

Multiple data sources fusion for enterprise quality improvement by a multilevel latent response model

SHUAI HUANG¹, JING LI^{2,*}, GERRI LAMB³, MADELINE H. SCHMITT⁴ and JOHN FOWLER²

¹*Department of Industrial and Management Systems Engineering, University of South Florida, Tampa, FL 33620, USA*

²*Industrial Engineering, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ 85201, USA*

³*College of Nursing & Health Innovation, Arizona State University, Tempe, AZ 85201, USA*

E-mail: jinglz@asu.edu

⁴*School of Nursing, University of Rochester, Rochester, NY 14627, USA*

Received June 2012 and accepted May 2013

Quality improvement of an enterprise needs a model to link multiple data sources, including the independent and interdependent activities of individuals in the enterprise, enterprise infrastructure, climate, and administration strategies, as well as the quality outcomes of the enterprise. This is a challenging problem because the data are at two levels—i.e., the individual and enterprise levels—and each individual's contribution to the enterprise quality outcome is usually not explicitly known. These challenges make general regression analysis and conventional multilevel models non-applicable to the problem. This article a new multilevel model that treats each individual's contribution to the enterprise quality outcome as a latent variable. Under this new formulation, an algorithm is developed to estimate the model parameters, which integrates the Fisher scoring algorithm and generalized least squares estimation. Extensive simulation studies are performed that demonstrate the superiority of the proposed model over the competing approach in terms of the statistical properties in parameter estimation. The proposed model is applied to a real-world application of nursing quality improvement and helps identify key nursing activities and unit (a hospital unit is an enterprise in this context) quality-improving measures that help reduce patient falls.

Keywords: Multi-data fusion, multilevel model, latent variable model, quality, health care

1. Introduction

An enterprise consists of individuals. Examples of an enterprise include a company or a company division, such as the technology division and marketing division, consisting of employees; a military team consisting of soldiers on a specific mission; and a hospital or a hospital unit, such as the emergency unit and surgical unit, consisting of medical professionals. Quality improvement for an enterprise is different from quality improvement for an individual in terms of the following aspects.

1. The quality of an enterprise is affected by the independent and interdependent activities of individual employees.
2. The quality of an enterprise is also affected by factors at the enterprise level such as the infrastructure, culture, and administration strategies.

3. Each individual contributes a certain amount to the quality of the enterprise; however, this individual contribution is usually not explicitly known.

For example, consider a hospital unit as an enterprise. Although the quality of a hospital unit has many dimensions, here we solely focus on one dimension—the quality of nursing in patient care; this can be measured in terms of the number of falls and medication errors made by the unit during a certain time period and also patient satisfaction with the care provided by the unit's staff. Because our focus is on the quality of nursing, we consider the individuals associated with the unit to be nurses, although a hospital unit also includes other medical professionals. It is not difficult to see that the aforementioned three aspects in the quality improvement of an enterprise fit perfectly in the context of nursing.

1. The quality of nursing in a unit is affected by the independent nursing activities of individual nurses as well as their interdependent activities in coordinating a patient's care (Lamb *et al.*, 2008).

*Corresponding author

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/uiie.

2. The quality of nursing in a unit is also affected by unit infrastructure, climate, and administration strategies.
3. Each nurse contributes a certain amount to the nursing quality of the unit. However, this contribution is usually not known because delivery of patient care requires team work and it is difficult to explicitly apportion the quality outcome of the unit to each individual nurse.

For quality improvement of an enterprise, three data sources may be utilized and they exist at two levels. Specifically, at the enterprise level, there are two data sources on quality outcomes and quality-affecting factors, such as enterprise infrastructure, climate, and administration strategies. At the individual level, there is one data source on individual independent and interdependent activities. To accomplish the goal of quality improvement of an enterprise, a statistical model is needed to link these multiple data sources. The model should be able to address the questions of how individual activities affect the quality of the enterprise, how factors at the enterprise level affect the quality, and how an individual's activities/performance interact with the enterprise-level factors to jointly affect the enterprise quality. Answers to these questions are keys to formulating action plans for quality improvement of the enterprise.

To address the above questions, regression models provide a potential tool. Compared with many "black-box" methods that focus primarily on best prediction of the outcome variable, regression models are a "white-box" method offering better interpretability by explicitly revealing the variables that lead to the best prediction for the outcome. This is especially important when the final objective is to formulate an action plan. Furthermore, parametric regression models allow for statistical inference in variable or model selection, which offers a rigorous way for generalizing the model estimated from a specific data set to the general population. Third, multi-level regression, as a specific type of regression model, was developed to handle multilevel data sources.

In the "language" of regression, the quality outcomes are called *responses*; the enterprise-level quality-affecting factors and the individual-level activity variables can both be called *predictors*. In general regression analysis, all of the responses and predictors should be at the same level. Therefore, to be able to apply general regression analysis to our data, an intuitive approach is to transform the data sources to a single level. Specifically, since out of the three data sources only the predictors of individual activities are at the individual level, we might consider aggregating these predictors to the enterprise level by using some pre-defined summary statistics, such as the sample mean and variance of the data of all of the individuals in each enterprise. Then, general regression analysis can be applied at the enterprise level. One drawback of this approach is loss of information, since the pre-defined summary statistics may not be able to capture the full spectrum of behavior of the individuals in

each enterprise. Another more severe drawback is aggregation bias (Robinson, 1950; Goldstein, 1995; Draper and Smith, 1998); i.e., a predictor at different levels may have different effects on the response. For example, in the nursing context, there is an important predictor called *exchanging*, which is originally defined for each individual nurse (i.e., an individual-level predictor) to characterize activities of exchanging information with others regarding patient care. If *exchanging* were aggregated to the unit level, then the resulting variable—i.e., the average level of information exchange of a unit—is a proxy measure of the unit's normative environment/facility in promoting information exchange and thus may have an effect on the nursing quality of the unit different from the effect of an individual nurse's activities in exchanging information.

When predictors are at different levels, multilevel regression provides a more appropriate tool than general regression analysis. Multilevel regression has been discussed in a diverse range of literature under a variety of titles. For example, they are referred to as *multilevel regression* or *multilevel linear models* in sociological research (Mason *et al.*, 1983; Goldstein, 1995), as *mixed-effects models* in biometric research (Elston and Grizzle, 1962; Laird and Ware, 1982; Singer, 1998), as *random efficient models* in econometrics (Rosenberg, 1973; Longford, 1993), and as *covariance components models* in statistics (Dempster *et al.*, 1981). A multilevel regression model allows the inclusion of predictors at two (e.g., individual and enterprise) or more levels. The basic idea is to build a separate regression model for each enterprise by linking the individual-level predictors with the response and then modeling the variation among enterprises by considering the regression coefficients as multivariate responses explained by enterprise-level predictors. However, existing multilevel models are not directly applicable to our problem because they require that the response variable be at the individual level.

To the best of our knowledge, there is no effective model to link enterprise-level quality responses with enterprise- and individual-level predictors for quality improvement of an enterprise. This research aims to bridge this gap by proposing a new multilevel regression model with enterprise-level responses, which treats an individual's contributions to the responses as latent variables. Furthermore, the proposed model is applied to a real-world application of nursing quality improvement. Most existing research in nursing quality is either qualitative or quantitative but only utilizes the unit-level information (Aiken *et al.*, 1994; Aiken *et al.*, 2002; Institute of Medicine, 2004; Kazanjian *et al.*, 2005; Lake *et al.*, 2006; Laschinger and Leiter, 2006; Aiken *et al.*, 2008; Friese *et al.*, 2008; Mark *et al.*, 2008), because the activities of individual nurses have long been considered to be non-quantifiable (Lamb *et al.*, 2008). This application analyzes the data collected by two co-authors (Lamb and Schmitt) in a Robert Wood Johnson Foundation (RWJF)-sponsored project, in which the research team designed the *first* instrument to measure nurses' independent activities and their interdependent activities in coordinating

patient care using a comprehensive collection of variables. The data also include measurements of unit-level quality-affecting factors and quality outcomes. By linking these multiple data sources together, the findings may have a profound impact on nursing quality improvement.

The remainder of the article is organized as follows: Section 2 presents the new multilevel model development; Section 3 performs simulation studies to assess the model performance; Section 4 presents the findings in applying the proposed model to nursing quality improvement; Section 5 concludes the article.

2. Proposed model—multilevel regression with enterprise-level response

2.1. Model formulation

Let i be the index for enterprises and m be the total number of enterprises; i.e., $i = 1, \dots, m$. Let j be the index for individuals and n_i be the total number of individuals in enterprise i ; i.e., $j = 1, \dots, n_i$. Let y_i denote the response (quality outcome) of enterprise i . Let \tilde{y}_{ij} denote the contribution of individual j to y_i and \tilde{y}_{ij} is latent. In this article; we focus on the situation when the contributions of individuals to the enterprise-level response are additive i.e., $y_i = \sum_{j=1}^{n_i} \tilde{y}_{ij}$. Please note that although model formulation (this section) and estimation and inference (next section) are discussed for this simple additive relationship between \tilde{y}_{ij} , $j = 1, \dots, n_i$, and y_i , they can be readily extended to address two other situations: one situation is that y_i is a weighted sum of the \tilde{y}_{ij} , $j = 1, \dots, n_i$, with known weights; the other situation is that y_i is a general function of \tilde{y}_{ij} , $j = 1, \dots, n_i$ —i.e., $y_i = f(\tilde{y}_{ij}, j = 1, 2, \dots, n_i)$ —with the function form, f , known and can be reasonably approximated by a linear function through the Taylor expansion. In this article, we focus on $y_i = \sum_{j=1}^{n_i} \tilde{y}_{ij}$ for better presentation and clarity and also considering that this relationship is appropriate for the real-world application in Section 4.

