Modeling local dependence in latent vector autoregressive models

Biostatistics (2019), **0**, 0, *pp.* 1–28
doi:10.1093/biostatistics/Article

Modeling local dependence in latent vector autoregressive models

Trung Dung Tran[1,*], Emmanuel Lesaffre[1], Geert Verbeke[1], and Joke Duyck[2]

[1]Department of Public Health and Primary Care, KU Leuven, 3000 Leuven, Belgium

[2]Department of OH Sciences, KU Leuven, 3000 Leuven, Belgium

trungdung.tran@kuleuven.be

Summary

We propose a Bayesian latent vector autoregressive (LVAR) model to analyze multivariate longitudinal data of binary and ordinal variables (items) as a function of a small number of continuous latent variables. We focus on the evolution of the latent variables while taking into account the correlation structure of the responses. Often local independence is assumed in this context. Local independence implies that, given the latent variables, the responses are assumed mutually independent cross-sectionally and longitudinally. However, in practice conditioning on the latent variables may not remove the dependence of the responses. We address local dependence by further conditioning on item-specific random effects. A simulation study shows that wrongly assuming local independence may give biased estimates for the regression coefficients of the latent vector autoregressive process as well as the item-specific parameters. Novel features of our proposal include (i) correcting biased estimates of the model parameters, especially the regression coefficients of the latent vector autoregressive process, obtained when local dependence is ignored

*To whom correspondence should be addressed.

and (ii) measuring the magnitude of local dependence. We applied our model on data obtained from a registry on the elderly population in Belgium. The purpose was to examine the values of oral health information on top of general health information.

Key words: Bayesian approach; Local dependence; Local independence; Oral health and general health; Vector autoregressive models.

1. INTRODUCTION

Various applications in medicine, sociology, psychology, etc. require analyzing multivariate longitudinal data. In these applications, a set of subjects is repeatedly measured over time and the subject's condition is expressed by a number of correlated variables (items). Frequently, these items are considered as manifestations of one or more underlying latent characteristic. Hence, to measure the change in a subject's condition, it might be useful to work with the lower dimensional latent scores, rather than to work with the original items. However, any proposed model for the change in the latent level should properly account for the complex correlation structure induced by repeatedly collecting multiple observed responses. This paper focuses on the evolution of the latent characteristics but simultaneously take into account the correlation structure of the observed responses.

Our motivating data set consists of oral health (OH) and general health (GH) indicators collected in the BelRAI registry on elderly people in Belgium. These indicators were obtained with a questionnaire that was planned every six months. The indicators are assumed to be the manifestations of the so-called latent OH and GH status, respectively. In this study, the clinical researchers were interested to know whether OH status can additionally predict GH status given the current GH status and vice-versa. The research question was triggered because many caregivers are not convinced of the value of OH information on top of GH information. In

addition, since in the future we are looking for a better OH questionnaire, it is of interest to see how much the new OH indicators will improve prediction of the medical condition of the subjects over the current OH indicators.

The above research questions suggest to use an item response theory (IRT) model (see e.g. De Ayala, 2009; Reckase, 2009) to link the binary and ordinal responses over time to the latent variables (factors) and a bivariate longitudinal model for the evolution of these variables. Since our interest is placed on the cross-lagged effects of the latent factors, a bivariate autoregressive process is argued to be appropriate. Borrowing ideas from time series literature on dynamic factor analysis models, Zhang and Nesselroade (2007), Hutton and Chow (2014), and Cui and Dunson (2014) proposed to use the vector autoregressive (VAR) process to model the evolution of the latent factors. However, these proposal make use of a standard assumption in IRT models, that is local independence, which may be too simplistic (Andrich, 2017).

For a LVAR model, this local independence assumption implies independence of (i) the responses of different items at the same time point, (ii) the responses of the same item over time, (iii) and the responses of different items at different time points given the latent variables. In other words, it is assumed that the latent factors account for all three sources of correlations: (i) cross-sectional correlations: correlations between the responses at a particular time point, (ii) auto-correlations: correlations between repeated measurements of the same response over time, and (iii) cross-lagged correlations: correlations between different responses at different time points.

When local independence does not hold, we speak of local dependence. Statistical models that do not properly take the local dependence into account may yield biased estimates, leading to possibly misleading inference (Andrich, 2017). Although the focus is on the evolution of the latent variables, adequate modeling of the associations of the responses over time by taking into account local dependence is necessary (Cagnone *and others*, 2009).

The recent literature dealing with local dependence falls into two following approaches: (i) adding random effects or (ii) modeling the conditional probability. Liu and Hedeker (2006) adds shared random effects for all items while de Jong *and others* (2007), Proust-Lima *and others* (2013), Verhagen and Fox (2013), and Fox *and others* (2017) incorporate item-specific random effects. Cagnone *and others* (2009) and Wang *and others* (2013) also include random effects when making use of the univariate autoregressive process for a single latent variable. For the latter approach, Olsbjerg and Christensen (2015) allow the conditional probability of a response of an item at a particular time point to depend on the response of that item at the previous time point. Alternatively, Andrich (2017) includes an extra parameter, so-called the magnitude of the dependence, to model the conditional probability. Unfortunately, these models are limited to the case of a single latent factor. Little effort has been made to deal with local dependence when vector autoregressive process for multiple latent variables is used. The generalized latent trait model, proposed by Dunson (2003), can be used for multiple latent factors. However, this proposal encounters substantial computational challenges with a poorly behaving sampler due to high posterior dependence in the model parameters (Ghosh and Dunson, 2009; Cui and Dunson, 2014).

In this paper, we propose a LVAR model that combines an IRT model for the observed responses and a VAR(1) process for the latent factors. We address the evolutions of the latent variables while taking into account local dependence. Our model allows for multiple latent factors. This proposal corrects for biased estimation of the regression coefficients of the VAR(1) process and item-specific parameters. In addition, we introduce a quantity to measure the magnitude of local dependence, ranging from 0 to 1 with a high value indicating local dependence. Novel aspects of our proposal include (i) an extension of Cui and Dunson (2014) by the introduction of subject-specific random effects to model local dependence, (ii) a generalization of Cagnone *and others* (2009) by assuming more than one latent variable, (iii) and an introduction of the

magnitude of local dependence, generalizing the magnitude of dependence proposed by Andrich (2017) for a very limited case: the Rasch model (one-parameter IRT model) with a single latent factor and only two time points.

The article proceeds as follows. In Section 2, we present our proposed model along with the identification, prior specification, and estimation procedure. We provide in Section 3 an extensive simulation study to investigate the performance of our proposal. We also compare our proposal with a model assuming local independence. The usefulness of our proposal is illustrated in Section 4 where we model the development of OH and GH indicators recorded in the BelRAI database. The paper concludes with some discussions in Section 5.

