METHODS

Using Multilevel Analysis in Patient and Organizational Outcomes Research

Sung-Hyun Cho

► ***Background:*** Outcomes research often compares patient and organizational outcomes across institutions, dealing with variables measured at different hierarchical levels. A traditional approach to analyzing multilevel data has been to aggregate individual-level variables at the institutional level.

► ***Objectives:*** To introduce the conceptual and statistical background of multilevel analysis and provide an example of multilevel analysis that was used to examine the relationship between nurse staffing and patient outcome.

► ***Methods:*** A two-level model was presented employing multilevel logistic regression analysis.

► ***Results:*** Outputs from multilevel analysis were interpreted. Other statistics were presented for model specification and testing.

► ***Conclusion:*** Researchers should consider multilevel modeling at the study design stage to select theoretically and statistically sound research methods.

► ***Key Words:*** hierarchical model • multilevel analysis • outcomes research

During the past decade, multilevel analysis has emerged as an analytical strategy in social and behavioral sciences and public health (Diez-Roux, 2000; Raudenbush & Bryk, 2002).. Using multilevel modeling allows researchers to examine simultaneously the effect of individual-level as well as group-level predictors on the dependent variable of interest. In nursing literature, Wu (1997) introduced multilevel linear models for meta-analysis in nursing research. Recently, Whitman, Davidson, Sereika, and Rudy (2001) applied a hierarchical longitudinal linear model (multilevel model), in examining the relationship between nurse staffing and the use of restraint. A growing use of multilevel analysis is expected in nursing research, especially in patient and organizational outcomes research. Outcomes research often compares patient and organizational outcomes across institutions. Data sets used in these outcomes studies commonly have a hierarchical or tiered-level structure. For example, a database that includes patients discharged across care units and hospitals may have a three-level data structure: patient (level 1), care unit (level 2), and hospital (level 3). This article introduces the conceptual and statistical background of multilevel analysis and provide an example of multilevel analysis that examines the relationship between nurse staffing and adverse patient outcomes (i.e., pneumonia).

Conceptual and Statistical Background of Multilevel Analysis

Multilevel models refer to analytic models that contain variables measured at different levels of the hierarchy (Kreft & De Leeuw, 1998). Research propositions tested in multilevel analysis refer to relationships between variables that are measured at different hierarchical levels. Because multilevel models acknowledge hierarchical data, researchers should not move aggregation or disaggregation variables to a single level (Hox, 1995). Thus, there are conceptual and statistical advantages in multilevel analysis; variables are analyzed at the level that they were defined and measured. For example, if job satisfaction is conceptualized and measured at the individual nurse level, it is theoretically correct to analyze the variable at the nurse level, not at a higher level (e.g., care unit, hospital). These advantages are particularly meaningful for comparing patient outcomes across hospitals because risk adjustment can be conducted at the patient level without aggregating risk factors at the hospital level.

The traditional approach to dealing with multilevel data in nursing outcomes research is to aggregate individual-level variables at the higher level. For example, patients who had pneumonia after surgery are aggregated at the hospital level by calculating an overall pneumonia rate for each hospital. Patient-level predictors (e.g., age, severity of illness) are also aggregated at the hospital level. Hos-

Sung-Hyun Cho, PhD, MPH, RN, is a researcher at Korea Institute for Health and Social Affairs, Seoul, Korea.

pital characteristics are then included in a multivariate model with the hospital as the unit of analysis. The literature reports several problems related to data aggregation. The ecological fallacy, which occurs when relationships between variables are examined using data aggregated at the group level, but conclusions are drawn at the individual level. Robinson (1950) concluded that correlations between aggregated variables cannot be used as substitutes for individual correlations. For example, when the mean age of patients in each hospital has an association with the overall pneumonia rate of hospitals, this relationship does not allow any inferences about the effect of age on the occurrence of pneumonia at the patient level. In this example, the problem of "shift of meaning" also arises (Snijders & Bosker, 1999). The shift of meaning holds that when a variable of individuals is aggregated to the group level, the meaning of the variable does not directly refer to the individual, but rather to the group. The mean age of patients may be an implication for the patient population of the hospital; and further, this meaning may be distinct from that of age at an individual level. The statistical issue may be another potential problem of data aggregation. In this instance, the process of aggregating to the higher level may inflate the estimates of the true relationship between variables because aggregated data eliminates within-hospital variance (Kreft & De Leew, 1998).

