Comparisons of Modeling Approaches for Evaluating the Longitudinal Association in a Clustered Healthcare Intervention Study

Yulan Liang*

Department of Family and Community Health, University of Maryland, Baltimore, MD, USA

Abstract

This paper addresses methodology issues related to evidence-based healthcare research, specifically when evaluating and analyzing the hospital practice environments (HPE) impacts on the patient health outcomes are conducted in longitudinal intervention survey studies. HPE include the spatially clustered hospital characteristics, including practice environment scale (PES) measures, hospital facilities, nursing staffing and nursing attributes. The longitudinal associations between HPE and patient smoking cessation counseling (SCC) activities, and patient heart failure (HF) outcomes are examined. Various longitudinal and hierarchical modeling are compared including linear mixed models with restricted maximum likelihood estimation, generalized estimating equations with quasi-likelihood estimation, hierarchical linear regression models with nonparametric generalized least squares estimations, and repeated ANOVA. Moreover, both pre-modeling including the items/dimension reduction issues for longitudinal item-response hospital survey data and post-modeling (the mediation analysis) are discussed and conducted. Results show some methodology and solution differences when including the spatial or temporal correlations of HPE simultaneously for examining the longitudinal effects of HPE on HF core outcome measures adjusted or potentially mediated by SCC and nurse staffing environmental variables. This may have implications and potential impact for healthcare decision-making. Patients can benefit from these research findings.

Keywords: Evidence based healthcare research; Spatial-temporal hospital practice environments; Longitudinal survey data; Patient heart failure outcomes; Linear mixed models; Generalized estimating equations

Introduction

The evidence based healthcare and medical researches, and implementation sciences have become more and more popular in healthcare, nursing, medicine, and public health fields, which is evident from the growing number of research articles in evidence based medicine and evidence based practice in healthcare [1-4]. Meanwhile, some barriers and challenges have risen in methodological points of view due to the fast growing big healthcare data resulting from mobile technology and electronic health record applications [5-7]. These methodology challenges may result at various data handling stages, including 1) longitudinal item-response survey data with redundancy, noise/outliers, measurement errors and missing/dropout; 2) multiple weak signals or small to moderate effects aggregated over time, and spatially clustered or correlated with intrinsic heterogeneity; 3) multivariate and multilevel correlated characteristics with various confounders; 4) the mixed quantitative and qualitative hospital data structures; 5) methodology differences for handling both unstructured or semi-structured qualitative healthcare information. These may in turn end up with inconsistent or potential non-producible or contradicting results [5-6]. These barriers eventually may prevent the research outcomes to be put into practice for healthcare and policy decision-making that patients could benefit from.

The purpose of this study is to examine the methodology issues and solution differences when evaluating the longitudinal associations of the hospital practice environment (HPE) and patient health outcomes (smoking cessation counseling (SCC) activities; heart failure (HF)) adjusted by other hospital characteristics (e.g., nurses attributes and nursing staffing variables) in the healthcare setting (e.g., rural hospitals). Specifically, through comparison of various statistical models and associated algorithms, either marginal population average models or hierarchical individualized conditional modeling, we examine the

potential issues and demonstrate the solution differences for analyzing the longitudinal survey intervention studies. In section 2, we present an example of randomized controlled trial of HPE intervention study from rural hospitals. The sample characteristics, and baseline information prior to the intervention are presented. In section 3, various statistical modeling approaches are conducted, which include hierarchical linear mixed model with restricted maximum likelihood estimation (REML), generalized estimating equations (GEE) with quasi-likelihood estimation, hierarchical linear regression model with nonparametric least squares estimations, repeated ANOVA. Both pre data/item-response dimension reductions and post mediation analysis were conducted to further examine some related challenge issues involved in the longitudinal multilevel hospital environmental factors linked to the patient outcomes. The results are presented in section 4, and some discussions are provided in section 5.