2. Proposed models

2.1 *Model specification*

Suppose that K items are repeatedly recorded on N individuals over time. Let Y_{ijk} be the response for the k^{th} item of the i^{th} individual at the j^{th} time point where $i = 1, ..., N$, $j = 1, ..., n_i$, $k = 1, ..., K$ with n_i ($n_i \leqslant T$) is the number of occasions for individual i and T is the largest number of occasions. In this paper, we assume that for each subject the measurements are taken on a subset of the set of T equidistant time points: $\{t_1, ..., t_T\}$. We then assume that the observed items are manifestations of R ($2 \leqslant R$) latent factors. Denote $\boldsymbol{\xi}_{ij} = (\xi_{ij1}, ..., \xi_{ijr}, ..., \xi_{ijR})^T$ the $R \times 1$ vector of latent factors for individual i at time point t_j.

A popular model, linking the responses to the latent variables, makes use of two-parameter IRT models for binary items (Fox, 2010) and polytomous IRT models for ordinal items (Ostini and Nering, 2005), namely,

$$h(P(Y_{ijk} \leqslant m)) = \theta_{km} + \boldsymbol{\beta}_k^T \boldsymbol{x}_{ij} + \boldsymbol{\lambda}_k^T \boldsymbol{\xi}_{ij}, \tag{2.1}$$

where $h(.)$ is a link function (typically a logit or probit function) and m ($0 \leqslant m \leqslant c_k - 2$) is some

score of item k with c_k is the number of categories. For a binary variable ($c_k = 2$ and $m = 0$), we model the probability being equal to 0 instead of 1 as typically done in the IRT literature. The parameters θ_{km} and $\boldsymbol{\lambda}_k$ are item-specific location (cut-point) and discrimination (factor loading) parameters, respectively. The cut-points $\{\theta_{km}\}$ are non-decreasing in m. The $K \times R$ matrix whose k^{th} row is $\boldsymbol{\lambda}_k^T$ is called the factor loading matrix and denoted by Λ. Furthermore, $\boldsymbol{\beta}_k$ is a $p \times 1$ vector of regression parameters and $\boldsymbol{x}_{ij}$ is a $p \times 1$ vector consisting of the values of p covariates for individual i at time point t_j.

We make use of a VAR(1) process to model the evolution of those latent factors:

$$\begin{aligned} \boldsymbol{\xi}_{i1} &\sim N(\boldsymbol{\mu}, \Omega), \\ \boldsymbol{\xi}_{ij} &\sim N(\boldsymbol{\mu} + \Gamma(\boldsymbol{\xi}_{i,j-1} - \boldsymbol{\mu}), \Omega - \Gamma\Omega\Gamma^T), \qquad \forall j \geqslant 2, \end{aligned} \tag{2.2}$$

where $\boldsymbol{\mu}$ is a $R \times 1$ vector and Γ and Ω are $R \times R$ matrices. The diagonal and off-diagonal entries of Γ are referred to as autoregressive and cross-lagged parameters, respectively. We here assume that the latent factors satisfy the Markovian property.

The combination of the level 1 and level 2 model yields a LVAR model. This model assumes local independence, i.e.

$$\begin{aligned} P(Y_{ijk} \leqslant m|\boldsymbol{x}_{ij}, \boldsymbol{x}_{ij'}, \boldsymbol{\xi}_{ij}, \boldsymbol{\xi}_{ij'}) P(Y_{ij'k} \leqslant m^{'}|\boldsymbol{x}_{ij}, \boldsymbol{x}_{ij'}, \boldsymbol{\xi}_{ij}, \boldsymbol{\xi}_{ij'}) \\ = P(Y_{ijk} \leqslant m, Y_{ij'k} \leqslant m^{'}|\boldsymbol{x}_{ij}, \boldsymbol{x}_{ij'}, \boldsymbol{\xi}_{ij}, \boldsymbol{\xi}_{ij'}) \end{aligned}.$$

The model specified by (2.1) and (2.2) is similar to the proposal by Cui and Dunson (2014) where the authors assume local independence. However, often local independence does not hold as we also encountered in our motivating data set (Section 4). We will show later that ignoring local dependence can give substantially biased estimates, especially for the regression coefficients of the VAR(1) process. Therefore, we need to address local dependence. We propose to extend model (2.1) with a random intercept, i.e. the model we propose is now

$$h(P(Y_{ijk} \leqslant m)) = \theta_{km} + \boldsymbol{\beta}_k^T \boldsymbol{x}_{ij} + \boldsymbol{\lambda}_k^T \boldsymbol{\xi}_{ij} + b_{ik}, \tag{2.3}$$

where b_{ik} is the random effect for item k of individual i, assumed to be normally distributed $N(0, \sigma_{bk}^2)$. Our proposal generalizes the approach by Cagnone *and others* (2009). More specifically,

Cagnone *and others* (2009) limits to the case that $R = 1$. Our proposal does not have that limitation.

The random effects and the latent factors jointly account for the longitudinal association of the responses of the same item over time. More specifically, the variability of the left hand side of (2.3) is explained by the latent factors and the random effects with the proportions

$$\frac{\boldsymbol{\lambda}_k^T Cov(\boldsymbol{\xi}_{ij})\boldsymbol{\lambda}_k}{\boldsymbol{\lambda}_k^T Cov(\boldsymbol{\xi}_{ij})\boldsymbol{\lambda}_k + \sigma_{bk}^2} \quad \text{and} \quad \frac{\sigma_{bk}^2}{\boldsymbol{\lambda}_k^T Cov(\boldsymbol{\xi}_{ij})\boldsymbol{\lambda}_k + \sigma_{bk}^2}, \tag{2.4}$$

respectively, where $Cov(\boldsymbol{\xi}_{ij})$ is the marginal covariance matrix of $\boldsymbol{\xi}_{ij}$. Following (2.2), we see that $Cov(\boldsymbol{\xi}_{ij}) = \Omega$. Indeed, $Cov(\boldsymbol{\xi}_{i1}) = \Omega$ and $Cov(\boldsymbol{\xi}_{ij}) = \Gamma Cov(\boldsymbol{\xi}_{ij-1})\Gamma^T + \Omega - \Gamma\Omega\Gamma^T = \Gamma\Omega\Gamma^T + \Omega - \Gamma\Omega\Gamma^T = \Omega$. Therefore (2.4) can be rewritten as

$$\frac{\boldsymbol{\lambda}_k^T \Omega\boldsymbol{\lambda}_k}{\boldsymbol{\lambda}_k^T \Omega\boldsymbol{\lambda}_k + \sigma_{bk}^2} \quad \text{and} \quad \frac{\sigma_{bk}^2}{\boldsymbol{\lambda}_k^T \Omega\boldsymbol{\lambda}_k + \sigma_{bk}^2}. \tag{2.5}$$

We call the second formula in (2.5) the magnitude of local dependence for item k, denoted by D_k. All D_k $(k = 1, ..., K)$ range from 0 to 1 with a high value indicating local dependence.