An alternative approach to nested data would be using a single regression model with patients as the unit of analysis, including dummy variables for hospitals in the model. This model, however, ignores clustering of patients and may exaggerate the precision of estimates. It also does not allow for inclusion of hospital-level variables (e.g., nurse staffing) (Gatsonis, Epstein, Newhouse, Normand, & McNeil, 1995).

Multilevel modeling is applicable to binary and count outcome variables as well as continuous variables. There are several statistical packages available to conduct multilevel analysis (i.e., HLM, Scientific Software International, Lincolnwood, IL; MLn, Institute of Education, London, UK; VARCL, iec ProGAMMA, Netherlands; MIXREG, MIXOR, & MIXNO, University of Illinois, Chicago, IL; SAS procedure MIXED & SAS GLIMMIX macro, SAS Institute, Cary, NC; and BMDP5-V, Statistical Solutions Ltd., Ireland, UK) (Snijders & Bosker, 1999). The HLM (version 5), developed by Raudenbush and his colleagues (2001), was used in preparing the example for this article.

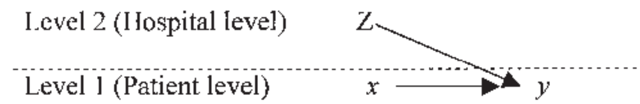

FIGURE 1. The structure of a multilevel proposition. x = patient characteristics (age, type of admission, number of diagnoses); y = probability of developing pneumonia; Z = hospital characteristics (nurse staffing, ownership).

An Example: Multilevel Logistic Model for Nurse Staffing and Pneumonia

Previous studies reported an inverse relationship between nurse staffing and pneumonia (American Nurses Association, 2002; Kovner & Gergen, 1998.. The example introduced here examines the probability of pneumonia among hospitals as it relates to nurse staffing. It is hypothesized that providing nursing hours per patient lowers the probability that a surgery patient will have pneumonia during hospitalization. This hypothesis contains the structure of a multilevel proposition (Figure 1). The hypothesis examines the effect of Level 2 variable Z (hospital characteristics) on the Level 1 variable y (pneumonia) while controlling for the Level 1 variable x (patient characteristics). The ownership variable was included as a hospital characteristic in the assumption that quality of care may differ by hospital characteristics (Keeler et al., 1992). Patient characteristics are also adjusted to isolate the effects of nurse staffing on pneumonia.

This example has been made artificially simple to illustrate the basics of multilevel modeling. Certain types of hospitals and patient groups were purposely selected to make the sample homogenous. The illustrative data include patients discharged in 1997 from 48 acute care, nonteaching, and nonrural California hospitals with 200–299 beds. Using diagnosis-related groups (DRGs), patients (N = 5,376) who had major small and large bowel (DRG 148, 149) or stomach, esophageal, and duodenal (DRG 154, 155) procedures were selected.

As presented in Figure 1, the dataset is assumed to have a two-level data structure: Level 1 (patient) and Level 2 (hospital). Because the outcome variable of pneumonia was defined as dichotomous, a multilevel logistic model was employed. The multilevel logistic model examines separate logistic regression models in each hospital, then examines the relation between the parameters of these models and hospital characteristics including nurse staffing. Thus, this multilevel regression decomposes the total variances into within-hospital and between-hospital components.