Also, we focus on a single response; a multi-response model is just a straightforward extension to the single-response model we propose. Furthermore, assume that there are a total of Q individual-level predictors and P enterprise-level predictors. Let x_{qij} denote the measurement on the q th individual-level predictor for individual j in enterprise i . Let s_{pi} be the measurement on the p th enterprise-level predictor for enterprise i .

The model formulation consists of two stages: At Stage 1, each individual's latent response is linked to a set of individual-level predictors. At Stage 2, the regression coefficients in the Stage 1 model for each enterprise are response variables that are hypothesized to be explained by enterprise-level predictors. Specifically,

The Stage 1 model is to regress the latent response on the individual-level predictors; that is,

$$\tilde{y}_{ij} = \beta_{0i} + \beta_{1i}x_{1ij} + \dots + \beta_{Q_i}x_{Q_{ij}} + \varepsilon_{ij}, \quad (1)$$

where $\varepsilon_{ij} \sim N(0, \sigma^2)$. This is to assume that the residual errors, ε_{ij} , $i = 1, \dots, m$, $j = 1, \dots, n_i$, are normal and independent and identically distributed (i.i.d.) across different individuals and enterprises. Note that model (1) keeps i (index for enterprises) in the subscript for the model coefficients, $\beta_{0i}, \dots, \beta_{Q_i}$. This is to acknowledge the uniqueness of each enterprise; i.e., we consider that the effect of individual-level predictors on the quality response may vary across enterprises.

The Stage 2 model is to regress the regression coefficients in Equation (1) on the enterprise-level predictors

$$\beta_{qi} = \gamma_{q0} + \gamma_{q1}s_{1i} + \dots + \gamma_{qp}s_{pi} + u_{qi}, \quad q = 0, 1, \dots, Q, \quad (2)$$

where $u_{qi} \sim N(0, \tau_{qq})$ and $\text{cov}(u_{qi}, u_{q'i}) = \tau_{qq'}$ for $q \neq q'$. Note that the distribution parameters, τ_{qq} and $\tau_{qq'}$, do not have enterprise index i in their subscripts. This is to assume that the random effects of different enterprises; i.e., u_{qi} , $i = 1, \dots, m$, are i.i.d. across the enterprises. This model aims to characterize how enterprise-level predictors may modify the effect that the individual-level predictors have on the response.

By inserting Equation (2) into Equation (1), a combined model can be obtained:

$$\begin{aligned} \tilde{y}_{ij} = & \gamma_{00} + \sum_{q=1}^Q \gamma_{q0}x_{qij} + \sum_{p=1}^P \gamma_{0p}s_{pi} \\ & + \sum_{p=1}^P \sum_{q=1}^Q \gamma_{qp}s_{pi}x_{qij} + u_{0i} + \sum_{q=1}^Q u_{qi}x_{qij} + \varepsilon_{ij} \end{aligned} \quad (3)$$

where γ_{00} is the grand mean; γ_{q0} is the fixed main effect from an individual-level predictor; γ_{0p} is the fixed main effect from an enterprise-level predictor; γ_{qp} is the fixed interaction effect between an individual-level predictor and an enterprise-level predictor; u_{0i} and u_{qi} are random effects; ε_{ij} is the residual error. The two-stage model formulation can be more clearly depicted by Fig. 1.

2.2. Model estimation

To make the discussion in this section easier, we adopt an equivalent but more succinct representation for the combined model in Equation (3); that is,

$$\tilde{y}_{ij} = \mathbf{s}_i^T \mathbf{\Upsilon} \mathbf{x}_{ij} + \mathbf{u}_i^T \mathbf{x}_{ij} + \varepsilon_{ij}, \quad (4)$$

where \mathbf{x}_{ij} is a vector of measurements on the Q individual-level predictors for individual j in enterprise i —i.e., $\mathbf{x}_{ij} = [1, x_{1ij}, \dots, x_{Q_{ij}}]^T$; \mathbf{u}_i is a vector of the random effects—i.e., $\mathbf{u}_i = [u_{0i}, u_{1i}, \dots, u_{Q_i}]^T$; \mathbf{s}_i is a vector of measurements on the P enterprise-level predictors for enterprise i —i.e., $\mathbf{s}_i = [1, s_{1i}, \dots, s_{pi}]^T$; $\mathbf{\Upsilon}$ is a $(1 + P) \times (1 + Q)$ matrix of the

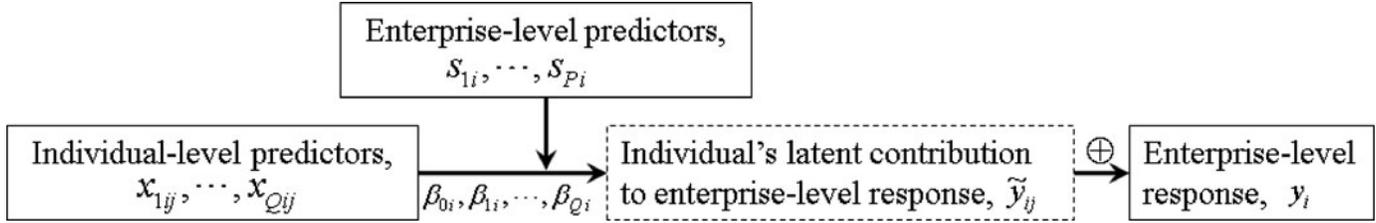


Fig. 1. Proposed multilevel model with enterprise-level response.

fixed effects; that is,

$$\mathbf{\Upsilon} = \begin{bmatrix} \gamma_{00} & \gamma_{10} & \cdots & \gamma_{Q0} \\ \gamma_{01} & \gamma_{11} & \cdots & \gamma_{Q1} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{0P} & \gamma_{1P} & \cdots & \gamma_{QP} \end{bmatrix}.$$

Furthermore, let $\mathbf{\Gamma}$ denote the covariance matrix of \mathbf{u}_i .

According to model (4), the parameters to be estimated are $\Psi = \{\mathbf{\Upsilon}, \mathbf{\Gamma}, \sigma^2\}$. The Maximum Likelihood Estimation (MLE) method can be employed. Specifically, it can be derived from Equation (4) that \tilde{y}_{ij} follows a normal distribution; i.e., $\tilde{y}_{ij} \sim N(\mathbf{s}_i^T \mathbf{\Upsilon} \mathbf{x}_{ij}, \mathbf{x}_{ij}^T \mathbf{\Gamma} \mathbf{x}_{ij} + \sigma^2)$. Furthermore, based on the relationship $y_i = \sum_{j=1}^{n_i} \tilde{y}_{ij}$, the distribution of y_i can be derived; that is,

$$y_i \sim N \left(\sum_{j=1}^{n_i} \mathbf{s}_i^T \mathbf{\Upsilon} \mathbf{x}_{ij}, \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \mathbf{\Gamma} \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) + n_i \sigma^2 \right). \quad (5)$$

Because the y_i , $i = 1, \dots, m$, are observable, the log-likelihood function of Ψ can be built upon the y_i variables; that is,

$$l(\Psi) \propto \sum_{i=1}^m \left\{ -\frac{1}{2} \log \left(\left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \mathbf{\Gamma} \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) + n_i \sigma^2 \right) - \frac{1}{2} \frac{(y_i - \sum_{j=1}^{n_i} \mathbf{s}_i^T \mathbf{\Upsilon} \mathbf{x}_{ij})^2}{\left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \mathbf{\Gamma} \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) + n_i \sigma^2} \right\}. \quad (6)$$

It is difficult to find closed-form expressions for the maximizer of the log-likelihood in Equation (6). Therefore, some iterative algorithm needs to be adopted. We adopt the Fisher Scoring (FS) algorithm (Longford, 1987). The FS algorithm begins with user-specified initial values for the parameters—i.e., $\Psi^{(0)} = \{\mathbf{\Upsilon}^{(0)}, \mathbf{\Gamma}^{(0)}, \sigma^{2(0)}\}$ in our case—and then continuously updates the parameters by $\Psi^{(k+1)} = \Psi^{(k)} + \alpha^{(k+1)} \mathbf{p}_k$ until some convergence condition is met; e.g., $\|\Psi^{(k+1)} - \Psi^{(k)}\|_2 < \varepsilon$, where $\varepsilon = 10^{-4}$ is a common choice. Clearly, the key in adapting the FS algorithm to our problem setting is to obtain \mathbf{p}_k and $\alpha^{(k+1)}$, called the *step direction* and *step size*, respectively. In what follows, we will first discuss how to obtain \mathbf{p}_k and then discuss how to select $\alpha^{(k+1)}$.