Methods

Examples of the multilevel longitudinal hospital practice environment on patient outcome study

Rural hospitals represent forty-one percent of nonfederal, short-term general and specialty hospitals, and provided care for nineteen percent of all discharges nationally in 2006 [8]. Patient populations are older with poor or fair health, more likely to be uninsured, and are more remote from health care services. This example is drawn from medical

Corresponding author: Yulan Liang, Department of Family and Community Health, University of Maryland, Baltimore, 655 W. Lombard Street, Baltimore, MD 21201-1579, USA, Tel: 410-706-4812; E-mail: liang@son.umaryland.edu

surgical registered nurses (RNs) that care for HF patients (N=591 RNs) from twenty-three rural hospitals in the eastern U.S [8-10]. The intervention study is a randomized, controlled trial with two cohorts: intervention group and control groups. The intervention includes two-day onsite HF collaborative training by 1) access to experts through shared practice strategies and benchmark 2) HF tool kit; 3) monthly conference calls (8). Two survey's data were collected after intervention for measuring the hospital quality and organizational context, which includes practice environment scale (PES) measures and smoking cessation counseling scale (SCC) measures. PES survey includes 31 items-responses, measured using Likert scales from 1-4 [11-13]; while the SCC survey includes 24 items-responses measured on a scale of 1-4 (Newhouse et al.). The nurse staffing attributes include RNs skill mix, RN hours per patient day (RN HPPD), RN turnover, and nursing characteristics such as work full or part time, highest degree, gender, age, ethnicity are also collected during the surveys (see Table 1 for descriptive statistics) [9,10]. Patient HF outcomes include four core HF measures: Compliance Discharge Instructions, Compliance Smoking Cessation, Compliance LFV, and Compliance with ACE Inhibitors which are repeatedly surveyed in seven consecutive quarters starting from September 2007 to March 2009.

Statistically this is a structural/panel longitudinal intervention survey example with both spatial (twenty-three rural hospitals) and temporal (repeated measured over time) information included, and it is complex design with more than one response from more than one cohort, from two-level hospital-nurse/patient survey data [14]. At the high level, practice environment scale (PES) and SCC from twenty-three rural hospitals (n=23) from the eastern U.S region are included. At the low level 1) RNs (N=591) characteristics and nursing staffing attributes are included; 2) HF outcomes measures from HF patients cared by RNs. In order to examine the longitudinal/temporal and spatial/hospital site effects of hospital PES, nurse staffing, nursing attributes on SCC activities and HF core measures in addition to group/cohort/intervention effect, the following methodology issues need be considered.

Dimension reduction issues for longitudinal item-response hospital survey data

When modeling the longitudinal relationship of the multiple item response survey data on repeated patient outcomes (HF), the first step is to reduce the potentially redundant items-responses (e.g., 31 items for PES, 24 items for SCC) using one or more latent variables or hidden factors for the dimension reduction prior to modeling the interested relationships. In healthcare or health service research fields or psychology, social behavioral sciences, one frequently used dimension reduction method is the "Item Parceling" method [15,16], which is considered as a coarse variant of factor score regression [17]. Item parceling involves summing or averaging item scores from two or more items, and uses these parcel scores (or "scale score" in personality psychology) as observed "latent variable scores" to estimate and evaluate the relationships between latent variables [15].

In the example survey data, PES with 31 items and SCC with

Independent variable	Minimum	Maximum	Mean	Std. Deviation
RN Skill Mix	-2.60	2.30	0.01	0.99
RN HPPD	-0.55	1.35	-0.14	0.27
RN Turnover	-0.48	1.19	-0.10	0.26
PES	-1.46	2.37	0.05	1.00

Table 1: Descriptive statistics of nurse staffing attributes and PES parcel score/index from rural hospitals.

24 items measured from scale 1-4, item parceling is employed by aggregating and summing the responses of the items with a summary index to create/measure the overall score of PES and SCC, respectively. The rationale for the use of item parcels in the follow-up modeling approaches were as follows: (1) the reliability of item parcels will be greater than the raw scales [16]; (2) when the data contain raw items that are non-normally distributed or/and coarsely categorized (e.g., roughly scaled from 1-4), item parcels based on a large number of items often can be regarded as normally distributed, and normal theory based maximum likelihood techniques are applicable to such data; (3) Item parceling reduces the number of variables/items in the follow up analysis in addition to the multicollinearity issues, thus also reduce the ratio of variables to subjects, which will lead to more stable estimates with better model fit than estimation using the raw items; (4) the variance shared between items (true score variance) is still preserved, while unshared variance (uniqueness) shrinks (5) the reduced correlated residuals or dual loading, the nuisance effects (not be of interest) are also minimized a follow-up model [15,18-20].

Moreover, five established PES composite sub-indexes were constructed and computed for 31 PES items-responses for item dimension reductions, which are Nursing Participation in Hospital Affairs (NPHA), Nursing Foundations for Quality of Care (NFQC), Nurse Manager Ability Leadership and Support of Nurses (NMALSN), Staffing and Resource Adequacy (SRA), and Collegial Nurse Physician Relations [13]. In addition, factor analysis was further conducted for item/dimension reductions for comparison purposes.

