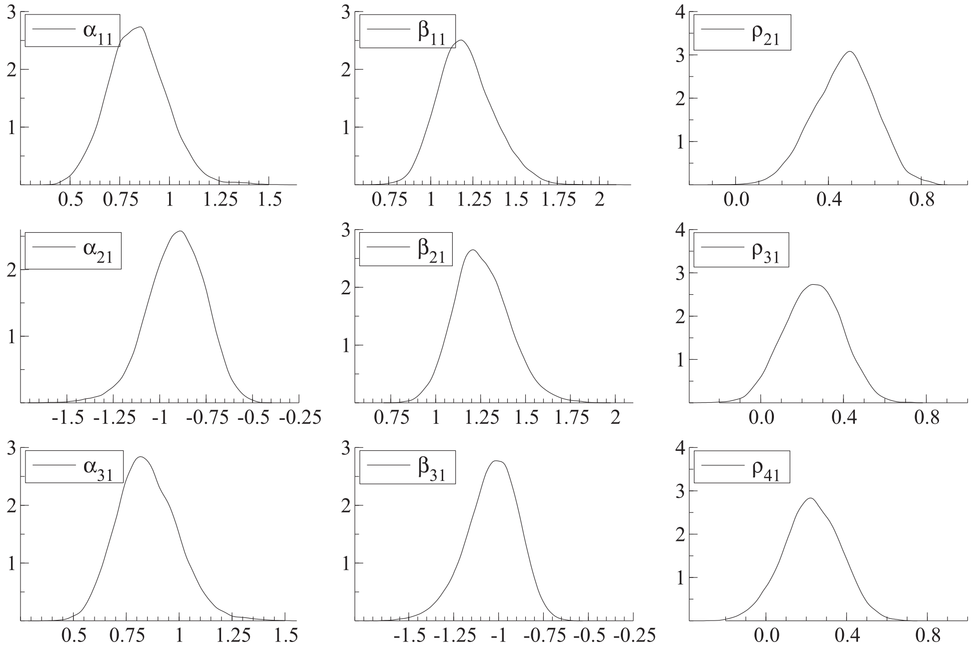


Fig. 5. Posterior densities for $\boldsymbol{\alpha}_1 = (\alpha_{11}, \alpha_{21}, \alpha_{31})'$, $\boldsymbol{\beta}_1 = (\beta_{11}, \beta_{21}, \beta_{31})'$ and ρ_{i1} ($i = 2, 3, 4$). True values: $\boldsymbol{\alpha}_1 = (1, -1, 1)'$, $\boldsymbol{\beta}_1 = (1, 1, -1)'$ and $\rho_{i1} = 0.4$ ($i = 2, 3, 4$).

Table 9
Summary statistics for the simulated data.

Variable	True	Mean	Std	95% Interval	IF
α_{11}	1	0.8400	0.1489	(0.5655, 1.1458)	151
α_{21}	-1	-0.9109	0.1555	(-1.2395,-0.6303)	143
α_{31}	1	0.8509	0.1440	(0.5934, 1.1538)	134
β_{11}	1	1.2082	0.1669	(0.9149, 1.5655)	125
β_{21}	1	1.2580	0.1538	(0.9820, 1.5861)	88
β_{31}	-1	-1.0356	0.1468	(-1.3567,-0.7786)	95
ρ_{21}	0.4	0.4691	0.1337	(0.1994, 0.7182)	137
ρ_{31}	0.4	0.2530	0.1395	(-0.0198, 0.5181)	123
ρ_{41}	0.4	0.2279	0.1418	(-0.0571, 0.4998)	122

References

Albert, J.H., Chib, S., 1993. Bayesian analysis of binary and polychotomous response data. Journal of the American Statistical Association 88 (422), 669–679.

Amemiya, T., 1981. Qualitative response models: A survey. Journal of Economic Literature 19 (4), 1483–1536.