The magnitude of local dependence, D_k, is similar to the intraclass correlation which can be defined as

$$ICC_k = \frac{\sigma_{bk}^2}{\boldsymbol{\lambda}_k^T \Omega\boldsymbol{\lambda}_k + \sigma_{bk}^2 + \pi^2/3}.$$

This definition is obtained by assuming that underlying each binary/ordinal item is a continuous variable whose relative value in comparison with (a) given threshold(s) indicates the value of the binary/ordinal variable (Goldstein *and others*, 2002; Browne *and others*, 2005). In addition, the correlation between the items and the latent variables is calculated as

$$\frac{\boldsymbol{\lambda_k}^T \boldsymbol{\omega}_r}{\sqrt{\boldsymbol{\lambda_k}^T \Omega\boldsymbol{\lambda_k} + \sigma_{bk}^2 + \sigma_0^2}},$$

where σ_0^2 equals $\pi^2/3$ or 1 if the logit or probit link function is used, respectively (see Section A in the Supplementary Material). Note that when $R = 1$, $\sigma_{bk} = 0$, and the probit link is used,

this correlation coefficient reduces to the biserial correlation coefficient (e.g. Johnson and Albert, 1999).

The inclusion of random effects allows for inter-individual variability, i.e. for each item two individuals having the same latent factors and observed values of covariates may still score differently related to unmeasured characteristics. For this reason the random effects are referred to as item-specific response tendencies. By the incorporation of random effects, we now assume that this model adequately takes into account the correlation structure of the observed responses.

2.2 *Identification issues*

Without adding further restrictions, model (2.3) is non-identifiable because of indeterminacy between the factor loadings ($\boldsymbol{\lambda}_k$), location parameter (θ_{km}), and the latent factors ($\boldsymbol{\xi}_{ij}$). A necessary condition for identifying this model is that the latent factors must have a scale and an origin (Oort, 2001). Scales of the latent factors can be imposed by fixing one factor loading per factor at a non-zero value or by fixing R factor variances at one occasion at a non-zero value. Fixing the origins for the latent factors can be done by fixing R item-specific location parameters (one location parameter per factor) or fixing R factor means at one occasion (Oort, 2001). In this paper, we fix the R factor means at 0 and the R factor variances at 1 at the first occasion, i.e. $\boldsymbol{\mu}=\mathbf{0}$ and Ω is a correlation matrix. Under this assumption, (2.2) is reduced to

$$\begin{aligned} &\boldsymbol{\xi}_{i1} \sim N(\mathbf{0}, \Omega), \\ &\boldsymbol{\xi}_{ij} \sim N(\Gamma\boldsymbol{\xi}_{i,j-1}, \Omega - \Gamma\Omega\Gamma^T), \qquad \forall j \geqslant 2. \end{aligned} \tag{2.6}$$

In addition to these restrictions, we must further impose at least $R(R-1)$ independent restrictions on Λ and/or Ω (Jöreskog, 1969). This can be done by requiring that certain elements of Λ and/or Ω have fixed values (Jöreskog, 1969), typically zero, or using priors with zero mean and small variance (Muthén and Asparouhov, 2012). We consider both approaches later.

2.3 *Prior distributions*

To complete the Bayesian specification of the LVAR model, prior distributions have to be given to the parameters. An informative prior may be chosen when prior knowledge or historical data is available and relevant. For example, prior knowledge about the structure of the latent factors should be taken into account. When there is no information or no intention to use prior information into the analysis, a non-informative prior is used. We specify the level 1 and level 2 model parameters separately. For both types of parameters a variety of prior distributions have been selected, basically all of them are vague but should take the above model considerations into account.

The level 1 model parameters are the location parameters θ_{km}, the discrimination parameters $\boldsymbol{\lambda}_k$ and the regression coefficients $\boldsymbol{\beta}_k$. In the literature, a variety of priors has been suggested for θ_{k0} in the binary case, varying from independent normals (Johnson and Albert, 1999), univariate hierarchical priors (Bradlow *and others*, 1999) and a multivariate hierarchical prior together with the discrimination parameters (Fox, 2010). For ordinal items, the prior must take into account their natural order ($\theta_{km} \leqslant \theta_{km'}$ when $m < m^{'}$). Again a number of priors has been proposed to impose that constraint, see e.g. Albert and Chib (1993) and Albert and Chib (2001). In the latter paper, a transformation turns the cut-points into unconstrained parameters. For the discrimination parameters, Johnson and Albert (1999) proposed a truncated normal, while Bradlow *and others* (1999) suggested a univariate hierarchical prior and, as indicated above, Fox (2010) suggested a joint hierarchical prior. Finally, for the logistic regression parameters $\boldsymbol{\beta}_k$, vague normal priors can be given. However, these priors do not protect against separation problems in the data, see e.g. Albert and Anderson (1984), Lesaffre and Albert (1989). One can take a normal prior with a realistic value for the prior (Lesaffre and Lawson, 2012) or a Cauchy(0, 2.5) or Cauchy(0, 5) prior (Gelman *and others*, 2008).

In this paper, we have taken a hierarchical prior for the cut-points (taking into account the

ordering) and another hierarchical prior for the discrimination parameters, i.e.,

$$\begin{array}{ll} \theta_{km} \mid \mu_\theta, \sigma_\theta & \overset{\text{iid}}{\sim} N(\mu_\theta, \sigma_\theta^2), \quad k = 1, ..., K, m = 0, ..., c_k - 2, \text{ given the order constraint,} \\ \lambda_{kr} \mid \sigma_\lambda & \overset{\text{iid}}{\sim} half\text{-}N(1, \sigma_\lambda^2), \quad k = 1, ..., K, r = 1, ..., R, \end{array}$$

where $\mu_\theta, \sigma_\theta$, and σ_λ are hyperparameters. Hierarchical priors have the advantage that the parameters are connected as much as the data allow for. Hence, they ensure some stability in the estimation process. For the regression coefficients, a Cauchy(0, 5) prior was taken. Finally, a half-normal or a half-Cauchy distribution can be given to σ_{bk} $(1 \leqslant k \leqslant K)$ (Gelman, 2006). For more details on the chosen priors, we refer to Section 3.

The level 2 model parameters are Γ and Ω. Vague priors such as a normal distribution with large variance can be given for the autoregressive and cross-lagged parameters in Γ. When $R = 2$, Ω has only one parameter, namely ρ, which can take a uniform prior on $[-1, 1]$. For more details on the chosen priors, we refer to Section 3.