Modeling for multi-level hospital survey data with longitudinal intervention design

The typical repeated measurement or longitudinal data analytical methods are linear mixed effect models (LMMs) or GEE in addition to repeated ANOVA (treat time as a factor) and hierarchical linear models [20-25]. Different types of models can be distinguished depending on whether the latent variables are assumed for the time dimension (e.g., ANOVA with time as a factor) and/or for the outcome repeated measurement dimension, and whether it measures and assesses how the association between the various outcomes evolves over time [21]. The use of latent variables allows for more flexible data structures but also has important implications with respect to the interpretation of the various model parameters in order to understand the association between the evolutions of all outcomes.

In our example, data was collected in a form of longitudinal panel data, i.e., the same hospital was observed over seven consecutive quarters with eighteen months periods and the survey data measured from the same hospitals were correlated. In order to model PES relationships on SCC activities adjusted by nurse attributes covariates, and longitudinal effects of PES on HF core outcome measures adjusted nurse staffing environmental variables or potentially mediated by SCC intervention, various types of models could be considered to incorporate such complex features and data structures: 1) adjusted effect with confounding/modifiers included versus unadjusted effects; 2) multi-level or hierarchical models versus single/one level model; 3) mixed models versus fixed or random effect models; 4) conditional models versus marginal models.

For comparison purposes, the following statistical methods are conducted for examining the primary PES effect and nurse attributes on SCC outcomes. First, conditional LMMs are applied with the inclusion of both a fixed and a random component [26]. The fixed component estimates the PES and nurse attributes (work full or part time, highest

degree, gender, age, ethnicity) effects, whereas the random component estimates the heterogeneity across hospital clusters in the regression coefficient. Specifically, in the high level, hospital sites (n=23) are assumed with random effect to incorporate the spatial correlations; while at the low level, PES and nurse attributes (work full or part time, highest degree, gender, age, ethnicity) are assumed as fixed effects. LMMs are used to analyze changes in individual response means, and are therefore appropriate for the modeling and prediction of individual response profiles by accounting for individual-specific heterogeneity and differences. REML is applied for unbiased or less estimations of both random and fixed effects compared with maximum likelihood estimations.

Next, population average marginal GEEs with quasi-Likelihood estimation are employed for the example data. Similar to LMMs, at the low level, PES and nurse attributes (work full or part time, highest degree, gender, age, ethnicity) are treated with fixed effects while high level correlated hospitals are assumed with random effect. The temporal correlations are assumed with exchangeable correlation structures after examining various correlation structures (e.g., independent, unstructured, etc.). Estimating the correct working correlation matrix provides efficiency parameter estimates, but if it isn't correctly estimated, the model parameters from GEEs tend to be consistent [27]. Another advantage of GEEs is the robust standard errors estimates to account for the within-subject correlations for repeated measurement data. Note that the marginal GEEs are appropriate if the research focus is on population-average, and the mean response depends only on the covariates of interest, not on any random effects or previous responses like LMMs, which focus on individual variability. In addition, one level repeated ANOVA and hierarchical linear models were also employed for comparison purposes [28-31]. Both models simultaneously include time/temporal/longitudinal information and spatial/hospital as either independent variables or latent factors in the modeling process. Similarly, PES and nurse attributes (work full or part time, highest degree, gender, age, ethnicity) are treated as fixed effects while hospital sites are treated as random effect in the hierarchical linear model.

Post mediation analysis with GEE for PES effect on HF core measures adjusted by SCC intervention and nursing staff factors

The major benefit of mediation analysis is that it can efficiently model both the direct and indirect effects of the mediation process to gain precisions by pooling all information together, especially if missing data appeared, unbalanced, and correlated multilevel longitudinal data. LMMs and structure equation models (SEMs) are commonly used methods for mediation analysis [32]. For instance, Blood et al. have conducted a study that compare LMMs with SEMs for the mediation analysis, the study showed that both models fit well to mediation analysis. Power for both models was good with a sample of size of 250 and a small to medium effect size [33]. Bias did not substantially increase for either model when data were generated from distributions that were both skewed and kurtotic. In settings where the goal is to evaluate the overall effects, the LMMs excluding mediating variables appear to have good performance with respect to power, bias and coverage probability relative to the SEM. Here we conduct the mediation analysis using GEEs instead of LMMs given our focus on the population average effect instead of prediction of individual response profiles for healthcare setting. Moreover, marginal effects are easy to understand and interpret for informed decision making. In GEEs mediation analysis, PES is primary interest, SCC Intervention is a mediator, nursing staffing variables and time effects from quarter to

quarter are effect modifiers and confounders, and HF core measures are outcome variables.