Barcella, W., Iorio, M.D., Malone-Lee, J., 2018. Modelling correlated binary variables: an application to lower urinary tract symptoms. Journal of the Royal Statistical Society: Series C (Applied Statistics) 67 (4), 1083–1100.

Barnard, J., McCulloch, R., Meng, X.L., 2000. Modeling covariance matrices in terms of standard deviations and correlations with application to shrinkage. Statistica Sinica 10, 1281–1311.

Berrett, C., Calder, C.A., 2012. Data augmentation strategies for the bayesian spatial probit regression model. Computational Statistics & Data Analysis 56 (3), 478–490.

Blair, G., Imai, K., Zhou, Y.-Y., 2015. Design and analysis of the randomized response technique. Journal of the American Statistical Association 110 (511), 1304–1319.

Burkill, S., Copas, A., Couper, M.P., Clifton, S., Prah, P., Datta, J., Conrad, F., Wellings, K., Johnson, A.M., Erens, B., 2016. Using the web to collect data on sensitive behaviours: A study looking at mode effects on the British national survey of sexual attitudes and lifestyles. PLoS One 11 (2), e0147983.

Chib, S., Greenberg, E., 1998. Analysis of multivariate probit models. Biometrika 85, 347–361.

Chong, A.C.Y., Chu, A.M.Y., So, M.K.P., Chung, R.S.W., 2019. Asking sensitive questions using the randomized response approach in public health research: An empirical study on the factors of illegal waste disposal. International Journal of Environmental Research and Public Health 16 (6), 970.

Chu, A.M.Y., So, M.K.P., Chan, T.W.C., Tiwari, A., 2020. Estimating the dependence of mixed sensitive response types in randomized response technique. Statistical Methods in Medical Research 29 (3), 894–910.

Chu, A.M.Y., So, M.K.P., Chung, R.S.W., 2018. Applying the randomized response technique in business ethics research: The misuse of information systems resources in the workplace. Journal of Business Ethics 151 (1), 195–212.

Chung, R.S.W., Chu, A.M.Y., So, M.K.P., 2018. Bayesian randomized response technique with multiple sensitive attributes: The case of information systems resource misuse. The Annals of Applied Statistics 12 (3), 1969–1992.

Cross, R., Bennett, P.N., Ockerby, C., Wang, W.C., Currey, J., 2017. Nurses' attitudes toward the single checking of medications. Worldviews on Evidence-Based Nursing 14 (4), 274–281.

Doyle, P., 1977. The application of probit, logit, and tobit in marketing: A review. Journal of Business Research 5 (3), 235–248.

Durante, D., 2019. Conjugate Bayes for probit regression via unified skew-normal distributions. Biometrika 106 (4), 765–779.

Fasano, A., Rebaudo, G., Durante, D., Petrone, S., 2021. A closed-form filter for binary time series. Statistics and Computing 31 (47), 1–20.

Frühwirth-Schnatter, S., Frühwirth, R., 2007. Auxiliary mixture sampling with applications to logistic models. Computational Statistics & Data Analysis 51, 3509–3528.

Gibbons, R.D., Wilcox-Gök, V., 1998. Health service utilization and insurance coverage: A multivariate probit analysis. Journal of the American Statistical Association 93 (441), 63–72.

Greenberg, B.G., Abul-Ela, A.-L. A., Simmons, W.R., Horvitz, D.G., 1969. The unrelated question randomized response model: Theoretical framework. Journal of the American Statistical Association 64 (326), 520–539.

Greenberg, B.G., Kuebler Jr, R.R., Abernathy, J.R., Horvitz, D.G., 1971. Application of the randomized response technique in obtaining quantitative data. Journal of the American Statistical Association 66 (334), 243–250.

Holmes, C.C., Held, L., 2006. Bayesian auxiliary variable models for binary and multinomial regression. Bayesian Analysis 1, 145–168.