The step direction, \mathbf{p}_k , can be expressed as $\mathbf{p}_k = -\mathbf{H}^{-1} \mathbf{F}$, where \mathbf{F} is the gradient $\partial l(\Psi) / \partial \Psi$ evaluated at $\Psi^{(k)}$ and \mathbf{H} is the expectation of the matrix $\partial l(\Psi)^2 / \partial \Psi \partial \Psi$, $E(\partial l(\Psi)^2 / \partial \Psi \partial \Psi)$, evaluated at $\Psi^{(k)}$. The derivation of $\partial l(\Psi) / \partial \Psi$ and $E(\partial l(\Psi)^2 / \partial \Psi \partial \Psi)$ can be found in the Appendix. One potential limitation in use of $\mathbf{p}_k = -\mathbf{H}^{-1} \mathbf{F}$ for computing \mathbf{p}_k is that we need to calculate the inverse of matrix \mathbf{H} , which may be computationally inefficient if the dimension of Ψ is large. To overcome this limitation, we propose the following strategy. Note that $E(\partial \mathbf{\Upsilon} \partial / \partial \mathbf{\Upsilon} \partial \sigma^2) = 0$ and $E(\partial l(\Psi)^2 / \partial \mathbf{\Upsilon} \partial \mathbf{\Gamma}) = 0$; i.e., the update of $\mathbf{\Upsilon}$ is independent of $\mathbf{\Gamma}$ and σ^2 in each iteration of the FS algorithm. This suggests another way to update $\mathbf{\Upsilon}$. Specifically, we can rewrite Equation (5) as

$$y_i = \text{vec} \left(\mathbf{s}_i^T \otimes \sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) \text{vec}(\mathbf{\Upsilon})^T + \xi_i, \quad (7)$$

where $\xi_i \sim N(0, (\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \mathbf{\Gamma} (\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2)$, and $\text{vec}(\cdot)$ is an operator that converts a matrix into a row vector by concatenating the columns of the matrix. Furthermore, Let \mathbf{V} be an $m \times m$ diagonal matrix with diagonal entry $\mathbf{V}_{ii} = (\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \mathbf{\Gamma} (\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2$. Let \mathbf{Z} be an $m \times (1+P)(1+Q)$ matrix with the i th row being $\text{vec}(\mathbf{s}_i^T \otimes \sum_{j=1}^{n_i} \mathbf{x}_{ij})$. Let $\mathbf{Y} = [y_1, \dots, y_m]^T$. Then, we can use the generalized least squares method to estimate the $\text{vec}(\mathbf{\Upsilon})^T$ in Equation (7); that is,

$$\text{vec}(\hat{\mathbf{\Upsilon}})^T = (\mathbf{Z}^T \mathbf{V} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{V} \mathbf{Y}. \quad (8)$$

The incorporation of this new strategy into the proposed FS algorithm results in the following new algorithm. At the k th iteration, the FS algorithm is used to obtain estimates for $\mathbf{\Gamma}$ and σ^2 ; i.e., $\mathbf{\Gamma}^{(k)}$ and $\sigma^{(k)}$, which are used to compute $\mathbf{V}^{(k)}$; then, the generalized least squares method is used to obtain an estimate for $\text{vec}(\mathbf{\Upsilon})^T$; i.e., $\text{vec}(\mathbf{\Upsilon}^{(k)})^T = (\mathbf{Z}^T \mathbf{V}^{(k)} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{V}^{(k)} \mathbf{Y}$. A complete description of this new algorithm is given in Fig. 2.

Next, we discuss how to select the step size, $\alpha^{(k+1)}$. The selection of $\alpha^{(k+1)}$ is a classic but challenging problem in optimization. In our case, the approach introduced in Rosenberg (1973) is recommended. Specifically, the FS algorithm can start with $\alpha^{(k+1)} = 1$. If $l(\Psi^{(k+1)}) > l(\Psi^{(k)})$, then accept $\alpha^{(k+1)} = 1$; otherwise, make $\alpha^{(k+1)} = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$, until $l(\Psi^{(k+1)}) > l(\Psi^{(k)})$. This approach works well

Input: data for individual- and enterprise-level predictors and enterprise-level response, i.e., \mathbf{x}_{ij} , \mathbf{s}_i , and y_i ; initial values for the parameters, $\Psi^{(0)}$; stopping criterion, ϵ . Denote $\Psi_{/\gamma} = \{\Gamma, \sigma^2\}$.

Initialize:
 Let $converge = false$;
 Let $k = 0$;

Repeat
 Calculate $\mathbf{F} = \frac{\partial l(\Psi_{/\gamma})}{\partial \Psi_{/\gamma}}$ and $\mathbf{H} = E\left(\frac{\partial^2 l(\Psi_{/\gamma})}{\partial \Psi_{/\gamma} \partial \Psi_{/\gamma}}\right)$ by (A-1) and (A-2), respectively, both evaluated at $\Psi_{/\gamma}^{(k)}$;
 Let $\mathbf{p}_k = -\mathbf{H}^{-1}\mathbf{F}$ and calculate $\Psi_{/\gamma}^{(k+1)} = \Psi_{/\gamma}^{(k)} + \alpha^{(k+1)}\mathbf{p}_k$;
 Calculate $\mathbf{Y}^{(k+1)}$ by (8).

If $\|\Psi^{(k+1)} - \Psi^k\|_2 \leq \epsilon$ **then**
 $converge = true$;
Else
 $converge = false$;
End if

Let $k = k + 1$;

Until $converge = true$
Output: $\Psi^{k+1} = \{\Psi_{/\gamma}^{(k+1)}, \mathbf{Y}^{(k+1)}\}$.

Fig. 2. The proposed algorithm for estimating the model parameters, $\Psi = \{\mathbf{Y}, \Gamma, \sigma^2\}$, of the proposed multilevel model with enterprise-level response.

empirically. However, a potential drawback is that sometimes it may be impossible to find a positive $\alpha^{(k+1)}$ that makes $l(\Psi^{(k+1)}) > l(\Psi^{(k)})$, so the algorithm will break down. A major reason is that the iterations in the FS algorithm may be far away from the optimal solution. To avoid this, the initial values, $\Psi^{(0)} = \{\mathbf{Y}^{(0)}, \Gamma^{(0)}, \sigma^{2(0)}\}$, should be well chosen.

To select good initial values, we adapt the method developed in Sun *et al.* (2007) for conventional multilevel regression with individual-level responses to our problem that considers the individual-level responses to be latent. Specifically, our method includes two steps. The first step is to select $\mathbf{Y}^{(0)}$ and $\sigma^{2(0)}$. For this purpose, we ignore the random effects, so Equation (5) becomes $y_i \sim N(\sum_{j=1}^{n_i} \mathbf{s}_i^T \mathbf{Y} x_{ij}, n_i \sigma^2)$. Estimates for \mathbf{Y} and σ^2 can be easily obtained by MLE; these estimates are used as $\mathbf{Y}^{(0)}$ and $\sigma^{2(0)}$, respectively.

The second step is to select $\Gamma^{(0)}$. Specifically, let $r_i = y_i - \sum_{j=1}^{n_i} \mathbf{s}_i^T \mathbf{Y}^{(0)} x_{ij}$. Then, according to Equation (4), we can write

$$r_i = \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \mathbf{u}_i + \varepsilon_i, \quad (9)$$

where $\varepsilon_i = \sum_{j=1}^{n_i} \varepsilon_{ij}$. Based on Equation (9), the least squares estimate for \mathbf{u}_i can be obtained; that is,

$$\hat{\mathbf{u}}_i = \left[\left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \right]^{-1} \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) r_i. \quad (10)$$

Inserting Equation (9) into Equation (10), $\hat{\mathbf{u}}_i = \mathbf{u}_i + [(\sum_{j=1}^{n_i} \mathbf{x}_{ij})(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T)]^{-1}(\sum_{j=1}^{n_i} \mathbf{x}_{ij})\varepsilon_i$. This further leads to $\sum_{i=1}^m \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i^T = \sum_{i=1}^m \mathbf{u}_i \mathbf{u}_i^T + \varepsilon_i^2 \sum_{i=1}^m [(\sum_{j=1}^{n_i} \mathbf{x}_{ij})(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T)]^{-1} + \sum_{i=1}^m [(\sum_{j=1}^{n_i} \mathbf{x}_{ij})(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T)]^{-1}(\sum_{j=1}^{n_i} \mathbf{x}_{ij})\varepsilon_i \mathbf{u}_i^T + \sum_{i=1}^m \varepsilon_i \mathbf{u}_i (\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T)[(\sum_{j=1}^{n_i} \mathbf{x}_{ij})(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T)]^{-1}$. Because ε_i is independent of any element in \mathbf{u}_i , the last two terms are negligible. Therefore,

$$\begin{aligned} & \frac{1}{m} \sum_{i=1}^m \mathbf{u}_i \mathbf{u}_i^T \\ &= \frac{1}{m} \left\{ \sum_{i=1}^m \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i^T - \varepsilon_i^2 \sum_{i=1}^m \left[\left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \right]^{-1} \right\} \\ &\approx \frac{1}{m} \left\{ \sum_{i=1}^m \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i^T - n_i \sigma^{(0)2} \sum_{i=1}^m \left[\left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \right]^{-1} \right\} \\ &\cdot \frac{1}{m} \sum_{i=1}^m \mathbf{u}_i \mathbf{u}_i^T, \end{aligned}$$

is used as $\Gamma^{(0)}$.

2.3. Model inference

After model parameters have been estimated, the next step is to perform hypothesis testing to check the statistical significance of the parameters. Three types of hypotheses usually need to be tested in our case: tests for fixed effects, tests for random effects, and tests for model comparison.