Results

Reliability statistics

Reliability statistics for PES with 31 items measured on scale 1-4 using two ways mixed Cronbach's Alpha is 0.941, while one-way random model, Cronbach's intra correlation coefficient (ICC) Alpha is 0.939. Both indicate that the hospital sample survey means for PES items have high reliabilities of 0.94 as an estimate of its unknown population mean. Similarly, for SCCs with 24 items measured on a scale of 1-4, Cronbach's Alpha for ICC is 0.951, which indicates the consistency and agreement among the measured items and provides the justification of using aggregated sum/composite index measures or un-weighted composite measures as either predictors or the outcome measures for handing multicollinearity issues of items for the following up models.

Dimension reduction for longitudinal item-response survey data

Table 1 displays the descriptive statistics of PES summary/composite index and the nursing staffing attributes. Table 2 displays descriptive statistics of aggregated summary scores index of HF core measures at baseline. The means from both tables show much smaller than the corresponding standard deviations (SDs), which may indicate the small or weak effect and large variations with more heterogeneity appeared in the PES of the survey data as we discussed earlier. The comparison between the item/unit-level analyses and aggregated analyses suggests that the aggregated summary index reduced the variations in the item-response variables so that their associations in the follow up analysis became detectable. Moreover, five established/popular PES composite sub-indexes (NPHA, NFQC, NMALSN, SRA, and CNPR) are also employed to 31 PES items for item dimension reductions. However, the correlations among the five aggregated composite sub-index is medium to large (Pearson correlation 0.35 to 0.71, see Table 3), which prevents including these five composite sub-index measures simultaneously in the follow-up modeling approach for examining the relationships of PES on patient outcomes (e.g., HF or SCC) due to potential multicollinearity issues.

Factor analysis is further conducted for item/dimension reductions for comparison purposes, in which the latent factor scores can be estimated by linear combinations of the discrete observed item into a few extracted factors. For 31 PES items, the first six extracted factors/principal components explained 59.3% total PES variations, while the first factor explained majority (36%) variation and the rest each only small percentage of variance (6%, 6%, 4%, 4%, 3% respectively). Similarly, for 24 SCC items, 4 factors explained 70% SCC variations. Although factor score regression analysis can be considered with these calculated factor scores for inferring the relationships between these latent variables and other external outcome variables. The issues with it are difficulties of the interpretations for the newly constructed factors

Dependent variable	Minimum	Maximum	Mean	Std. Deviation
Compliance Discharge Instructions	-2.98	0.83	0.09	0.78
Compliance Smoking Cessation	-2.97	1.36	0.10	0.73
Compliance LFV	-2.86	0.56	0.11	0.72
Compliance with ACE Inhibitors	-2.24	0.74	0.08	0.83

Table 2: Descriptive statistics of patient composite heart failure outcome core measures at baseline.

	NPHA	NFQC	NMALSN	SRA	CNPR
Nursing Participation in Hospital Affairs	1	.710**	.691**	.501**	.486**
Nursing Foundations for Quality of Care	.710**	1	.635**	.531**	.517**
Nurse Manager Ability Leadership and Support of Nurses	.691**	.635**	1	.514**	.436**
Staffing and Resource Adequacy	.501**	.531**	.514**	1	.375**
Collegial Nurse Physician Relations	.486**	.517**	.436**	.375**	1
Note. ** p<0.01 (2-tailed).					

Table 3: The correlation matrix of five established PES composite sub-index summarized from 31 items.

Methods	Standardized estimated beta	95% CI
LMM (REML) (adjusted)	0.297	(0.178, 0.416)
GEE (adjusted)	0.283	(0.177, 0.389)
Hierarchical LM or repeated ANOVA (adjusted)	0.295	(0.179, 0.412)
Repeated ANOVA with Variance component model (unadjusted)	0.273	(0.161, 0.386)

Table 4: Comparisons of PES effect estimates from four methods on patient SCC activities adjusted by the nursing attributes.

and also potential bias issue [18]. The focus for this study is comparison of the longitudinal modeling from both spatially and temporary correlations survey data, to make the comparisons of the modeling approaches consistent and easy to interpret, the uni-dimensional summary index of both SCC and PES are applied for the follow-up modeling comparisons.