Level 1 Model

The Level 1 model is specified to compare patients with pneumonia who were discharged from the same hospitals (Equation 1). Here $PNEU_{ij}$ is the response for a patient i in hospital j.

$$\text{Logit}\ (PNEU_{ij}) = \beta_{0j} + \beta_{1j}AGE_{ij} + \beta_{2j}ADMIT_{ij} + \beta_{3j}NDX_{ij} \quad (1)$$

The dependent variable pneumonia (PNEU) distinguishes between those who had (1) and did not have pneumonia (0). Within each hospital, the probability of pneumonia for a patient is modeled as a function of patient age (AGE), type of admission (ADMIT), and the number of diagnoses at admission (NDX). To make the intercept of the regression line meaningful, AGE and NDX were centered to the grand mean of the whole sample. Type of admission (1 = scheduled, 0 = unscheduled) is a binomial variable. An average patient is defined as one that has zero values for the three variables (AGE, ADMIT, NDX) and that, in other words, is

TABLE 1. Estimates for Multilevel Logistic Regression of Pneumonia as a Function of Patient and Hospital Characteristics

	Model 1		Model 2		Model 3	
Fixed Effect	**Estimate**	**Standard Error**	**Estimate**	**Standard Error**	**Estimate**	**Standard Error**
γ_{00} = intercept	−3.1194*	.1173	−2.9275*	.1369	−2.8776*	.5366
γ_{10} = coefficient of AGE			0.0291*	.0052	0.0292*	.0052
γ_{20} = coefficient of ADMIT			−0.9765*	.1660	−0.9774*	.1676
γ_{30} = coefficient of NDX			0.0543*	.0204	0.0546*	.0205
γ_{01} = coefficient of OWN					0.1578	.2886
γ_{02} = coefficient of STAFF					−0.0146	.0873
Random Effect	**Variance Component**	**Standard Deviation**	**Variance Component**	**Standard Deviation**	**Variance Component**	**Standard Deviation**
τ_{00} = var(U_{0j}) intercept variance	.3551*	.5959	.3792*	.6158	.3747*	.6121
Deviance	11,705.93		11,592.17		11,591.67	
Number of parameters	2		5		7	

Note. AGE = patient age; ADMIT = type of admission; NDX = number of diagnoses at admission; OWN = hospital ownership; STAFF = nurse staffing.
*$p < .05$

with age at the grand mean, the number of diagnoses at the grand mean and an unscheduled admission. Thus, the intercept (β_{0j}) corresponds to the log odds of pneumonia that an average patient had pneumonia while in the hospital *j*.

Level 2 Model

The Level 2 model takes into account the differences between hospitals and explains these differences in terms of hospital characteristics. Equation 2 indicates that the intercepts from Level 1 (β_{0j}) were modeled as a function of hospital ownership (OWN) and nurse staffing (STAFF) with a random effect (U_{0j}). This modeling means that within-hospital intercepts of each hospital vary systematically with hospital ownership and nurse staffing. Ownership is included as a binomial variable (1 = investor-owned, 0 = nonprofit). Nurse staffing indicates nursing hours worked by registered nurses per patient day. Nurse staffing is considered to be a hospital-level variable in this model because, in most cases, the number of nursing hours provided to a specific patient is unknown. In cases where the number of nursing hours provided is known or it is possible to estimate, the staffing variable may be treated as a patient-level variable. In many studies, nursing hours were measured and analyzed at the care unit level.

$$\beta_{0j} = \gamma_{00} + \gamma_{01}OWN_j + \gamma_{02}STAFF_j + U_{0j}$$

$$\beta_{1j} = \gamma_{10}, \quad \beta_{2j} = \gamma_{20}, \quad \beta_{3j} = \gamma_{30} \quad (2)$$

To make the model simple, the other coefficients of the Level 1 (β_{1j}, β_{2j}, β_{3j}) were modeled as a fixed slope without random effect. If these slopes were assumed to be random, their random effects (U_{1j}, U_{2j}, U_{3j}) would be included in the model. This example also assumes that the slopes at Level 1 are not related to hospital characteristics. If researchers hypothesized that the relationships between patient characteristics and pneumonia differ by hospital characteristics, cross-level interaction effects (e.g., AGE*STAFF) should also be included in the model. In addition, if the number of patients for each hospital is assumed to affect the occurrence of pneumonia, it can be included in the model with a fixed effect. The variation in group size does not pose a problem for the application of multilevel analysis; however, the smaller groups will have a smaller influence on the results than the larger groups (Snijders & Bosker, 1999).