Specifically, to test a fixed effect, γ_{qp} —i.e., $H_0 : \gamma_{qp} = 0$ versus $H_1 : \gamma_{qp} \neq 0$ —the test statistic $t = \hat{\gamma}_{qp} / \sqrt{\text{var}(\hat{\gamma}_{qp})}$ can be used, $q = 0, 1, \dots, Q$, $p = 0, 1, \dots, P$. $\hat{\gamma}_{qp}$ is an MLE estimate for γ_{qp} , and $\text{var}(\hat{\gamma}_{qp})$ can be asymptotically approximated by the corresponding element in matrix \mathbf{H}^{-1} . Recall that both $\hat{\gamma}_{qp}$ and \mathbf{H} have been obtained from the model estimation method proposed in Section 2.2. This test statistic is asymptotically standard normal; this property can be used to calculate the P -value of the test. Furthermore, to test the random effects is equivalent to testing the covariance matrix of the random effects, Γ ; i.e., $H_0 : \Gamma = \Gamma_0$ versus $H_1 : \Gamma = \Gamma_1$. For example, to test whether a random effect u_{qi} exists, $q = 0, 1, \dots, Q$, we can set Γ_0 to be a reduced form of Γ_1 by making the q th row and column of Γ_1 to be null. Denote the maximum log-likelihood values under H_0 and H_1 by l_0 and l_1 . Then, the test statistic is $-2(l_0 - l_1)$, which has an approximate χ^2 distribution with d degrees of freedom, where d is the difference between the number of unique parameters in Γ_1 and that in Γ_0 . This is a typical Likelihood Ratio Test (LRT), which can also be used to simultaneously test multiple fixed and random effects; i.e., LRT can be used in model comparison.

3. Simulation studies

An intuitive, competing approach to the proposed model is to aggregate the individual-level predictors to the enterprise level by using some summary statistics and then perform general regression analysis at the enterprise level. Limitations of this aggregate model have been conceptually discussed in Section 1. In this section, we will use simulation data to compare the performance of the aggregate model and the proposed model in terms of the statistical properties in fixed and random effect estimation.

To generate the simulation data, the following true model is used. Consider that an individual-level response, y_{ij} , is linked to one individual-level predictor, x_{ij} , by a regression $y_{ij} = \beta_i x_{ij} + \varepsilon_{ij}$ and β_i is linked to one enterprise-level predictor, s_i , by a regression $\beta_i = \gamma s_i + u_i$. These two regressions can be combined into one; i.e., $y_{ij} = \gamma s_i x_{ij} + u_i x_{ij} + \varepsilon_{ij}$. The parameters of the true model include the fixed effect, γ , variance of the random effect, $\tau = \text{var}(u_i)$, and variance of the residual error, $\sigma^2 = \text{var}(\varepsilon_{ij})$. γ , τ , and σ^2 are all assumed to be one. Furthermore, data are generated from the true model for each unit. Specifically, for unit i , the following steps are performed: (i) draw two sample from $N(0, 1)$ and make them the values for u_i and s_i , respectively, and compute β_i ; (ii) draw n_i samples from $N(0, 1)$ for $\varepsilon_{i1}, \dots, \varepsilon_{in_i}$, and another n_i samples from $N(0, 1)$ for x_{i1}, \dots, x_{in_i} , and compute y_{i1}, \dots, y_{in_i} ; (iii) compute $y_i = \sum_{j=1}^{n_i} y_{ij}$.

Based on the data, y_i , x_{ij} , and s_i , $j = 1, \dots, n_i$, $i = 1, \dots, m$, the proposed model can be applied to estimate the true model parameters γ , τ , and σ^2 . Alternatively, the aggregate model can also be applied, which builds an ordinary regression of y_i on predictors s_i , $x_i = \sum_{j=1}^{n_i} x_{ij}$, and $s_i x_i$. In this aggregate model, γ is estimated by the coefficient of predictor $s_i x_i$. The aggregate model is not able to separate between- and within-enterprise variations by providing separate estimates for τ and σ^2 . Instead, it estimates the overall variation by the residual variance. In addition, as data on the individual-level response, y_{ij} , are available in simulation, conventional multilevel regression is also applied to estimate γ , τ , and σ^2 . These estimates can be used as the gold standard to assess the impact of treating the individual-level response as latent by the proposed model. The results of comparison between the three models are presented as follows.

3.1. Fixed effect estimation

For each model, the average estimate for γ over 100 repetitions of the simulation was obtained. Furthermore, the deviation of the average estimate from the true $\gamma = 1$ was computed to assess the bias in the estimation. The deviations for the gold standard, proposed, and aggregate models are 0.028, 0.021, 0.023, respectively, when $m = 10$ (number of enterprises) and $n_i = n = 50$ (enterprise sample

size). The deviations are small, which is also true for other m and n values. This implies that the proposed model gives unbiased estimators for fixed effects. In fact, this property of the proposed model can be theoretically proved. Specifically, according to Equation (10), the estimates for the fixed effect are $\text{vec}(\hat{\mathbf{Y}})^T = (\mathbf{Z}^T \mathbf{V} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{V} \mathbf{Y}$. Then,

$$\begin{aligned} (\text{vec}(\hat{\mathbf{Y}})^T) &= \{(\mathbf{Z}^T \mathbf{V} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{V}\} E(\mathbf{Y}) \\ &= \{(\mathbf{Z}^T \mathbf{V} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{V}\} \mathbf{Z} \text{vec}(\mathbf{Y})^T = \text{vec}(\mathbf{Y})^T, \end{aligned}$$

where the second “=” follows from Equation (8). The gold standard model also gives unbiased estimators for fixed effects, which is a well-known property for conventional multilevel regression. In the aggregate model, the data on the response variable—i.e., the y_i —are independent but non-identically distributed, because $\text{var}(y_i) = x_i^2 \tau + n \sigma^2$. Even though the data are not i.i.d., ordinary least squares estimation can still give unbiased estimators for the regression coefficients (Demidenko, 2004), so the estimates for fixed effects by the aggregate model are also unbiased.

Furthermore, we compare the three models in terms of the standard error of the fixed effect estimate. Figure 3(a) shows the standard error averaged over 100 repetitions of the simulation by each model (y -axis) with respect to the number of enterprises, m (x -axis). The enterprise sample size, n , was fixed to be 50. It can be seen that the standard error by the proposed model is very close to that by the gold standard model, whereas that by the aggregate model is much larger. A large standard error leads to the risk of mis-detecting significant fixed effects. Also, Fig. 3(a) shows that increasing the number of enterprises, m , can significantly reduce the standard errors for all three models. Furthermore, we varied the enterprise sample size by generating the sample size of each enterprise, n_i , from a uniform distribution on the interval [1, 10] (Rosenberg, 1973; Lamb *et al.*, 2008). The simulation was repeated and the results are shown in Fig. 4. Comparing Fig. 4 with Fig. 3, it can be seen that smaller and unbalanced enterprise sample sizes increase the standard errors but only slightly. In other words, the enterprise sample size influences the standard errors much less than the number of enterprises. This observation is consistent with existing knowledge about multilevel regression (Draper and Smith, 1998).

3.2. Random effect estimation

The aggregate model does not include any random effects. Therefore, the comparison is between the proposed and gold standard models. For each model, the deviation of the average estimate for τ from the true $\tau = 1$ was computed to assess the bias in the estimation. The deviations for the gold standard and proposed models are -0.015 and -0.024 , respectively, when $m = 100$ and $n = 50$. The magnitudes of these deviations become smaller when m increases. This empirically implies that both models might give unbiased estimators for τ . Furthermore, we compared the two

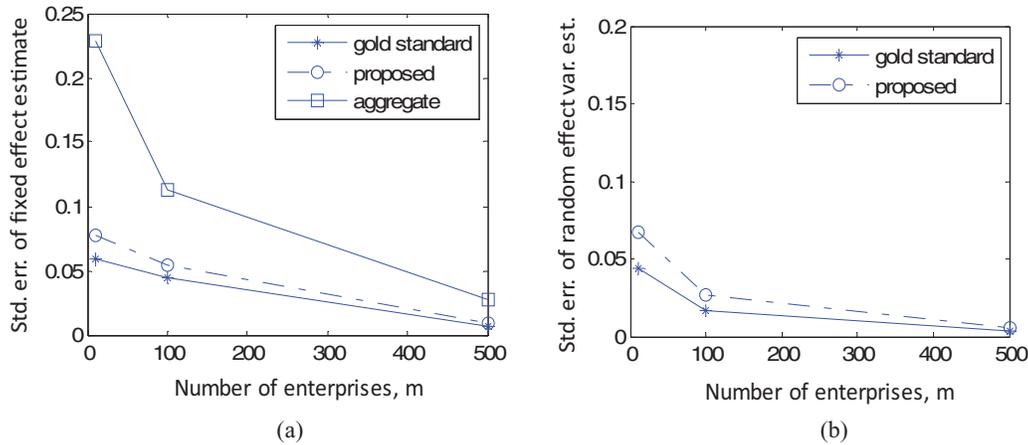


Fig. 3. (a) Standard error of the estimate for the fixed effect, γ , versus the number of enterprises, m , and (b) standard error of the estimate for the random effect variance, τ , versus the number of enterprises, m . In both figures, enterprise sample size is fixed to be $n = 50$.

models in terms of the standard error of the estimate for τ . The result is given in Fig. 3(b), from which it is clear that the standard error by the proposed model is very close to that by the gold standard model. Increasing the number of enterprises, m , can significantly reduce the standard errors for both models, whereas increasing n does not have this effect (results not shown here).

3.3. Residual error estimation

For each model, the deviation of the average estimate for σ from the true $\sigma = 1$ was computed to assess the bias in the estimation. The deviations for the gold standard, proposed, and aggregate models are -0.003 , -0.177 , 13.394 , respectively, when $m = 100$ and $n = 50$. The magnitudes of these deviations only display small changes when m and n in-

crease. This implies that the proposed model may be biased in estimating the residual error σ ; fortunately, the magnitude of the bias is small. In contrast, the aggregate model over estimates σ with a large bias. The large bias is due to the fact that the aggregate model cannot separate within- and between-enterprise variation, so that the estimate for the residual error is a combination of the two variation sources. Furthermore, we compared the three models in terms of the standard error of the estimate for σ . The same phenomena/trends were observed as for the standard error of the fixed effect estimate.