Comparisons of statistical methods/models for the effects of PES on SCC activities adjusted by nursing staff and hospital sites

Prior to examining the effects of PES and nursing attributes on SCC activities, the hospital sites/cluster effects are examined using variance component analysis to explore how outcomes vary with hospital (higher level) rather than nursing characteristics (lower level) [28]. The intra correlation coefficient of hospital site effect is 0.07, indicating seven percent of variance explained by the hospital sites in occurrence of SCC activities and there are small variations among the hospitals. As Cohen suggests, if the ICC is higher than five point nine percent, the cluster effect (repeat panel data) within the same aggregate unit (hospital) in estimating the regression model cannot be neglected [34].

The comparisons among LMMs, GEEs, Hierarchical linear regression and repeated ANOVA for PES on the SCC activities adjusted by the nursing attributes (work full or part time, highest degree, gender, age, ethnicity) also show the differences of the estimated effects and precisions (standard errors) (Table 4). Marginal GEEs take into account the population-averaged relationship, while LMMs express the relationships on inter-individuals via random effects. The differences between parameter estimates of the two models may depend on the between-individual heterogeneity over time, which are described by the random intercepts and random slopes (time-before, and time-after) variances in the LMM (random effects) [35-40]. This inter-individual heterogeneity shows the differences between the parameters estimate of the marginal GEE and the random/conditional effect LMMs. In our study, the effect of PES on SCC using multilevel LMMs, the beta is 0.297 (t=4.9, p<0.001), 95% CI: (0.178, 0.416). With marginal GEEs, PES effect beta is 0.283 (Wald t=27.4 p<0.001) 95% CI (0.177, 0.389).

Compared LMMs, GEEs with hierarchical linear regression model or repeated ANOVA, e.g., in the hierarchical linear model, high level hospital sites with random effects, PES and nurse attributes as fixed effects, the estimated PES effect on SCC adjusted by the nurse attributes (work full or part time, highest degree, gender, age, ethnicity), beta=0.295 (t=4.99 p<0.0001), 95% CI (0.179, 0.412), which is very similar to LMM estimates since both methods are conditional models

with inclusion of random effects for the heterogeneity in addition to the advantage of handling missing data. In the repeated ANOVA, PES explained 5.9% variance on SCC activity outcome adjusted by the nursing attributes. While PES only explained 4.9% variance on SCC without adjusted nursing attributes and the estimated PES is also smaller (beta=0.273, t=4.77, p<0.001, 95% CI (0.161, 0.386) compared with above adjusted effects (Table 4).

Besides above estimates differences, the interpretations among these compared models are also different. For instance, in marginal GEEs, the interpretation for the estimated regression coefficients represent population-averaged values and describes how the average values for the response variable may be changed in the studied population, and relates to the sub-population that includes the covariate comparing the sub-population not including that covariate. In the LMMs, the interpretation for the regression coefficient describes how any variables for any subject are subject to change and is for a person that has a covariate, when compared to the same person not having a covariate [41]. In our example, for GEEs, PES effect on SCC is 0.283 adjusted by the nurse attributes indicating that every one unit increase in PES, the average SCC will increase 0.283 for the studied population; while for LMMs, PES effect on SCC is 0.297, indicating one unit increases in PES, SCC increase 0.297 for the given subject after adjusting the covariates. These results show that the aggregated PES scores are associated with nurses who implement more smoking cessation interventions in addition to the significant hospital site/cluster differences. The nurse attributes (work full or part time, highest degree, gender, age, ethnicity) are not associated with more smoking cessation counseling (SCC) activities, which might be partially due to these nurse attributes being measured at the individual/low level with larger variations while PES scores are aggregated at the hospital/higher level.

The longitudinal effects of PES on heart failure core measures adjusted by nurse staffing variables

Since marginal GEEs have merits of easy understanding with population average effect interpretations, which also include the group/cluster effect adjusted with more precise estimates (smaller SE, less wide 95% CI) based on Table 4, we further conduct marginal GEEs to examine the longitudinal effects of PES on four HF core measures adjusted by nurse staffing variables (RNs skill mix, RN hours per patient day (RN HPPD), RN turnover). After testing various correlation structures including independent, exchangeable, AR (1) etc., AR (1) correlation is used for obtaining the estimates based on the best model selection criteria [27]. Table 5 displays the estimated effects. PES shows significant effect on Compliance Left Ventricular Function

	Compliance discharge instructions		Compliance smoking cessation		Compliance left ventricular function		Compliance with ACE inhibitors	
	Standardized beta	95% CI	Standardized beta	95% CI	Standardized beta	95% CI	standardized beta	95% CI
RN Skill Mix								
RN HPPD								
RN Turnover	-1.042**	(-1.777, -.307)	-1.148*	(-2.180, -.117)	-.893*	(-1.784, -.002)	-1.044*	(-1.820, -.269)
PES					.217**	(.054, .379)		

Note. Study periods from Sept., 07 to March 09; n=591 nurses; N=23 hospitals with GEE (AR (1) correlation).