2.4 *Estimation*

Estimating the model parameters of the suggested LVAR model is quite challenging (see Section B of the Supplementary Material), and requires a Markov Chain Monte Carlo (MCMC) procedure. To deal with missing values, we assume that the missingness mechanism is missing at random (MAR). Hence, the likelihood was obtained by integrating out the missing values (Little and Rubin, 2002):

$$L(\boldsymbol{\eta}|Y_{obs}) = \int f(Y_{obs}, Y_{miss}|\boldsymbol{\eta}) dY_{mis}$$

where Y_{obs} and Y_{mis} are the observed and missing parts of the data, respectively, and $\boldsymbol{\eta}$ are the model parameters. From $L(\boldsymbol{\eta}|Y_{obs})$, we defined the posterior distribution by multiplying it with the priors.

Sampling from the posterior distribution was done using the software package Stan (Carpenter *and others*, 2017). To assess convergence, we checked the trace plots and used the Gelman-Rubin

diagnostic to ensure that the estimated potential scale reduction factor (Rhat) for all parameters was smaller than 1.1. The converged chains were run until the effective sample size was large enough for the parameters of interest (e.g. Monte Carlo error/posterior SD $\leqslant 0.05$). Finally, we used Watanabe's information criterion WAIC (Watanabe, 2010) for model comparison as suggested by Luo and Al-Harbi (2017).

3. Simulation study

To assess the performance of the proposed model, we performed six simulation studies. Simulation study 1 compares our proposed model assuming local dependence, i.e. level 1 is based on model (2.3), (referred as model DEP) and a model assuming local independence, i.e. level 1 is based on model (2.1), (referred as model INDEP) in four scenarios where the magnitude of local dependence is 0, 0.3, 0.6, and 0.9, respectively. Simulation study 2 investigates the effect of misspecification of the distribution for the subject-specific random effects. The third one examines the effect of ignoring important covariates and the fourth one investigate how the magnitude of local dependence can reflex misspecification with respect to the number of latent variables. We also examine the proposal with a small sample size in simulation study 5. Finally, we did a sensitivity analysis for the magnitude of local dependence under different priors.

The results from the simulation studies show that model DEP and model INDEP perform fairly similarly when local independence exits and model DEP outperforms model INDEP in case of local dependence. Specifically, model INDEP provides substantially biased estimates. Biased estimates are also observed when important covariates are ignored. Another result is that local dependence can be used to detect misspecification with respect to the number of latent variables. Finally, our proposed model can work with relatively small datasets and in this case, the magnitude of local dependence does not change much under different sets of priors. We here present the results for the simulation study 1, scenario 2, i.e. comparing model DEP and

INDEP where magnitude of local dependence is around 0.3. For the other simulation studies and scenarios, the readers are referred to Section C of the Supplementary Material.

The following setting was taken. We fixed the number of individuals at $N = 400$. For each individual, the number of repeated measurements was randomly sampled in $\{2, \ldots, 12\}$ (i.e. $2 \leqslant n_i \leqslant T = 12 \ \forall i$). The sampling procedure is set so that there are more individuals with fewer observations. The number of latent factors was fixed at two and the number of items at seven with three binary and four ordinal items representing the first and the second factor, respectively. Two elements of the factor loading matrix, Λ, are fixed at zero, i.e.,

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \\ \lambda_{31} & 0 \\ 0 & \lambda_{42} \\ \lambda_{51} & \lambda_{52} \\ \lambda_{61} & \lambda_{62} \\ \lambda_{71} & \lambda_{72} \end{pmatrix}. \tag{3.7}$$

We have taken two covariates: a binary covariate assuming 1 with probability 0.5, and a $N(0, 1)$ distributed covariate. With this setting, 100 simulated datasets are generated. From each complete data set, item values were put as missing using a missing at random mechanism where the probabilities of being missing depend on the covariates and the previous response of that item, starting at the second time point. The sampling procedure for n_i, the true values of the model parameters, and the parameters for coding missing values are given in Section C of the Supplementary Material.

Denote $\Gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}$, the following prior distributions are then specified:

$$\begin{aligned}
&\theta_{km} \mid \mu_\theta, \sigma_\theta && \overset{\text{iid}}{\sim} N(\mu_\theta, \sigma_\theta^2), k = 1, ..., K, m = 0, ..., c_k - 2, \text{ given the order constraint,} \\
&\mu_\theta && \sim Cauchy(0, 5), \\
&\sigma_\theta && \sim half{-}Cauchy(0, 5), \\
&\lambda_{kr} \mid \sigma_\lambda && \overset{\text{iid}}{\sim} half{-}N(1, \sigma_\lambda^2), \quad k = 1, ..., K, r = 1, ..., R, \\
&\sigma_\lambda && \sim half{-}Cauchy(0, 5), \\
&\boldsymbol{\beta}_k && \overset{\text{iid}}{\sim} Cauchy(0, 5), \quad k = 1, ..., K, \\
&\gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22} && \overset{\text{iid}}{\sim} N(0, 100), \\
&\rho && \sim Uniform(-1, 1), \\
&\sigma_{bk} && \overset{\text{iid}}{\sim} half{-}Cauchy(0, 5), \quad k = 1, ..., K.
\end{aligned}$$

For each model, three chains with 5 000 iterations were run and the last 2 500 iterations of each chain were retained for posterior summaries. Convergence was checked as explained above.

All fittings passed the convergence check. For each parameter, we computed the average relative bias ((estimated-true)/true) (RB), the mean squared error (MSE) of the estimate, and the coverage probability (CP), i.e. the proportion where the 95% credible interval (CI) covers the true value. The results are provided in Table 1.

The results show that the estimators from model DEP are unbiased. Except for γ_{11} and γ_{22} and some others, the average relative biases are negative in the model INDEP with values ranging around -0.33. This means that model INDEP shrinks these parameters towards 0 with about 0.33. For the bivariate AR process, model INDEP overestimates the autoregressive effects while it substantially underestimates the cross-lagged effects. The relatively poor performance of model INDEP compared to model DEP is also seen in the MSE, which is often higher for the INDEP model. Finally, the coverage of the 95% CI's for model INDEP is low, especially for the autoregressive and cross-lagged effects while around the nominal level for model DEP. Thus, ignoring local dependence can give substantially biased estimates. In particular, it downgrades the cross-lagged effects, which in various cases are of the main interest.

4. Application to BelRAI database

4.1 *BelRAI dataset*

From 2010 onwards, the Belgian National Institute for Health and Disability Insurance (NIHDI) funded initiatives to reduce the risk of institutionalization for older people, i.e keeping them longer at home by providing home care. The individuals who were aged 65 and older, frail, and at risk of institutionalization were referred to health care agencies by the physicians, social services, or nurses but the decision is upon the subject and/or their family members. Health care agencies delivered at home the intervention programs, which were classified into different types: case man-

agement, day care, night care, occupational therapy, psychological support, and others. During the delivery of interventions, the professional caregivers, such as nurses, occupational therapists, physiotherapists, or psychologists completed the BelRAI Home Care instrument (questionnaire), under the consent of their clients and/or family members. The BelRAI instrument represents the Belgian version of the interRAI instruments, which are comprehensive assessments that evaluate at regular intervals the physical, clinical, psychological and social condition of an elderly person. The questionnaire was filled in at baseline and every six months afterwards. This information can be used to draw up a personalized care plan. More details about the BelRAI database can be found in Almeida Mello *and others* (2016) and references therein.