Furthermore, we performed simulation studies with more individual- and enterprise-level predictors. Specifically, in the first set of simulations, we included Q individual-level predictors ($Q > 1$), while keeping the number of enterprise-level predictors to be one. Therefore,

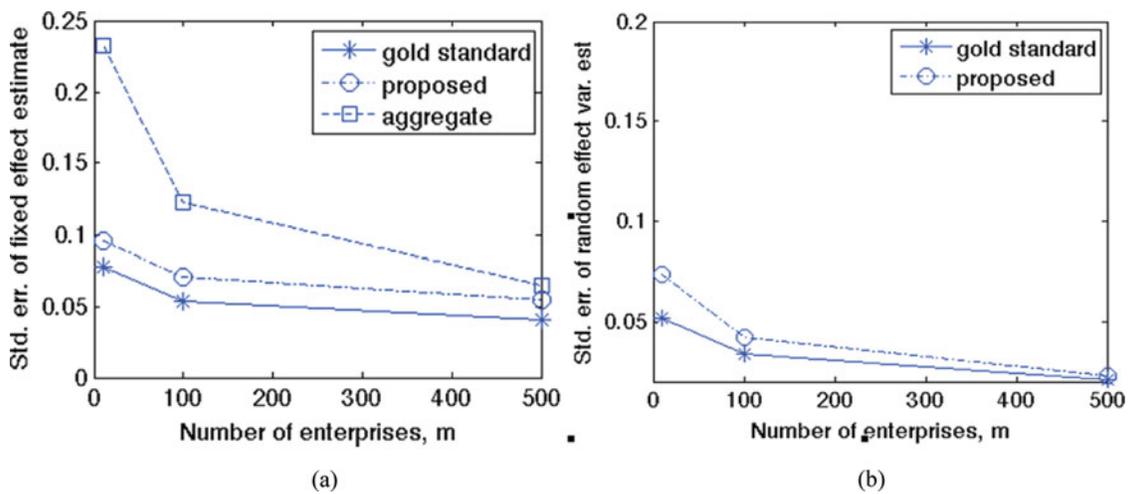


Fig. 4. (a) Standard error of the estimate for the fixed effect, γ , versus number of enterprises, m , and (b) standard error of the estimate for the random effect variance, τ , versus number of enterprises, m . In both figures, the sample size of each enterprise is sampled from uniform[1, 10].

Table 1. A comparison of the standard error of fixed effect estimate for the proposed, gold standard, and aggregate models in terms of the statistical properties in model estimation ($m = 100, n = 50$)

	$Q = 1, P = 1$	$Q = 8, P = 1$	$Q = 1, P = 8$
Gold standard	0.045	0.328	0.304
Proposed	0.054	0.554	0.613
Aggregate	0.113	1.012	0.719

the true model used to generate the simulation data was $y_{ij} = \sum_{q=1}^Q \gamma_q s_i x_{qij} + \sum_{q=1}^Q u_{qi} x_{qij} + \varepsilon_{ij}$, where $\gamma_q = 1$, $\tau_{qq} = \text{var}(u_{qi}) = 1$, and $\sigma^2 = \text{var}(\varepsilon_{ij}) = 1$. s_i and x_{qij} were sampled from the $N(0, 1)$ distribution. Based on the simulation data, the proposed, aggregate, and gold standard models were applied to estimate the Q fixed effects, variances of the Q random effects (the aggregate model cannot estimate these), and the residual variance. Because all three models have been theoretically proven to give an unbiased estimator for each fixed effect, the simulation result for assessing the bias of each model in fixed effect estimation is not shown here. The models were compared in terms of the standard errors of the fixed effects estimates, the biases, and standard errors of the random effect variance estimates and the bias and standard error of the residual variance estimate. Due to page limits, instead of showing the standard error of each fixed effect estimate, we show the average over the standard errors of the estimates for the Q fixed effects. A similar consideration applies to the random effects; i.e., we show the average over the biases/standard errors of the estimates for the Q random effects' variances. The results for $Q = 8$ are shown in the " $Q = 8, P = 1$ " column of Tables 1–5.

In the second set of simulations, we included P enterprise-level predictors ($P > 1$), while keeping the number of individual-level predictors at one. Therefore, the true model used to generate the simulation data was $y_{ij} = \sum_{p=1}^P \gamma_p s_{pi} x_{ij} + u_i x_{ij} + \varepsilon_{ij}$, where $\gamma_p = 1$, $\tau = \text{var}(u_i) = 1$, and $\sigma^2 = \text{var}(\varepsilon_{ij}) = 1$. s_{pi} and x_{ij} were sampled from the $N(0, 1)$ distribution. The three models were applied to the simulation data. The results for $P = 8$ are in the " $Q = 1, P = 8$ " column of Tables 1 to 5. Note that because there are eight fixed effects, only the average over the standard errors of the estimates for these fixed effects is

Table 2. A comparison of the standard error of random effect variance estimate for the proposed, gold standard, and aggregate models in terms of the statistical properties in model estimation ($m = 100, n = 50$)

	$Q = 1, P = 1$	$Q = 8, P = 1$	$Q = 1, P = 8$
Gold standard	0.017	0.614	0.086
Proposed	0.027	1.245	0.257
Aggregate	N/A	N/A	N/A

Table 3. A comparison of the bias of random effect variance estimate for the proposed, gold standard, and aggregate models in terms of the statistical properties in model estimation ($m = 100, n = 50$)

	$Q = 1, P = 1$	$Q = 8, P = 1$	$Q = 1, P = 8$
Gold standard	-0.015	0.044	0.064
Proposed	-0.027	0.214	0.092
Aggregate	N/A	N/A	

shown in Table 1 due to page limits. In addition, the results for one individual-level predictor and one enterprise-level predictor, discussed previously, are copied here for the purpose of comparison.

The following observations can be drawn from the performed simulations:

1. The proposed model is consistently better than the aggregate model in terms of all the statistical properties chosen for comparison and regardless of the number of predictors.
2. The proposed model performs close to the gold-standard model, especially when the number of individual-level predictors, Q , is small. This is because Q represents the number of random effects; a large Q increases the variation sources, introducing more uncertainty in the model estimation.
3. With a fixed number of enterprises—i.e., $m = 100$ in Table 6—the fewer the number of parameters to be estimated, the better the estimation. This is true for all three models. Furthermore, the enterprise sample size, n , influences the model performance to a much smaller extent than the number of enterprises.

4. Application in nursing quality improvement

4.1. Data collection and selection of predictors and response variable

The case study in this section uses the data collected by two co-authors in an RWJF-sponsored project. Data were collected mainly in the format of surveys handed out to 614 nurses in 32 hospital units (human subject approval

Table 4. A comparison of the standard error of the residual standard deviation (σ) estimate for the proposed, gold standard, and aggregate models in terms of the statistical properties in model estimation ($m = 100, n = 50$)

	$Q = 1, P = 1$	$Q = 8, P = 1$	$Q = 1, P = 8$
Gold standard	0.000 01	0.005	0.002
Proposed	0.02	0.281	0.047
Aggregate	3	26.82	20.368

Table 5. A comparison of the bias of the residual standard deviation (σ) estimate for the proposed, gold standard, and aggregate models in terms of the statistical properties in model estimation ($m = 100, n = 50$)

	$Q = 1, P = 1$	$Q = 8, P = 1$	$Q = 1, P = 8$
Gold standard	0.003	0.308	0.144
Proposed	0.177	3.228	0.131
Aggregate	13.394	22.174	20.598

was obtained). Using the “language” of this article, nurses are the “individuals” and units are the “enterprises.”

4.1.1. Selection of individual-level predictors

The survey included 11 questions designed to measure independent nursing activities and interdependent activities in coordinating patient care. Examples of the 11 questions include “I organize the supplies that I need to be able to keep the care of my patients on track,” “I initiate actions to get my nursing team members to do what is needed to keep my patients on their plan of care,” and “I communicate information to my interdisciplinary team members they need to know to carry out their patient care activities or to make changes in their plan of care.” Each question included four aspects for capturing the nurses’ perceptions of (i) the amount of time spent on the activity in a usual shift; (ii) the priority placed on the activity for a usual shift; (iii) the amount of time spent on the activity in the last shift worked; and (iv) the amount of time spent on this activity compared with the perception of the amount of time needed in the last shift worked. In the survey, the nurses were asked to respond to the four aspects for each question. The response to each aspect was on a 1 to 5 numerical scale, where 1 represents the least amount of time spent (for aspects (i), (iii), and (iv))/the lowest priority (for aspects (ii)) and 5 represents the most amount of time spent/the highest priority. In the data analysis described in this section, we took the average over the responses to the four aspects and considered the average as the response to each question. In this way, the response to each question was a combined measure for the amount of time spent and priority of the activity to which this question corresponds.

The 11 questions were designed to be indicators for six underlying constructs, including organizing one’s own activities and resources, checking patient progress and response, doing the work of others to keep care on track, assisting each other’s work, mobilizing people and resources, and exchanging information with team members. The six constructs are called organizing, checking, backfilling, assisting, mobilizing, and exchanging, or denoted by o , c , b , a , m , and e in short in this section. Note that constructs o , c , and b correspond to the independent nursing activities, whereas a , m , and e correspond to the interdependent nurs-

ing activities; i.e., activities that need coordination between nurses and with other healthcare professionals.