**p<0.01; * p<0.05. Only significant effects with beta estimates and 95% CI were reported.

Table 5: PES effects on heart failure core measures (compliance discharge instructions, compliance smoking cessation, compliance left ventricular function, compliance with ACE Inhibitors) adjusted by of nurse staffing variables.

	Compliance discharge instructions		Compliance smoking cessation		Compliance left ventricular function		Compliance with ACE inhibitors	
	Standardized beta	95% CI	Standardized beta	95% CI	Standardized beta	95% CI	Standardized beta	95% CI
RN Skill Mix								
RN HPPD								
RN Turnover			-.724*	(-1.424, -.023)	-.501*	(-.921, -.081)		
Practice Environment (PES)			.133*	(.002, .264)	.262**	(.086, .437)		
SCC Intervention/Control								
Time (Quart7 vs. quart1)					.334**	(.156, .513)		
Time (Quart6 vs. quart1)								
Time (Quart5 vs. quart1)	.341*	(.014, .667)			.420**	(.171, .668)		
Time (Quart4 vs. quart1)								
Time (Quart3 vs. quart1)								
Time (Quart vs. quart1)								

Note. Study Periods (7 quarters From Sept., 07 (quarter 1) to March 09 (Quarter 7); n=591 nurses; N=23 hospitals)

**p<0.01 *p<0.05. Only significant effects with beta estimates and 95% CI are reported.

Table 6: Mediation analyses with GEE for longitudinal effects of PES and SCC on HF adjusted by context nurse staffing variables.

core measure (beta=0.217, 95% CI (.054, .379)). Moreover, nursing staffing variable, RN Turnover also shows significant negative effects on all four-core measures, indicating reductions of RN Turnover would have positive impacts on the patient HF outcomes.

Mediation analysis with GEE for PES effect on HF core measures mediated by SCC intervention and adjusted by nurse staffing

Although existing studies indicate adjusted LMMs appear to be the more suitable rather than the GEEs in some situations [41], we performed mediation analysis using GEE for PES effects on HF in order to be compared to above analysis results. In GEE mediation analysis, PES is our primary interest (rather than the intervention), SCC intervention serves as a mediator, nursing staffing variables are confounders, and HF core measures are outcome variables. Time effect from quarter to quarter is also included as effect modifier to help identify whether there are delayed, transient, or cumulative effects during/after the interventions [5]. Results shows that although SCC shows no effect on HF core measures, the estimated PES effects are slightly different than without SCC intervention mediator from 4.4, e.g., on Compliance Left Ventricular Function core measure, either with or without SCC, PES has significant effects (beta=0.262 versus beta=0.217); while with SCC mediator, PES shows also significant effects on Compliance Smoking Cessation (beta=0.133, 95% CI is (.002, .264) (Table 6). In addition, both PES and nursing Turnover found consistent associations on HF outcomes either with or without SCC intervention, both at the overall time trend and seasonality/quarterly levels.

Discussions and Conclusion

Mixed quantitative and qualitative healthcare data structures (continuous, ordinal/Likert scale, nominal) are common in healthcare and evidence based medical researches [35]. Analyzing mixed longitudinal hospital survey data is a challenge task due to some of the special features, e. g, multilevel correlated characteristics with heterogeneity and various confounders, small or weak aggregated effects on patient outcomes both spatially and temporally, and redundant, measurement errors/noises or missing data often appearing in the survey panel data. In this paper, several longitudinal models including LMMs, GEEs, hierarchical linear models, and repeated ANOVA are applied and compared to examine the methodology differences for the longitudinal relationships of the hospital practice environments (PES) on the patient outcomes adjusted by other hospital attributes/ characteristics in the healthcare intervention studies. The distinction between GEEs (marginal) and LMMs (conditional), and hierarchical linear regression and repeated ANOVA depends on the specific scientific question to be addressed with different interpretations. The results show the differences regarding the coefficient estimates and the associated standard errors, e.g., PES effects on SCC from the marginal GEE are smaller and precision is higher than those from multiple level conditional LMM and hierarchical LM adjusted for the hospital confounders. When the cluster effects are significant and non-ignorable, the longitudinal multi-level data with either LMM or GEE is appropriate with fewer samples, better power, and more precise estimates that incorporate both temporal correlations with repeated measurements and spatial correlations with clustered hospitals effects.