Our study involves $N = 5\,420$ participants with 13 199 observations. We were interested in the relation between OH and GH status over time. Common manifestations of general health deterioration in older people include limited physical functioning, cognitive impairment, and depression (Ganguli, 2009) whereas common oral health issues are chewing difficulty, dryness of the mouth (Razak *and others*, 2014), and non-intact teeth (De Visschere *and others*, 2016). Therefore, three binary OH and four ordinal GH indicators mentioned above were internationally chosen with an intention that they can capture all relevant information to evaluate the OH and GH status (http://www.interrai.org/). The OH status was evaluated via three binary OH-related items: non-intact teeth (NT), chewing difficulty (CD), and dry mouth (DM). Four interRAI validated scales, evaluating the functional, cognitive, and mental condition of the subjects and their stability, represent the GH status: Activities of Daily Living (ADL), Cognitive Performance Scale (CPS), Depression Rating Scale (DRS), and Changes in Health, End-Stage Disease, Signs, and Symptoms Scale (CHESS). We collapsed the original classes (7, 7, 15, and 6 respectively) of the GH indicators into four and coded each GH indicator from 0 to 3 where lower values indicate a better health condition. These items are supposed to represent the true OH and GH status, but to reveal the true states probably a more extensive examination is required. Additionally,

we included age (at baseline), gender (70% is female), living status (53% living alone), and type of intervention (70% is case management) into the analysis. It is of interest to describe the development of OH and GH status over time and to assess the importance of the cross-lagged effects, i.e. the additional information that the current OH (resp. GH) status provides on the future GH (resp. OH) status (Figure 1).

4.2 *Results*

All models in this section take the same prior specification as specified in Section 3. Three chains with 5 000 iterations were run for each model and the last 2 500 iterations of each chain were retained for posterior summaries. WAIC was computed for model comparison. The implemented formula for WAIC computation is provided in Section D of the Supplementary Material. We performed a posterior predictive check (PPC) as a goodness-of-fit test with the following discrepancy function (van der Linden, 2016):

$$\chi(\boldsymbol{y}, \boldsymbol{\eta}) = \sum_{i=1}^{N} \sum_{j=1}^{n_i} \sum_{k=1}^{K} \frac{(Y_{ijk} - E(Y_{ijk}|\boldsymbol{\eta}))^2}{Var(Y_{ijk}|\boldsymbol{\eta})},$$

where $E(Y_{ijk}|\boldsymbol{\eta}) = \sum_{m=0}^{c_k-1} m \times P(Y_{ijk} = m|\boldsymbol{\eta})$, $Var(Y_{ijk}|\boldsymbol{\eta}) = \sum_{m=0}^{c_k-1} (m - E(Y_{ijk}|\boldsymbol{\eta}))^2 \times P(Y_{ijk} = m|\boldsymbol{\eta})$, and $\boldsymbol{\eta}$ is the set of all parameters. Using this discrepancy function, we computed the posterior predictive p-value (sometimes called Bayesian p-value), which is the proportion that the function, computed from the original data, is larger than its value, computed from the generated data, i.e.

$$\text{estimated PPP-value} = \frac{1}{K} \sum_{i=1}^{K} I[\chi(\boldsymbol{y}, \boldsymbol{\eta}^k) \geqslant \chi(\tilde{\boldsymbol{y}}^k, \boldsymbol{\eta}^k)]$$

where $\boldsymbol{\eta}^1, ..., \boldsymbol{\eta}^K$ is a converged Markov chain from $p(\boldsymbol{\eta} \mid \boldsymbol{y})$, $\tilde{\boldsymbol{y}}^k$ a replicated data generated from $p(\boldsymbol{y} \mid \boldsymbol{\eta}^k)$. For a well chosen model, PPP-value will be around 0.5.

We start with a model DEP where the factor loading matrix is specified as:

$$\Lambda = \begin{pmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \\ 0 & \lambda_{72} \end{pmatrix}.$$

Since the posterior predictive p-value (PPP-value) of this model is 0.45, we can continue with this model. For a PPP-value close to 0 or 1, this model can be extended by allowing the cross-loadings to be estimated. In this case, a prior distribution with mean zero and small variance should be used otherwise the model will be unidentified (Muthén and Asparouhov, 2012). As an illustration, we specified $\lambda_{kr} \sim N(0, 0.01)$ for the cross-loadings. The WAIC for model DEP with and without the cross-loadings is 73692 and 73693, respectively, with PPP-values of 0.46 and 0.45, respectively. As a result, no extension with regard to the cross-loadings is needed and the model DEP without the cross-loadings is chosen.

We also fitted the corresponding model INDEP to the data. Table 2 gives parameter estimates (95% credible interval) for the autoregressive and cross-lagged parameters as well as WAIC for model DEP and model INDEP (See Table A9 in the Supplementary Material for the other regression parameters). Since WAIC is much smaller for model DEP, we conclude that local dependence exists in the data structure. In fact, the magnitude of local dependence (95% CI) for the seven items are 0.84 (0.80, 0.89), 0.29 (0.18, 0.39), 0.61 (0.54, 0.68), 0.86 (0.83, 0.89), 0.70 (0.64, 0.75), 0.74 (0.69, 0.78), 0.68 (0.62, 0.74), respectively. As indicated by the simulation study, these values of magnitude of local dependence are high and local dependence should be taken into account. Hence, between these two models, model DEP should be used for statistical inference.

Table 3 provides estimates and 95% CI's for the correlation coefficients between the OH, GH indicators and the latent OH, GH status. Moderate and high correlation coefficients indicate that the observed indicators can capture the latent status. Specifically, chewing difficulty associates

the most with the latent OH status, while cognition and depression are strongly associated with the latent GH status. In addition, chewing difficulty is also highly correlated with the latent GH status.

The results from model DEP in Table 2 indicate that the current OH status provides additional information in predicting the future value of GH status given the current GH status. Moreover, higher values for current OH correspond to higher values of future GH. In other words, the importance of the cross-lagged effect means that having a poor OH status is additionally predictive of a poor GH status in the future.

The clinical consequence of the findings is that the caregivers should not ignore OH status when preparing personalized care plans. The cross-lagged effect from OH to GH in the BelRAI population suggest that the presence of OH problem can be considered as a symptom of GH problem in the future.