To verify the designed/hypothetical correspondence between the 11 questions and the six constructs, we performed factor analysis with a rotation method called Procrustes (Harman, 1976). This method identifies the factors that underly the questions, such that the correspondence between the factors and the questions maximally overlaps with the designed/hypothetical correspondence between the six constructs and the questions. Six factors were identified and their correspondence with the 11 questions is almost the same as the designed/hypothetical correspondence, except that a is identified to correspond to questions Q5 and Q6, whereas it was designed to correspond to only Q5. Question Q6 was “I communicate information to my nursing team members that they need to know to carry out their patient care activities or to make changes in the plan of care.” It is reasonable to believe that this question is an indicator for both e and a . Therefore, we used the six factors identified as individual-level predictors in our model.

4.1.2. Selection of unit-level predictors

The survey also included 30 questions capturing the nurses’ perceptions about the infrastructure, climate, and administration strategies in their respective units. Examples of the 30 questions include “Our information technology helps me to find the information I need quickly,” “Physicians respond quickly when we call them for a change in an order or change in patient status,” and “The physical layout of the unit allows us to get the supplies we need easily.” The response to each question was on a 1 to 5 numerical scale, where 1 to 5 represents “strongly disagree” to “strongly agree.” Note that although these questions asked for an individual nurse’s response, it is more appropriate to include them as unit-level predictors rather than individual-level predictors. Therefore, we took the average over the responses from all of the nurses in each unit for a question. Furthermore, we performed principal component analysis (Jolliffe, 2001) on the 30 unit-level predictors and kept the first Principle Component (PC) as the final unit-level predictor included in our model. This had two purposes: (i) to reduce the number of unit-level predictors; (ii) the first PC is a linear combination of the 30 unit-level predictors, thus hypothetically serving as an overall measure for the extent to which each unit has characteristics that facilitate nursing quality improvement.

4.1.3. Selection of response variable/quality outcome

Unit-level data for measuring the quality of nursing in each unit were collected separately from the survey. We included one quality measure, the total number of falls per 1000 patient days, or “falls” in short, as the response variable in our model. A summary of the predictors and response variable that were selected is given in Table 6.

Table 6. Description of the predictors and response variable included in the model

	Description	Notes
<i>Individual-level predictors</i>		
<i>o</i>	A factor measuring the level of nursing activity in organizing one's own activities and resources	<i>o</i> , <i>c</i> , <i>b</i> , <i>a</i> , <i>m</i> , and <i>e</i> are all standardized variables. Each is standardized by the respective global (i.e., across all units) mean and global standard deviation. Larger values in a predictor indicate higher levels of activity this predictor describes. "Level" is a combined measure for the time spent in this activity and the priority of this activity.
<i>c</i>	A factor measuring the level of nursing activity in checking patient progress and response	
<i>b</i>	A factor measuring the level of nursing activity in doing the work of others to keep care on track	
<i>a</i>	A factor measuring the level of nursing activity in assisting each other's work	
<i>m</i>	A factor measuring the level of nursing activity in mobilizing people and resources	
<i>e</i>	A factor measuring the level of nursing activity in exchanging information with team members	
<i>Unit-level predictors</i>		
<i>s</i>	A composite measure based on 30 variables measuring unit infrastructure, climate, and administration strategies	Larger values in <i>s</i> indicate that the unit has characteristics that facilitate nursing quality improvement, hypothetically.
<i>Unit-level response variable</i>		
Falls	Number of falls per 1000 patient days	Smaller values indicate better nursing quality.

4.2. Modeling and results

We generated a scatterplot for each predictor in Table 6 with respect to the response. For an individual-level predictor, we plotted the unit average because the response was at the unit level. The scatterplots (see online supplement) show linear trends. This confirms the validity of using a regression model for this particular application. Furthermore, we applied the proposed multi-level model to the data. The Stage 1 model included all six individual-level predictors; i.e., $\tilde{y}_{ij} = \beta_{0i} + \beta_{1i}(o)_{ij} + \beta_{2i}(c)_{ij} + \beta_{3i}(b)_{ij} + \beta_{4i}(a)_{ij} + \beta_{5i}(m)_{ij} + \beta_{6i}(e)_{ij} + \varepsilon_{ij}$. The Stage 2 model regresses each Stage 1 model coefficient on the unit-level predictor, i.e., $\beta_{qi} = \gamma_{q0} + \gamma_{q1}s_i + u_{qi}$, $q = 0, 1, \dots, 6$. The combined model is

$$\begin{aligned} \tilde{y}_{ij} = & \gamma_{00} + \gamma_{10}(o)_{ij} + \gamma_{20}(c)_{ij} + \gamma_{30}(b)_{ij} + \gamma_{40}(a)_{ij} \\ & + \gamma_{50}(m)_{ij} + \gamma_{60}(e)_{ij} + \gamma_{01}s_i + \gamma_{11}(s-o)_{ij} + \gamma_{21}(s-c)_{ij} \\ & + \gamma_{31}(s-b)_{ij} + \gamma_{41}(s-a)_{ij} + \gamma_{51}(s-m)_{ij} + \gamma_{61}(s-e)_{ij} \\ & + u_{0i} + u_{1i}(o)_{ij} + u_{2i}(c)_{ij} + u_{3i}(b)_{ij} + u_{4i}(a)_{ij} \\ & + u_{5i}(m)_{ij} + u_{6i}(e)_{ij} + \varepsilon_{ij}, \end{aligned} \quad (11)$$

where $(s-o)_{ij} = s_i \times (o)_{ij}$ and the coefficient γ_{11} reflects the interaction effect between *s* and *o*. \tilde{y}_{ij} is the latent contribution of nurse *j* in unit *i* to the total number of falls in this unit. $y_i = \sum_{j=1}^{n_i} \tilde{y}_{ij}$ is the total number of falls in unit *i* and is observable.

By applying the model estimation method proposed in Section 2, we obtained estimates for the fixed effects denoted by the γ coefficients in Equation (11) and an estimate for the covariance matrix Γ of the random effects

denoted by the *u* coefficients. Here, considering the sample size limitation, we assumed that Γ was diagonal. Therefore, the proposed method actually gave an estimate for the variance/standard deviation of each random effect. In addition, the proposed method gave an estimate for σ^2 that is the variance of the residual ε_{ij} . These estimates are shown in Table 7.

It can be seen that the model in Table 7 (called the full model hereafter) has many fixed effects with large *P*-values and random effects with small standard deviations. This implies that the model may be simplified. We performed model selection using the LRT suggested in Section 2.3. The selected model was further subjected to model adequacy checks and there were no apparent violations of the model assumptions (please see online supplement for details). However, unit 26 was found to be an outlier. After removing unit 26, the final model was obtained, as shown in Table 8. The R^2 value for this model is 0.94, showing a good fit.

To facilitate the interpretation of the final model, we now write out the Stage 1 and Stage 2 models corresponding to the final model; i.e., $\tilde{y}_{ij} = \beta_{0i} + \beta_{4i}(a)_{ij} + \beta_{6i}(e)_{ij} + \varepsilon_{ij}$ (Stage-1); $\beta_{0i} = \gamma_{00} + \gamma_{01}s_i + u_{0i}$, $\beta_{4i} = \gamma_{40} + \gamma_{41}s_i$, $\beta_{6i} = u_{6i}$ (Stage 2). Some interesting conclusions can be drawn.

1. In general, because the final model only includes the individual-level predictors, *a* and *e*, it indicates that other nursing activities, *o*, *c*, *b*, and *m*, do not have a significant impact on the number of falls. Note that *a* and *e* are both interdependent nursing activities in coordinating

Table 7. Estimated effects of individual-level and unit-level predictors on the number of falls

Fixed effect	Estimate		SE		P value		
Intercept, γ_{00}	3.436		0.187		0.000		
o , γ_{10}	-0.094		0.182		0.610		
c , γ_{20}	-0.079		0.181		0.670		
b , γ_{30}	0.162		0.222		0.475		
a , γ_{40}	-0.509		0.225		0.036		
m , γ_{50}	-0.218		0.269		0.427		
e , γ_{60}	-0.165		0.327		0.620		
s , γ_{01}	-0.262		0.146		0.088		
$s.o$, γ_{11}	0.181		0.154		0.255		
$s.c$, γ_{21}	0.127		0.140		0.377		
$s.b$, γ_{31}	0.031		0.144		0.835		
$s.a$, γ_{41}	-0.246		0.181		0.189		
$s.m$, γ_{51}	0.000		0.116		1.000		
$s.e$, γ_{61}	-0.079		0.303		0.798		
	Intercept, std dev of u_{0i}	o , std dev of u_{1i}	c , std dev of u_{2i}	b , std dev of u_{3i}	a , std dev of u_{4i}	m , std dev of u_{5i}	e , std dev of u_{6i}
Random effect	1.260e-6	1.472e-6	6.247e-7	0.425	3.555e-6	1.338e-6	0.917
$\sigma =$	1.957e-05						

patient care. This indicates that the coordination between nurses and with other healthcare professionals may be more important for reducing falls, compared with independent nursing activities.

- β_{0i} is the mean number of falls in unit i when the levels of nursing activities a and e are equal to their respective global means. β_{0i} is affected by s according to the Stage 2 model; also, the coefficient for s , γ_{01} , is negative. This implies that a unit with a high level of s will have a fewer mean number of falls than a unit with a low level of s , even though the nurses in the two units have the same levels of “assisting” and “exchanging” activities. Recall that s is a composite measure for the extent to which the unit has characteristics that facilitate nursing quality improvement, hypothetically. Our finding confirms this hypothesis. In addition, β_{0i} consists of a random effect, u_{0i} , whose variance is significant. This implies that hospital units vary in their mean number of falls even

after controlling for s . Furthermore, we can estimate the average of the mean number of falls across the population of units with the same s by $\gamma_{00} + \gamma_{01}s$. For example, if we focus on the population of units with $s = 0.0317$ (this number is the average value of s over the 32 units in the data), then on average these units will have a mean number of falls equal to $\gamma_{00} + 0.0317\gamma_{01} = 3.53$.