Moreover, despite some debates involved the item parceling

approach for the item-response reduction, we utilized the aggregated uni-dimensional un-weighted summary index for the longitudinal modeling approach comparisons. In addition, five popular hospital sub-index measures and factors analysis (weighted sums of individual items) are compared and discussed. The advantage with item parceling is its potentials for minimizing the errors and sources of variance from correlated residual or dual loading and diminishing the size of nuisance effects from item-level solution. In addition to handling the multicollinearity issues for the modeling approaches and eliminating the weak signal or redundant/duplicated effects of each item. Population effects in the item-level model might be made smaller in the parcel-level model [41,42]. Other hierarchical factor analysis for the construction the composite index will be explored in future work by evaluations of weighting algorithms in determining weights that take into account mixed quantitative and/or qualitative measures, and the multicollinearity issues. Other popular latent variable modeling for both longitudinal mediation analyses together with the hierarchical factor analysis will be compared, e.g., using the upper level and lower factor scores in a larger SEM or GEE. Last, but also important to mention that current dominant method for mixed quantitative and qualitative healthcare data structures (continuous, ordinal/Likert scale, nominal) is treating Likert scale data as continuous data rather than ordinal/categorical data. This is relevant to how we model such data either categorical data analysis methodology or classical continuous normal theory, and corresponding statistical solutions for longitudinal survey data.

References

1. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB (2000) Evidence-based medicine: How to practice and teach EBM 2. Churchill Livingstone, London.

2. Melnyk BM, Fineout-Overholt E (2005) Making the case for evidence-based practice. In: Melnyk BM, Fineout-Overholt E (eds.) Evidence-based practice in nursing & healthcare. A guide to best practice. Lippincot Williams & Wilkins, pp: 3-24.

3. Spring B, Hitchcock K (2010) Evidence-based practice in psychology. In: Weiner IB, Craighead WE (ed.) Corsini's Encyclopedia of Psychology and Behavioral Science, Wiley.

4. Peile E (2004) Reflections from medical practice: balancing evidence-based practice with practice based evidence. In: Thomas G, Pring R (eds.) Evidence-based Practice in Education. Open University Press, pp: 102-116.

5. Lynn P (2009) Methodology for longitudinal Surveys, Wiley.

6. Madigan D, Stang P, Berlin J, Schuemie M, Overhage J, et al. (2014) A Systematic Statistical Approach to Evaluating Evidence from Observational Studies. Annual Review of Statistics and Its Application 1: 11-39.

7. Seago JA, Williamson A, Atwood C (2006) Longitudinal analyses of nurse staffing and patient outcomes: more about failure to rescue. J Nurs Adm 36: 13-21.

8. Newhouse RP, Dennision C, Morlock M, Frick K, Pronovost P, et al. (2013) A Phased Cluster-randomized Trial of Rural Hospitals Testing a Quality Collaborative to Improve Heart Failure Care: Organizational Context Matters. Medical Care 51: 396-403.

9. Newhouse RP, Dennision C, Liang Y (2011) Psychometric Testing of the Smoking Cessation Counseling (SCC) Scale. Journal of Nursing Scholarship 43: 405-411.

10. Newhouse RP, Dennision C, Liang Y (2010) Evidence-based Smoking Cessation Counseling by Registered Nurses: Description and Predictors in Rural Hospitals. The American Nurse Today.

11. Flynn L, Liang Y, Dickson G, Aiken L (2010) Effects of the Nursing Practice Environment on Quality Outcomes in Nursing Homes. Journal of American Geriatric Society 58: 2401-2406.

12. Flynn L, Liang Y, Dickson G, Xie M, Boyer C, et al. (2012) Nurses Practice Environments, Error Interception Practices, and Inpatient Medication Errors. Journal of Nursing Scholarship 44: 180-186.

13. Aiken L, Clarke S, Sloane D, Sochalski J, Silber J (2002) Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. Journal of Nursing Scholarship 288: 1987-1993.

14. Mark BA, Salyer J (1999) Methodological issues in treatment effectiveness and outcomes research. Outcomes Manag Nurs Pract 3: 12-18.