Combined Model

The combined model is described in Equation 3 by substituting Equation 2 into Equation 1. The first part of the equation, $\gamma_{00} + \gamma_{10}AGE_{ij} + \gamma_{20}ADMIT_{ij} + \gamma_{30}NDX_{ij} + \gamma_{01}OWN_j + \gamma_{02}STAFF_j$, is called the fixed part of the model and does not include random effect. γ_{00} is the average intercept and γ_{10}, γ_{20}, γ_{30}, γ_{01}, and γ_{02} are the average regression coefficients of the predictors. The rest of the model (U_{0j}) is the random part that only consists of random effect.

$$\text{Logit}(PNEU_{ij}) = \gamma_{00} + \gamma_{10}AGE_{ij} + \gamma_{20}ADMIT_{ij} + \gamma_{30}NDX_{ij} + \gamma_{01}OWN_j + \gamma_{02}STAFF_j + U_{0j} \quad (3)$$

Estimates for the combined model are presented in Table 1 as Model 3. In many respects a multilevel regression is interpreted in the same way as an ordinary regression (Paterson, 1991). The estimates for Model 3 show that γ_{00} (− 2.8776) is the expected log-odds of pneumonia for an average patient as previously defined. This log-odds corresponds to a probability of $[e^{-2.8776}/(1 + e^{-2.8776})] = .053$ (Allison, 1999). The fixed effects of all Level 1 predictors are statistically significant. For example, the coefficient of ADMIT (γ_{20} = −.9774) indicates that the estimated odds of pneumonia for a patient with scheduled admission is about 38% ($e^{-.9774}$= .376) of the odds of pneumonia for a patient with unscheduled admission. Although the two hospital predictors at Level 2 did not reach statistical significance, patients discharged from investor-owned hospitals have a greater tendency to present with pneumonia. Nurse staffing, which is the variable of interest, appeared to have a negative fixed effect on the occurrence of pneumonia with the coefficient of γ_{02} (−.0146).

Model Specification and Testing

Like other multivariate analyses, multilevel models should be specified based on theory and the hypotheses to be tested. Some statistics may help for model specification and testing. Table 1 compares three different models that include different sets of random and fixed effects. Model 1 includes a random intercept without predictors, which is called an empty model. Model 2 includes three Level 1 patient characteristics without hospital characteristics, and Model 3 is specified as Equation 3.

A hypothesis test of random variance is useful to assess the presence of hierarchically structured data. The result from Model 1 (Table 1) shows that the null hypothesis (H_0: τ_{00} = 0) is rejected, suggesting that some significant covariance exists between patients in the same hospitals. With the existence of covariance, applying traditional regression models violates the assumption of independence of observations and increases type I error (Kreft & De Leew, 1998). Another way to examine clustered data is to compute intraclass correlation, which is a measure of the degree of dependence of patients belonging to the same hospital (Kreft & De Leew, 1998). A greater intraclass correlation indicates that patients are more likely to share common experiences (i.e., pneumonia). The intraclass correlation can be also interpreted as the fraction of total variability that is due to the hospital level (Kreft & De Leew, 1998). There are different ways to define intraclass correlation for

Outcomes research often compares patient and organizational outcomes across institutions

multilevel logistic models. One definition is presented in Equation (Snijders & Bosker, 1999). In a multilevel logistic model, σ^2 is fixed to $\pi^2/3$ (= 3.29). Using this formula, the intraclass correlation coefficient of Model 1 is .097. Introducing three Level 1 variables (Model 2) increased the intraclass correlation to .103.

$$\rho = \tau_{00} / (\tau_{00} + \sigma^2) = \tau_{00} / (\tau_{00} + \pi^2/3) \quad (4)$$

where ρ = intraclass correlation coefficient, τ_{00} = between-hospital variance, σ^2 = within-hospital variance.