- β_{4i} reflects the strength of association between a nurse’s assisting activity in unit i and the number of falls. β_{4i} is affected by s according to the Stage 2 model; also, the coefficient for s , γ_{41} , is negative. This implies that in a unit with a high level of s , increasing the assisting activity of nurses will reduce the number of falls more than in a unit with a low level of s . In addition, β_{4i} does not include a random effect. This implies that after controlling for s , hospital units behave similarly in terms of the strength of association between nurses’ assisting activities and the number of falls; i.e., little variability in the strength of association remains to be explained. Furthermore, $\gamma_{40} + \gamma_{41}s$ can be used to estimate the average strength of association between nurses’ assisting activities and the number of falls across the population of units with the same s . Note that because $\gamma_{40} < 0$, $\gamma_{41} < 0$, and $s > 0$, the strength of association is always negative, implying that nurses’ assisting activities will reduce the number of falls regardless of the unit to which the nurses belong.
- β_{6i} reflects the strength of association between a nurse’s exchanging activity in unit i and the number of falls. β_{6i} includes a random effect, implying that hospital units vary in terms of the strength of association between nurses’ exchanging activities and the number of falls. However, this variability cannot be accounted for by s . Furthermore, as β_{6i} does not include any fixed effect, it

Table 8. Estimated effects of individual-level and unit-level predictors on the number of falls in the final model

Fixed effect	Estimate	SE	P value
Intercept, γ_{00}	3.538	0.129	0.000
a , γ_{40}	-0.402	0.185	0.038
s , γ_{01}	-0.241	0.076	0.004
$s.a$, γ_{41}	-0.395	0.081	0.000
	Intercept, std dev of u_{0i}	e , std dev of u_{6i}	
Random effect	0.238	0.869	
$\sigma =$	0.354		

implies that on average there is little association between nurses' exchanging activities and the number of falls.

5. Conclusions

This article proposed a multilevel model to link individual- and enterprise-level predictors with an enterprise-level quality outcome for enterprise quality improvement. Unlike conventional multilevel regression, which requires the outcome be at the individual level, the proposed model treats each individual's contribution to the enterprise quality outcome as a latent variable. An algorithm was proposed to estimate the model parameters, which integrates the FS algorithm and generalized least squares estimation. Simulation studies were conducted to assess the performance of the proposed model, in comparison with the aggregate model that which aggregates the individual-level predictors to the enterprise level and the gold standard model, which assumes that each individual's contribution to the enterprise quality outcome is explicitly known. These studies showed that the proposed model performs close to the gold standard model in terms of the biases and standard errors of the estimates for the fixed effects, variances of the random effects and residual variance. In contrast, the aggregate model leads to much larger standard errors of the estimates for the fixed effects and much larger bias and standard error of the residual variance estimate; also, the aggregate model cannot separate the within- and between-enterprise variations by providing separate estimates for the random effects variances and residual variance.

The proposed model was applied to a real-world application of nursing quality improvement. Our finding showed that the interdependent nursing activities in coordinating patient care, especially the assisting and exchanging activities, have a significant impact on reducing patient falls. Also, our finding confirmed that the unit infrastructure, climate, and administration strategies that are hypothesized to improve nursing quality do help significantly reduce falls. In addition, the assisting activity of each nurse and the quality-improving infrastructure, climate, and administration strategies of the nurse's unit promote each other in reducing falls.

Finally, we point out several future research directions. Multiple quality outcomes, such as falls, medication errors, and patient satisfaction, may be considered together, leading to a multilevel model with multiple latent responses to be developed. Also, a generalized multilevel model may be more appropriate considering that the response variables may not all be strictly normal. Furthermore, robust model estimation with a large number of predictors and limited sample sizes may be studied. In addition to the methodological development, it would be of interest to investigate the difference between hospital units in terms of how the individual- and unit-level predictors affect nursing quality outcomes and formulate specific quality improvement

plans for each unit. Also, instead of using a composite measure as the unit-level predictor, it may reveal more insights to include each specific quality-assuring measure as a predictor in order to assess the effectiveness of each measure.

Funding

This work is partly supported by the Interdisciplinary Nursing Quality Research Initiative program at the Robert Wood Johnson Foundation and the National Science Foundation under grants CMMI-0825827 and CMMI-1069246.

Supplemental Material

Supplemental data for this article can be accessed at the publisher's website.

References

- Aiken, L.H., Clarke, S.P., Sloane, D.M., Lake, E.T. and Cheney, T. (2008) Effects of hospital care environment on patient mortality and nurse outcomes. *The Journal of Nursing Administration*, **38**, 223–229.
- Aiken, L.H., Clarke, S.P., Sloane, D.M., Sochalski, J.A. and Silber, J.H. (2002) Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *Journal of the American Medical Association*, **288**(16), 1987–1993.
- Aiken, L.H., Smith, H.L. and Lake, E.T. (1994) Lower Medicare mortality among a set of hospitals known for good nursing care. *Medical Care*, **32**, 771–787.
- Demidenko, E. (2004) *Mixed Models, Theory and Applications*, John Wiley, New York, NY.
- Dempster, A.P., Rubin, D.B. and Tsutakawa, R.K. (1981) Estimation in covariance components models. *Journal of the American Statistical Association*, **76**, 341–353.
- Draper, N.R. and Smith, H. (1998) *Applied Regression Analysis*, third edition, Wiley, New York, NY.
- Elston, R.C. and Grizzle, J.E. (1962) Estimation of time response curves and their confidence bands. *Biometrics*, **18**, 148–159.
- Friese, C.R., Lake, E.T., Aiken, L.H., Silber, J. and Sochalski, J.A. (2008) Hospital nurse practice environments and outcomes for surgical oncology patients. *Health Services Research*, **43**(4), 1145–1163.
- Goldstein, H. (1995) *Multilevel Statistical Models*, second edition, John Wiley, New York, NY.
- Harman, H.H. (1976) *Modern Factor Analysis*, third edition, The University of Chicago Press, Chicago, IL.
- Institute of Medicine, Committee on Quality of Health Care in America. (2004) *Keeping Patients Safe: Transforming the Work Environment of Nurses*, National Academies Press, Washington, DC.
- Jolliffe, I.T. (2001) *Principal Component Analysis*, second edition, Springer, New York, NY.
- Kazanjian, A., Green, C., Wong, J. and Reid, R. (2005) Effects of the hospital nursing environment on patient mortality: a systematic review. *Journal of Health Services Research and Policy*, **10**(2), 111–117.
- Laird, N.M. and Ware, H. (1982) Random-effects models for longitudinal data. *Biometrics*, **38**, 963–974.
- Lake, E.T. and Friese, C.R. (2006) Variations in nursing practice environments: relation to staffing and hospital characteristics. *Nursing Research*, **55**(1), 1–9.
- Lamb, G.S., Schmitt, M.H., Edwards, P., Sainfort, F., Duva, I. and Higgins, M. (2008) Measuring staff nurse care coordination in the hospital. Invited presentation, National State of the Science Congress on Nursing Research, Washington, DC.

Laschinger, H.K.S. and Leiter, M.P. (2006) The impact of nursing work environments on patient safety outcomes. *Journal of Nursing Administration*, **36**(5), 259–267.

Longford, N. (1987) A fast scoring algorithm for maximum likelihood estimation in unbalanced models with nested random effects. *Biometrika*, **74**(4), 817–827.

Longford, N. (1993) *Random Coefficient Models*, Clarendon Press, Oxford, UK.

Mark, B.A., Hughes, L.C., Belyea, M., Bacon, C.T., Chang, Y. and Jones, C.B. (2008) Exploring organizational context and structure as predictors of medication errors and falls. *Journal of Patient Safety*, **4**(2), 66–77.

Mason, W.M., Wong, G.M. and Entwistle, B. (1983) Contextual analysis through the multilevel linear model, in *Sociological Methodology*, Leinhardt, S. (ed.), Jossey-Bass, San Francisco, CA, pp. 72–103.

Robinson, W.S. (1950) Ecological correlations and the behavior of individuals. *American Sociological Review*, **15**, 351–357.

Rosenberg, B. (1973) Linear regression with randomly dispersed parameters. *Biometrika*, **60**, 61–75.

Singer, J.D. (1998) Using SAS PROC MIXED to fit multilevel models, hierarchical models: issues and methods. *Journal of Educational and Behavioral Statistics*, **23**(4), 323–355.

Sun, Y., Zhang, W.Y. and Tong, H. (2007) Estimation of the covariance matrix of random effects in longitudinal studies. *The Annals of Statistics*, **35**(6), 2795–2814.