5. Discussion

We have proposed a Bayesian LVAR model utilizing a vector autoregressive process for the latent level while taking local dependence into consideration. The proposal is particularly useful for analyzing the evolution of the latent factors. Our contribution includes (i) extending the approach by Cui and Dunson (2014) where the authors assume local independence, (ii) generalizing the approach by Cagnone *and others* (2009) where that approach limits to the case of a single latent variable, and (iii) introducing a quantity to measure the magnitude of local dependence proposed by Andrich (2017) for a limited case. We applied this model to analyze the changes of the OH and GH status over time. The results offer a persuasive evidence to caregivers about the usefulness of OH information on top of GH information.

Several consequences can occur when ignoring the local dependence. Biased estimation when assuming local independence can result in a misleading inference. In our application, ignoring local

independence would have affected considerably the clinical conclusions made from the BelRAI data. In fact, compared to model DEP, the effects of OH on future GH in model INDEP is insignificant. Although the proposal by Andrich (2017) can correct for the biasedness, it limits to the Rasch model, i.e one-parameter IRT model, with a single latent factor and only two time points. Our proposal is applicable to a broader setting with multiple latent factors.

A potential application of the LVAR framework is to predict future item values for subjects included in the study. Wang *and others* (2017) propose a two-stage prediction procedure for a setting with one latent factor but that procedure can be extended to the case of multiple latent variables. First, one predicts the latent factors using, e.g. forecasting techniques from the time series literature (e.g. Brockwell and Davis, 2016). The second step is to predict the item values via the link from the latent factors. The precision of the prediction depends upon the quality of prediction from the two steps. Again, ignoring the local dependence, which affects the parameter estimates in both steps, might give wrong predicted values.

A noticeable feature of the proposed model is that it can serve as an indicator for checking the assumption that the observed responses well represent the latent variables. The magnitude of local dependence can indicate how well the observed items represent the latent variables because a high magnitude means a low proportion of the explained variability, contributed by the latent variables. For example, in the BelRAI application, although the observed indicators can capture the latent variables, high values of magnitude of local dependence suggest that the observed OH and GH indicators do not well represent the latent OH and GH status.

We argue that our model is useful in a more broadly sense than in the current example. Indeed, examples can be found in clinical trials of a similar type. For example, the effect of ceftriaxone, a treatment to slow the disease progression of amyotrophic lateral sclerosis (ALS) which is a neurodegenerative disease, on ten ALS functional rating scales was evaluated on three latent factors, bulbar, upper limb, and lower limb functions (Wang and Luo, 2017). In behavioral

science, Hutton and Chow (2014) examine the daily emotional status of individuals and study the individuals' regulation of positive affect and negative effect, each represented by a battery of items, comparing the laboratory versus the field settings. Our proposal can also be applied to the study described by Dunson (2003) where the neurobehavioral development in rodents exposed to the pesticide methoxychlor was examined, where the neurologic function and behavior are measured by a number of tests.

Although our focus was on two latent factors, it is relatively straightforward to generalize our model to the case of more than two factors. Most of the above specification remains the same except that some attention has to be paid on the VAR(1) process. Extensions for larger R require some notice on the prior distribution for the correlation matrix Ω. Another generalization, which is to relax the assumption of constant time difference, is the topic of our current research. In this case, the autoregressive and cross-lagged parameters need to be adjusted in order to reflect the effect of time length.

As suggested by Oort (2001), we keep the factor loadings of each item invariant over time in order to keep the meaning of that item constant. We assume invariant factor loadings because the interpretation becomes difficult if the meaning of the indicators changes over time. However, in other applications, this assumption can be relaxed by allowing the factor loadings of an item varying around an average value (e.g. de Jong *and others*, 2007; Fox, 2010).

Depending on the research question, a particular model can be chosen for the latent factors. In our paper when interest is placed on the cross-lagged effects, a bivariate autoregressive process is argued to be appropriate since the cross-lagged parameters directly address those effects. This choice, which is useful for testing the cross-lagged effects at the latent level, can therefore further serve as an exploratory tool supporting the establishment of the causal relationship between the latent factors (Eichler, 2013). When the effect of covariates on the latent factors or when the cluster nature of the latent variables is of interest, multivariate linear mixed models (e.g. Verbeke

and others, 2014; Wang and Luo, 2017) or multilevel models (e.g. de Jong and Steenkamp, 2010; Fox, 2010) may be more appropriate for the latent factors.

One assumption is that the design should be roughly balanced. However, for subjects, it does not require the observed repeated measurements are equidistant because MAR mechanism can deal with intermittent missingness. For unbalanced designs and/or informative missingness (responses and/or covariates), further extensions are needed. The former can be tackled by allowing the parameters of the AR process to depend on time distance and the latter can be tackled by simultaneously modeling the missingness mechanism.

We limited to the comparison between model DEP and model INDEP. However, other comparisons can be made. For example, local dependence can be extended so that the responses at the same time point can be correlated stronger than implied by the latent factors. To do so, we could assume a multivariate distribution for the random effects. However, this may create computational issues (Verbeke *and others*, 2014). In addition, this extension lacks the attractive feature of our proposal, that is dimension reduction. Another deviation that we are currently studying is to compare our proposal with the latent linear mixed model (Wang and Luo, 2017), which focuses on the mean structure of the latent variables. We will report the results soon.

6. SOFTWARE

Software in the form of R code, together with a simulated input data set and complete documentation is available on Github at https://github.com/tdt01/local_dependence

7. SUPPLEMENTARY MATERIAL

Supplementary Material is available online at http://biostatistics.oxfordjournals.org.

ACKNOWLEDGMENTS

The authors obtain financial support from the funding C24/15/034 of KU Leuven, Belgium. *Conflict of Interest*: None declared.

References

Albert, A. and Anderson, J.A. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika* **71**(1), 1–10.

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

Albert, James H. and Chib, Siddhartha. (2001). Sequential ordinal modeling with applications to survival data. *Biometrics* **57**(3), 829–836.

Almeida Mello, Johanna, Declercq, Anja, Ces, Sophie, Van Durme, Therese, Van Audenhove, Chantal and Macq, Jean. (2016). Exploring home care interventions for frail older people in Belgium: A comparative effectiveness study. *Journal of the American Geriatrics Society* **64**(11), 2251–2256.

Andrich, David. (2017). Controlling response dependence in the measurement of change using the rasch model. *Statistical Methods in Medical Research* **0**(0), 1–17.

Bradlow, Eric T., Wainer, Howard and Wang, Xiaohui. (1999). A Bayesian random effects model for testlets. *Psychometrika* **64**(2), 153–68.

Brockwell, P.J. and Davis, R.A. (2016). *Introduction to Time Series and Forecasting*, Springer Texts in Statistics. Springer International Publishing.

Browne, Wj, Subramanian, SV, Jones, K and Goldstein, H. (2005). Variance partitioning in multilevel logistic models that exhibit overdispersion. *Journal Of The Royal Statistical Society Series A-Statistics In Society* **168**, 599–613.

Cagnone, Silvia, Moustaki, Irini and Vasdekis, Vassilis. (2009). Latent variable models for multivariate longitudinal ordinal responses. *British Journal of Mathematical and Statistical Psychology* **62**(2), 401–415.