The deviance test can be used as a method to test the model statistically. Deviance is defined as minus twice the natural logarithm of the likelihood and can be regarded as a measure of lack of fit between the model and the data (Snijders & Bosker, 1999). HLM (version 5) produces the deviance statistics using Laplace approximation that is one of the computational algorithms for binary outcome models (Raudenbush & Bryk, 2002; Raudenbush, Bryk, Cheong, & Congdon, 2001). The deviance cannot be interpreted directly, but rather is compared between models that are fit to the same data set. Suppose that two models (M_0, M_1) have deviances by D_0 and D_1, with m_0 and m_1 parameters, respectively. The difference of the deviance (D_0 - D_1) can be used as a test statistic having a chi-squared distribution with ($m_1 - m_0$) degrees of freedom. For instance, compared to Model 1, Model 2 shows better fit, having ($D_0 - D_1$) = 113.7596 with ($m_1 - m_0$) = 3, P <.001. Model 3 has the deviance statistic very close to that of Model 2 with an insignificant chi-squared test. This result may suggest that nurse staffing and ownership are not good predictors to explain the variations in the occurrence of pneumonia between hospitals. Researchers may prefer Model 2, which is more parsimonious than Model 3, as long as it incorporates into the theory.

Models 2 and 3 indicate that patient characteristics (AGE, ADMIT, NDX) had strong relationships with pneumonia. Contrary to previous studies, results from Model 3 did not support the hypothesis that nurse staffing would have an inverse relationship with pneumonia. This inconsistent finding may be attributed to the use of different patient groups or statistical analysis techniques. Further analysis is needed to compare results from traditional regression models with aggregated data and those from multilevel models.

This report supports the use of multilevel analysis when dealing with variables measured at different hierarchical levels. Consideration of multilevel modeling at the study design stage may help researchers select theoretically and statistically sound research methods.

Accepted August 14, 2002.

This project was supported by grant number R03 HS11397 from the Agency for Healthcare Research and Quality.

Corresponding author: Sung-Hyun Cho, PhD, MPH, RN, Korea Institute for Health and Social Affairs, Department of Health Policy, San 42-14 Bulgwang-dong Eunpyeong-gu, Seoul, Korea 122-705 (e-mail: shcho@kihasa.re.kr).

References

Allison, P. D. (1999). *Logistic regression using the SAS system: Theory and application.* Cary, NC: SAS Institute.

American Nurses Association. (2000). *Nurse staffing and patient outcomes in the inpatient hospital setting.* Washington, DC: American Nurses Publishing.

Diez-Roux, A. V. (2000). Multilevel analysis in public health research. *Annual Review of Public Health, 21*, 171-192.

Gatsonis, C., Epstein, A. M., Newhouse, J. P., Normand, S. L., & McNeil, B. J. (1995). Variations in the utilization of coronary angiography for elderly patients with an acute myocardial infarction: An analysis using hierarchical logistic regression. *Medical Care, 33*(6), 625-642.

Hox, J. J. (1995). *Applied multilevel Analysis.* Amsterdam: TT-Publikaties.

Keeler, E. B., Rubenstein, L. V., Kahn, K. L., Draper, D., Harrison, E. R., McGinty, M. J., et al. (1992). Hospital characteristics and quality of care. *JAMA, 268*(13), 1709-1714.

Kovner, C., & Gergen, P. J. (1998). Nurse staffing levels and adverse events following surgery in U.S. hospitals. *Image: Journal of Nursing Scholarship, 30*(4), 315-321.

Kreft, I., & De Leeuw, J. (1998). *Introducing multilevel modeling.* London: Sage Publications.

Paterson, L. (1991). An introduction to multilevel modelling. In S. W. Raudenbush & J. D. Willms (Eds.), *Schools, classrooms, and pupils: International studies of schooling from a multilevel perspective* (pp. 13-24). San Diego: Academic Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2001). *HLM5: Hierarchical linear and nonlinear modeling.* Lincolnwood: Scientific Software International.

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

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* London: Sage Publications.

Whitman, G. R., Davidson, L. J., Sereika, S. M., & Rudy, E. B. (2001). Staffing and pattern of mechanical restraint use across a multiple hospital system. *Nursing Research, 50*(6), 356-362.

Wu, Y. W. (1997). An application of hierarchical linear models to meta-analysis in nursing research. *Nursing Research, 46*, 295-298.