Appendix

Derivation of $\partial l(\Psi)/\partial \Psi$ and $E\partial l(\Psi)^2/\partial \Psi \partial \Psi$ in the FS algorithm

Through some matrix algebra, we derive the first derivatives, $\partial l(\Psi)/\partial \Psi$, as follows:

$$\begin{aligned} \frac{\partial l(\Psi)}{\partial \Upsilon} &= \sum_{i=1}^m \frac{(\sum_{j=1}^{n_i} \mathbf{s}_i \mathbf{x}_{ij}^T)(y_i - \sum_{j=1}^{n_i} \mathbf{s}_i^T \Upsilon x_{ij})}{(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2}, \\ \frac{\partial l(\Psi)}{\partial \sigma^2} &= -\frac{1}{2} \sum_{i=1}^m \frac{n_i}{(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2} \\ &\quad + \frac{1}{2} \sum_{i=1}^m \left\{ \frac{n_i (y_i - \sum_{j=1}^{n_i} \mathbf{s}_i^T \Upsilon x_{ij})^2}{((\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2)^2} \right\}, \\ \frac{\partial l(\Psi)}{\partial \Gamma} &= -\frac{1}{2} \sum_{i=1}^m \left\{ \frac{(\sum_{j=1}^{n_i} \mathbf{x}_{ij})(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T)}{(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2} \right. \\ &\quad - \frac{(y_i - \sum_{j=1}^{n_i} \mathbf{s}_i^T \Upsilon x_{ij})^2}{((\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2)^2} \\ &\quad \left. \times \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \right\}, \end{aligned} \tag{A1}$$

Based on the first derivatives, the second derivatives can be further obtained:

$$\frac{\partial l(\Psi)^2}{\partial \Upsilon \partial \Upsilon} = -\sum_{i=1}^m \frac{1}{(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2}$$

$$\begin{aligned} &\times \left(\sum_{j=1}^{n_i} \mathbf{s}_i \mathbf{x}_{ij}^T \right) \otimes \left(\sum_{j=1}^{n_i} \mathbf{s}_i \mathbf{x}_{ij}^T \right), \\ \frac{\partial l(\Psi)^2}{\partial \Upsilon \partial \sigma^2} &= -\sum_{i=1}^m n_i \frac{(\sum_{j=1}^{n_i} \mathbf{s}_i \mathbf{x}_{ij}^T)(y_i - \sum_{j=1}^{n_i} \mathbf{s}_i^T \Upsilon x_{ij})}{((\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2)^2}, \\ \frac{\partial l(\Psi)^2}{\partial \Upsilon \partial \Gamma} &= -\sum_{i=1}^m \left\{ \frac{(y_i - \sum_{j=1}^{n_i} \mathbf{s}_i^T \Upsilon x_{ij})}{((\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2)^2} \right. \\ &\quad \left. \times \left(\sum_{j=1}^{n_i} \mathbf{s}_i \mathbf{x}_{ij}^T \right) \otimes \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \right\}, \\ \frac{\partial l(\Psi)^2}{\partial \sigma^2 \partial \sigma^2} &= \frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{((\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2)^2} \\ &\quad - \sum_{i=1}^m \left\{ n_i^2 \frac{(y_i - \sum_{j=1}^{n_i} \mathbf{s}_i^T \Upsilon x_{ij})^2}{((\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2)^3} \right\}, \\ \frac{\partial l(\Psi)^2}{\partial \sigma^2 \partial \Gamma} &= \frac{1}{2} \sum_{i=1}^m n_i^2 \left\{ \frac{(\sum_{j=1}^{n_i} \mathbf{x}_{ij})(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T)}{((\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2)^2} \right. \\ &\quad \left. - \frac{2(y_i - \sum_{j=1}^{n_i} \mathbf{s}_i^T \Upsilon x_{ij})^2 (\sum_{j=1}^{n_i} \mathbf{x}_{ij})(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T)}{((\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2)^3} \right\}, \\ \frac{\partial l(\Psi)^2}{\partial \Gamma \partial \Gamma} &= \frac{1}{2} \sum_{i=1}^m \left\{ \frac{1}{((\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2)^2} \right. \\ &\quad \times \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \otimes \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \\ &\quad - 2 \frac{(y_i - \sum_{j=1}^{n_i} \mathbf{s}_i^T \Upsilon x_{ij})^2}{((\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2)^3} \\ &\quad \left. \times \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \otimes \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij} \right) \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T \right) \right\}, \end{aligned}$$

where \otimes is the Kronecker operator (Demidenko, 2004). Next, we calculate $E\partial l(\Psi)^2/\partial \Psi \partial \Psi$ with respect to $y_i, i = 1, \dots, m$, which gives

$$\begin{aligned} E \left(\frac{\partial l(\Psi)^2}{\partial \Upsilon \partial \Upsilon} \right) &= -\sum_{i=1}^m \frac{1}{(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T) \Gamma(\sum_{j=1}^{n_i} \mathbf{x}_{ij}) + n_i \sigma^2} \\ &\quad \times \left(\sum_{j=1}^{n_i} \mathbf{s}_i \mathbf{x}_{ij}^T \right) \otimes \left(\sum_{j=1}^{n_i} \mathbf{s}_i \mathbf{x}_{ij}^T \right), \\ E \left(\frac{\partial l(\Psi)^2}{\partial \Upsilon \partial \sigma^2} \right) &= 0, \end{aligned}$$

$$\begin{aligned}
E\left(\frac{\partial l(\Psi)^2}{\partial \Upsilon \partial \Gamma}\right) &= 0, \\
E\left(\frac{\partial l(\Psi)^2}{\partial \sigma^2 \partial \sigma^2}\right) &= -\frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{\left(\left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T\right) \Gamma \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}\right) + n_i \sigma^2\right)^2}, \\
E\left(\frac{\partial l(\Psi)^2}{\partial \sigma^2 \partial \Gamma}\right) &= -\frac{1}{2} \sum_{i=1}^m n_i^2 \frac{\left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}\right) \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T\right)}{\left(\left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T\right) \Gamma \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}\right) + n_i \sigma^2\right)^2}, \\
E\left(\frac{\partial l(\Psi)^2}{\partial \Gamma \partial \Gamma}\right) &= -\frac{1}{2} \sum_{i=1}^m \left\{ \frac{1}{\left(\left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T\right) \Gamma \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}\right) + n_i \sigma^2\right)^2} \right. \\
&\quad \times \left. \begin{aligned} &\left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}\right) \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T\right) \otimes \left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}\right) \\ &\times \left.\left(\sum_{j=1}^{n_i} \mathbf{x}_{ij}^T\right) \right\}. \quad (\text{A2})
\end{aligned}
\right.
\end{aligned}$$

Biographies

Shuai Huang is an Assistant Professor in Industrial and Management Systems Engineering at the University of South Florida. He received his B.S. in Statistics from the University of Science and Technology of China in 2007 and his Ph.D. in Industrial Engineering from the Arizona State University in 2012. His research interests are statistical learning and data mining with applications in healthcare and manufacturing.

Jing Li is an Associate Professor in Industrial Engineering at Arizona State University. She received an M.A. in Statistics and a Ph.D. in Industrial and Operations Engineering from the University of Michigan in 2005 and 2007, respectively. Her research interests are applied statistics and data mining for modeling, analysis, and control of data-rich complex systems. The application domains of her research include production systems, healthcare, and communication networks. Her research is sponsored by the NSF, NIH, DOD, and Arizona State. She is an NSF CAREER Awardee. She is also a recipient of the Best Paper Award from the Industrial Engineering Research Conference (twice). She has been the Chair for Data Mining Subdivision of INFORMS, an Associate Ed-

itor for *Journal of Chinese Institute of Industrial Engineers*, and a Guest Editor for *IIE Transactions* and *Quality and Reliability Engineering*.

Gerri Lamb is an Associate Professor in the College of Nursing and Health Innovation and the Herberger Institute for Design and the Arts at Arizona State University. She was the Principal Investigator for the Nurse-Sensitive Measurement of Hospital Care Coordination grant funded by The Robert Wood Johnson Foundation that generated the data and framework used in this research. She is deeply involved in numerous national quality initiatives focused on care coordination with the National Quality Forum the American Board of Internal Medicine, the Agency for Health Care Research and Quality, the American Academy of Nursing, and the Nursing Alliance for Quality Care.

Madeline H. Schmitt is Professor Emerita at the University of Rochester, School of Nursing in Rochester, New York. She was a major consultant to the Nurse-Sensitive Measurement of Hospital Care Coordination grant, including activities related to this analysis. She is a nurse-sociologist who, since the 1970s, has focused her scholarship on outcomes of inter-professional collaborative practice models and inter-professional education. She was one of two U.S. members for the WHO Task Force 2010 report *Framework for Action in Inter-professional Education and Collaborative Practice*. In 2010–2011, she chaired the expert panel for *Core Competencies for Inter-professional Collaborative Practice*, a report commissioned by six national professional education associations (AACN, AACOM, AACP, AAMC, ADEA, ASPH). She is an Editor Emerita of the *Journal of Interprofessional Care* and remains active in consultation, research, and publication, as well as limited teaching, with regard to outcomes of collaborative care and inter-professional education.

John Fowler is the Motorola Professor and Chair of the Supply Chain Management department at Arizona State University (ASU). He is also a Professor of Industrial Engineering (IE) and was previously the program chair for IE at ASU. His research interests include discrete event simulation, deterministic scheduling, and multi-criteria decision making. He has published over 95 journal articles and over 100 conference papers. He was the Program Chair for the 2002 and 2008 Industrial Engineering Research Conferences and the 2008 Winter Simulation Conference. He is currently serving as Editor-in-Chief for a new Institute of Industrial Engineers (IIE) journal focused on healthcare delivery systems entitled *IIE Transactions on Healthcare Systems Engineering*. He is also an Editor of the *Journal of Simulation*. He is a Fellow of the Institute of Industrial Engineers and currently serves as the IIE Vice President for Continuing Education, is a former INFORMS Vice President, and is an SCS representative on the Winter Simulation Conference Board of Directors.

Copyright of IIE Transactions is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.