Carpenter, Bob, Gelman, Andrew, Hoffman, Matthew, Lee, Daniel, Goodrich, Ben, Betancourt, Michael, Brubaker, Marcus, Guo, Jiqiang, Li, Peter and Riddell, Allen. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software* **76**(1), 1–32.

Cui, Kai and Dunson, David B. (2014). Generalized dynamic factor models for mixed-measurement time series. *Journal of Computational and Graphical Statistics* **23**(1), 169–191.

De Ayala, R. J. (2009). *The Theory and Practice of Item Response Theory*, Methodology in the social sciences. New York: Guildord Press.

de Jong, MG and Steenkamp, Jbem. (2010). Finite mixture multilevel multidimensional ordinal irt models for large scale cross-cultural research. *Vox Sanguinis* **98**, 3–32.

de Jong, Martijn G., Steenkamp, Jan-Benedict E.M. and Fox, Gerardus J.A. (2007). Relaxing cross-national measurement invariance using a hierarchical irt model. *Journal of consumer research* **34**(2), 260–278.

De Visschere, Luc, Janssens, Barbara, De Reu, Griet, Duyck, Joke and Vanobbergen, Jacques. (2016, Nov). An oral health survey of vulnerable older people in belgium. *Clinical Oral Investigations* **20**(8), 1903–1912.

Dunson, David B. (2003). Dynamic latent trait models for multidimensional longitudinal data. *Journal of the American Statistical Association* **98**(463), 555–563.

Eichler, Michael. (2013). Causal inference with multiple time series: principles and problems. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **371**(1997).

Fox, Jean-Paul. (2010). *Bayesian Item Response Modeling: Theory and Applications*, Statistics for Social and Behavioral Sciences. New York: Springer.

Fox, Jean-Paul, Mulder, Joris and Sinharay, Sandip. (2017). Bayes factor covariance testing in item response models. *Psychometrika* **82**(4), 979–1006.

Ganguli, Mary. (2009). Depression, cognitive impairment and dementia: Why should clinicians care about the web of causation? *Indian journal of psychiatry* (51), 29–34.

Gelman, Andrew. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis* **1**(3), 515–534.

Gelman, Andrew, Jakulin, Aleks, Pittau, Maria Grazia and Su, Yu-Sung. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics* **2**(4), 1360–1383.

Ghosh, Joyee and Dunson, David B. (2009). Default prior distributions and efficient posterior computation in Bayesian factor analysis. *Journal of Computational and Graphical Statistics* **18**(2), 306–320.

Goldstein, Harvey, Browne, William and Rasbash, Jon. (2002). Partitioning variation in multilevel models. *Understanding Statistics* **1**(4), 223–231.

Hutton, R. Shane and Chow, Sy-Miin. (2014). Longitudinal multi-trait-state-method model using ordinal data. *Multivariate Behavioral Research* **49**(3), 269–282.

Johnson, Valen E and Albert, James H. (1999). *Ordinal Data Modeling*, Statistics for Social Science and Public Policy. Berlin: Springer.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* **34**(2), 183–202.

Lesaffre, E and Albert, A. (1989). Partial separation in logistic discrimination. *Journal of the Royal Statistical Society Series B-Methodological* **51**(1), 109–116.

Lesaffre, E and Lawson, A B. (2012). *Bayesian Biostatistics*, Statistics in Practice. Wiley.

Little, Roderick J. A and Rubin, Donald B. (2002). *Statistical analysis with missing data*, 2nd ed. edition., Wiley series in probability and statistics. London: Wiley.

Liu, Li C. and Hedeker, Donald. (2006). A mixed-effects regression model for longitudinal multivariate ordinal data. *Biometrics* **62**(1), 261–268.

Luo, Yong and Al-Harbi, Khaleel. (2017). Performances of loo and waic as irt model selection methods. *Psychological Test and Assessment Modeling* **59**(2), 183–205.

Muthén, Bengt O. and Asparouhov, Tihomir. (2012). Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychological methods* **17 3**, 313–35.

Olsbjerg, Maja and Christensen, Karl. (2015). Modeling local dependence in longitudinal IRT models. *Behavior Research Methods* **47**(4), 1413–1424.

Oort, FJ. (2001). Three-mode models for multivariate longitudinal data. *British Journal of Mathematical & Statistical Psychology* **54**, 49–78.

Ostini, Remo and Nering, Michael L. (2005). *Polytomous Item Response Theory Models*, Quantitative Applications in the Social Sciences 144. California: Sage.

Proust-Lima, Ccile, Amieva, Hlne and JacqminGadda, Hlne. (2013). Analysis of multivariate mixed longitudinal data: A flexible latent process approach. *British Journal of Mathematical and Statistical Psychology* **66**(3), 470–487.

Razak, P Abdul, Richard, K M Jose, Thankachan, Rekha P, Hafiz, K A Abdul, Kumar, K Nanda and Sameer, K M. (2014). Geriatric oral health: a review article. *Journal of international oral health : JIOH* **6**(6).

Reckase, M.D. (2009). *Multidimensional Item Response Theory*, Statistics for Social and Behavioral Sciences. New York: Springer.

van der Linden, W.J. (2016). *Handbook of Item Response Theory, Volume One: Models*, Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences. CRC Press.

Verbeke, Geert, Fieuws, Steffen, Molenberghs, Geert and Davidian, Marie. (2014). The analysis of multivariate longitudinal data: A review. *Statistical Methods in Medical Research* **23**(1), 42–59.

Verhagen, A. J. and Fox, J. P. (2013). Bayesian tests of measurement invariance. *British Journal of Mathematical and Statistical Psychology* **66**(3), 383–401.

Wang, Jue and Luo, Sheng. (2017). Multidimensional latent trait linear mixed model: an application in clinical studies with multivariate longitudinal outcomes. *Statistics in Medicine* **36**(20), 3244–3256.

Wang, Jue, Luo, Sheng and Li, Liang. (2017). Dynamic prediction for multiple repeated measures and event time data: an application to parkinson's disease. *The annals of applied statistics* **11 3**, 1787–1809.

Wang, Xiaojing, Berger, James O. and Burdick, Donald S. (2013). Bayesian analysis of dynamic item response models in educational testing. *Ann. Appl. Stat.* **7**(1), 126–153.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research* **11**, 3571–3594.

Zhang, Zhiyong and Nesselroade, John R. (2007). Bayesian estimation of categorical dynamic factor models. *Multivariate Behavioral Research* **42**(4), 729–756.