Imai, K., Van Dyk, D.A., 2005. A Bayesian analysis of the multinomial probit model using marginal data augmentation. Journal of Econometrics 124 (2), 311–334.

Keers, R.N., Williams, S.D., Cooke, J., Ashcroft, D.M., 2013. Causes of medication administration errors in hospitals: A systematic review of quantitative and qualitative evidence. Drug Safety 36 (11), 1045–1067.

Kim, J., Bates, D.W., 2013. Medication administration errors by nurses: Adherence to guidelines. Journal of Clinical Nursing 22 (3-4), 590–598.

Kwan, S.S.K., So, M.K.P., Tam, K.Y., 2010. Applying the randomized response technique to elicit truthful responses to sensitive questions in IS research: The case of software piracy behavior. Information Systems Research 21 (4), 941–959.

Laffont, C.M., Vandemeulebroecke, M., Concordet, D., 2014. Multivariate analysis of longitudinal ordinal data with mixed effects models, with application to clinical outcomes in osteoarthritis. Journal of the American Statistical Association 109 (507), 955–966.

Lee, S.-Y., Song, X.-Y., Cai, J.-H., 2010. A Bayesian approach for nonlinear structural equation models with dichotomous variables using logit and probit links. Structural Equation Modeling 17 (2), 280–302.

Liu, X., 2008. Parameter expansion for sampling a correlation matrix: An efficient GPX-RPMH algorithm. Journal of Statistical Computation and Simulation 78, 1065–1076.

Liu, X., Daniels, M., 2006. A new algorithm for simulating a correlation matrix based on parameter expansion and re-parameterization. Journal of Computational and Graphical Statistics 15, 897–914.

Llewellyn, R., Gordon, P., Wheatcroft, D., Lines, D., Reed, A., Butt, A., Lundgren, A., James, M., 2009. Drug administration errors: A prospective survey from three South African teaching hospitals. Anaesthesia & Intensive Care 37 (1), 93.

Manski, C.F., 1977. The structure of random utility models. Theory and Decision 8 (3), 229–254.

McFadden, D., 1980. Econometric models for probabilistic choice among products. Journal of Business S13–S29.

Muthén, B., 1979. A structural probit model with latent variables. Journal of the American Statistical Association 74 (368), 807–811.

O'brien, S.M., Dunson, D.B., 2004. Bayesian multivariate logistic regression. Biometrics 60 (3), 739–746.

Polson, N.G., Scott, J.G., Windle, J., 2013. Bayesian inference for logistic models using Pólya-Gamma latent variables. Journal of the American Statistical Association 108 (504), 1339–1349.

Qin, Q., Hobert, J.P., 2019. Convergence complexity analysis of Albert and Chib's algorithm for Bayesian probit regression. The Annals of Statistics 47 (4), 2320–2347.

Roy, V., Hobert, J.P., 2007. Convergence rates and asymptotic standard errors for Markov chain Monte Carlo algorithms for Bayesian probit regression. Journal of the Royal Statistical Society: Series B 69 (4), 607–623.

Sheu, S.-J., Wei, I.-L., Chen, C.-H., Yu, S., Tang, F.-I., 2009. Using snowball sampling method with nurses to understand medication administration errors. Journal of Clinical Nursing 18 (4), 559–569.

Song, X.-Y., Lee, S.-Y., 2005. A multivariate probit latent variable model for analyzing dichotomous responses. Statistica Sinica 645–664.

Talhouk, A., Doucet, A., Murphy, K., 2012. Efficient Bayesian inference for multivariate probit models with sparse inverse correlation matrices. Journal of Computational and Graphical Statistics 21, 739–757.

Warner, S.L., 1965. Randomized response: A survey technique for eliminating evasive answer bias. Journal of the American Statistical Association 60 (309), 63–69.

Zhang, X., Boscardin, W.J., Belin, T.R., 2006. Sampling correlation matrices in Bayesian models with correlated latent variables. Journal of Computational and Graphical Statistics 15, 880–896.