15. Bandalos DL (2002) The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. Structural Equation Modeling 9: 78-102.

16. Kishton JM, Widaman KF (1994) Unidimensional versus domain representative parceling of questionnaire items: An empirical example. Educational & Psychological Measurement 54: 757-765.

17. Skrondal A, Laake P (2001) Regression among Factor Scores. Psychometrika 66: 563-576.

18. Hoshino T, Bentler P (2013) Bias in Factor Score Regression and a Simple Solution, Analysis of Mixed Data Methods & Applications Edited by Alexander R . de Leon and Keumhee Carrière Chough Chapman and Hall/CRC.

19. Thompson B, Melancon, JG (1996) Using item "Testlest"/"Parcels" in confirmatory factor analysis: An example using the PPSDQ-78 ED 404 349.

20. Little TD, Rhemtulla M, Gibson K, Schoemann AM (2013) Why the items versus parcels controversy needn't be one. Psychol Methods 18: 285-300.

21. Verbeke G, Fieuws S, Molenberghs G, Davidian M (2014) The analysis of multivariate longitudinal data: a review. Stat Methods Med Res 23: 42-59.

22. Fieuws S, Verbeke G (2004) Joint modelling of multivariate longitudinal profiles: Pitfalls of the random-effects approach. Stat Med 23: 3093-3104.

23. Fieuws S, Verbeke G (2006) Pairwise fitting of mixed models for the joint modelling of multivariate longitudinal profiles. Biometrics 62: 424-431.

24. Fieuws S, Verbeke G, Molenberghs G (2007) Random-effects models for multivariate repeated measures. Statistical Methods in Medical Research 16: 387-393.

25. Hubbard AE, Ahern J, Fleischer NL, Van der Laan M, Lippman SA, et al. (2010) To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. Epidemiology 21: 467-474.

26. Diggle P, Heagerty P, Liang K (2002) Analysis of longitudinal data. Oxford: Clarendon Press.

27. Cillessen AH1, Borch C (2006) Developmental trajectories of adolescent popularity: a growth curve modelling analysis. J Adolesc 29: 935-959.

28. Silber JH, Rosenbaum PR, Ross RN (1995) Comparing the contributions of groups of predictors: Which outcomes vary with hospital rather than patient characteristics. J Am Stat Assoc 90: 7-18.

29. Shin JH (2009) Application of repeated-measures analysis of variance and hierarchical linear model in nursing research. Nurs Res 58: 211-217.

30. Fox MT, Brathwaite AC, Sidani S (2004) Evaluating the effectiveness of interventions: exploration of two statistical methods. Can J Nurs Res 36: 20-30.

31. Adewale AJ, Hayduk L, Estabrooks CA, Cummings GG, Midodzi WK, et al. (2007) Applications of hierarchical linear models for evaluations of health interventions: demystifying the methods and interpretations of multilevel models. Nurs Res 56 (4): S40-46.

32. Skrondal A, Rabe-Hesketh S (2004) Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models. Boca Raton: Chapman & Hall/CRC.

33. Blood EA, Cabral H, Heeren T, Cheng DM (2010) Performance of mixed effects models in the analysis of mediated longitudinal data. BMC Med Res Methodol 10: 16.

34. Cohen J (1960) A coefficient for agreement for nominal scales, in Education and Psychological Measurement 20: 37-46.

35. Bauer DJ, Curran PJ (2004) The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. Psychological Methods 9: 3-29.

36. Carrière I, Bouyer J (2002) Choosing marginal or random-effects models for

longitudinal binary responses: application to self-reported disability among older persons. BMC Med Res Methodol 2: 15.

37. Faes C, Aerts M, Molenberghs G, Geys H, Teuns G, et al. (2008) A high-dimensional joint model for longitudinal outcomes of different nature. Stat Med 27: 4408-4427.

38. West BT (2009) Analyzing longitudinal data with the linear mixed models procedure in SPSS. Eval Health Prof 32: 207-228.

39. Omar RZ, Wright EM, Turner RM, Thompson SG (1999) Analysing repeated measurements data: a practical comparison of methods. Stat Med 18: 1587-1603.

40. Catyopadhyay S, Ganguli B, Chatterjee A (2011) A review of multivariate longitudinal data analysis. Statistical Methods in Medical Research 20: 299-330.

41. Aktas Samur A, Coskunfirat N, Saka O (2014) Comparison of predictor approaches for longitudinal binary outcomes: application to anesthesiology data. PeerJ 2: e648.

42. Nasser F, Wisenbaker J (2003) A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. Educational and Psychological Measurement 63: 729-757.