[*Received xxx xx, xxxx; revised xxx xx, xxxx; accepted for publication xxx xx, xxxx*]

Table 1. Simulation study 1, scenario 2: Average relative bias (RB), mean squared error (MSE), and coverage probability (CP) for the selected parameters

		Model DEP			Model INDEP		
Parameter	True value	RB	MSE	CP	RB	MSE	CP
θ_2	0.60	-0.053	0.113	96	-0.384	0.104	81
θ_3	0.90	-0.003	0.111	97	-0.341	0.154	73
θ_{41}	-3.00	0.003	0.076	98	-0.373	1.285	0
θ_{42}	-1.00	-0.022	0.066	94	-0.399	0.185	30
θ_{43}	2.70	0.031	0.079	95	-0.342	0.886	0
θ_{51}	-4.00	-0.039	0.292	95	-0.454	3.420	0
θ_{52}	-1.50	-0.076	0.218	94	-0.474	0.588	20
θ_{53}	2.60	0.032	0.187	96	-0.412	1.223	7
θ_{71}	-2.50	0.001	0.039	96	-0.232	0.360	2
θ_{72}	-1.00	-0.001	0.034	94	-0.247	0.082	48
θ_{73}	1.40	0.018	0.030	95	-0.200	0.097	41
λ_{11}	1.20	0.009	0.016	96	-0.193	0.065	32
λ_{12}	0.30	0.067	0.010	95	0.186	0.013	78
λ_{21}	4.00	0.031	0.233	97	-0.298	1.662	17
λ_{22}	0.50	0.101	0.044	99	-0.208	0.055	66
λ_{31}	4.10	0.009	0.235	94	-0.255	1.347	27
λ_{42}	3.10	0.016	0.047	98	-0.348	1.206	0
λ_{51}	0.60	0.116	0.065	95	-0.456	0.143	39
λ_{52}	5.20	-0.030	0.252	93	-0.392	4.484	2
λ_{61}	0.40	0.140	0.032	94	0.004	0.036	67
λ_{62}	3.00	0.003	0.050	94	-0.348	1.140	0
λ_{71}	0.30	0.100	0.017	93	0.093	0.019	69
λ_{72}	1.70	-0.006	0.014	96	-0.268	0.219	3
β_{11}	0.30	0.142	0.034	97	0.008	0.026	93
β_{12}	0.50	0.005	0.010	95	-0.124	0.012	86
β_{21}	0.10	0.476	0.233	95	0.025	0.114	92
β_{22}	0.20	0.138	0.074	95	-0.203	0.034	92
β_{31}	-0.10	-0.120	0.303	96	-0.338	0.146	91
β_{32}	0.20	0.030	0.057	98	-0.322	0.029	97
β_{41}	-0.20	-0.064	0.131	97	-0.453	0.062	87
β_{42}	0.40	-0.047	0.034	96	-0.402	0.041	61
β_{51}	0.30	0.311	0.350	95	-0.307	0.134	95
β_{52}	-0.30	0.055	0.128	93	-0.410	0.062	83
β_{61}	-0.10	-0.509	0.133	97	-0.643	0.068	92
β_{62}	-0.20	0.098	0.040	94	-0.248	0.020	92
β_{71}	-0.20	-0.032	0.047	96	-0.247	0.035	94
β_{72}	-0.10	0.339	0.015	95	0.057	0.009	92
σ_{b1}	0.80	-0.007	0.020	94	-	-	-
σ_{b2}	2.50	0.030	0.149	95	-	-	-
σ_{b3}	2.70	0.036	0.189	92	-	-	-
σ_{b4}	2.10	0.012	0.033	92	-	-	-
σ_{b5}	3.30	-0.009	0.089	95	-	-	-
σ_{b6}	1.90	0.002	0.020	98	-	-	-
σ_{b7}	1.30	-0.002	0.013	96	-	-	-
γ_{11}	0.90	-0.006	0.000	97	0.071	0.004	0
γ_{12}	0.10	0.029	0.000	98	-0.612	0.004	1
γ_{21}	0.12	0.055	0.000	95	-0.572	0.005	0
γ_{22}	0.85	-0.010	0.000	95	0.098	0.007	0
ρ	0.40	-0.045	0.003	96	-0.157	0.009	73

Table 2. BelRAI data: Posterior means (95% credible intervals) for auto-regressive, cross-lagged parameters, and WAIC for model DEP, which assumes the local dependence, and model INDEP, which assumes the local independence.

	model DEP			model INDEP		
	Mean	95% CI		Mean	95% CI	
γ_{11}	0.978	0.959	0.996	0.995	0.985	1.003
γ_{12}	-0.003	-0.033	0.027	-0.005	-0.031	0.024
γ_{21}	0.058	0.025	0.094	-0.005	-0.034	0.022
γ_{22}	0.878	0.843	0.909	0.964	0.947	0.980
WAIC	73693	72775	74611	114921	113935	115908

Table 3. Correlation coefficients between the OH, GH indicators and the OH, GH status

	OH			GH		
Item	Mean	95% CI		Mean	95% CI	
Non-intact teeth	0.39	0.33	0.45	0.22	0.18	0.26
Chewing difficulty	0.84	0.78	0.90	0.47	0.42	0.52
Dry mouth	0.62	0.56	0.67	0.35	0.30	0.40
ADL	0.20	0.17	0.23	0.35	0.31	0.38
CPS	0.30	0.27	0.34	0.54	0.49	0.58
DRS	0.27	0.24	0.31	0.48	0.44	0.52
CHESS	0.23	0.20	0.27	0.41	0.37	0.45

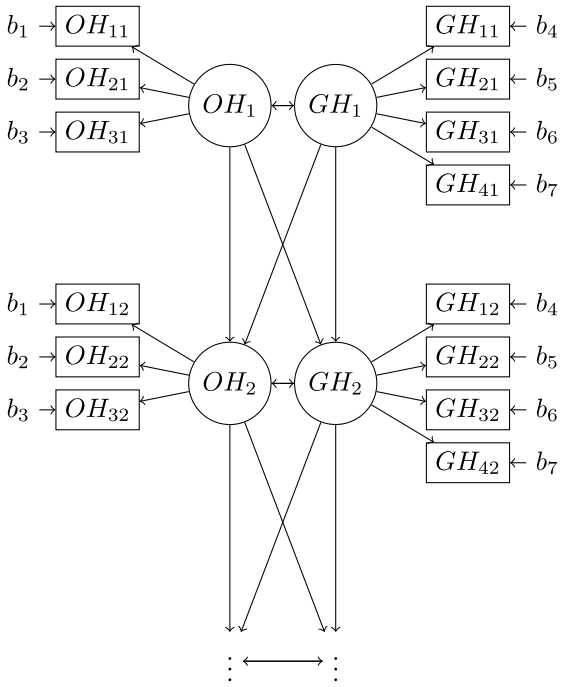

Fig. 1. The hypothesized relationship of the oral health (OH) and general health (GH) status. Three binary OH items and four ordinal GH items are selected to represent OH and GH status, respectively. The individual is measured at the first visit and then every 6 months afterwards. The random effects are added to take into account the local dependence. The index i for subject i is suppressed.