2012-01-01

Generalized Linear Latent Mixed Modeling Of Functional Independent Measures And Patient Outcomes

Maduranga Kasun Dassanayake
University of Texas at El Paso, mkdassanayake@miners.utep.edu

GENERALIZED LINEAR LATENT MIXED MODELING OF FUNCTIONAL INDEPENDENT MEASURES AND PATIENT OUTCOMES

MADURANGA KASUN DASSANAYAKE

Department of Mathematical Sciences

APPROVED:

Amy Wagler, Chair, Ph.D.

Joan Staniswalis, Ph.D.

Rebecca Reed-Jones, Ph.D.

Benjamin C. Flores, Ph.D
Interim Dean of the Graduate School

to my

MOTHER and FATHER

with love

GENERALIZED LINEAR LATENT MIXED MODELING OF FUNCTIONAL INDEPENDENT MEASURES AND PATIENT OUTCOMES

by

MADURANGA KASUN DASSANAYAKE

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Mathematical Sciences

THE UNIVERSITY OF TEXAS AT EL PASO

August 2012

Acknowledgements

At the very outset, I wish to place on record my deep gratitude to Dr. Amy Wagler; who gave inspiration to this thesis and as adviser, offered guidance, suggestions and advice throughout this task. I would also like to thank my committee members, Dr. Staniswalis and Dr. Reed-Jones for their valuable comments and suggestions. My heartiest appreciation should go to all the lectures in the Department of Statistics who have heralded a new wave of knowledge and understanding in me. Their kind support and guidance not only during this thesis, throughout my stay in the university, have in many ways, modeled me into who I am today. Last but not least, I express deepest gratitude to my family and friends who were always with me in need and helped me in many ways.

Abstract

The Functional Independent Measure (FIM) is one of the most widely accepted functional assessment measures used in the rehabilitation community. Past research studies have investigated the relationship between place of discharge, admission FIM scores or FIM difference scores, and patients' characteristics and found relationships between those variables. However, most of these studies fail to account for the multi-layered multidimensionality of the FIM and the measurement error associated with the FIM items. This study utilizes Generalized Linear Latent Mixed Models (GLLAMM) and Structural Equation Models (SEM) to assess which patient characteristics are associated with FIM difference scores and the structural relationship between admission FIM and place of discharge.

With regard to the models predicting place of discharge using FIM, it is found that orthopedic patients have at least a 50% chance of being discharged home if their mean cognition score is higher than 2.5. Similarly, stroke patients have 50% or more chance of being discharged home if their mean cognition score is higher than 4. Both stroke and orthopedic patients have higher odds of being discharged home if they were admitted from home than anywhere else. Now, with regard to the models for the FIM difference, patients that were readmitted tend to increase the motor dimension of FIM by over half a point. Even though the variable age is statistically significant for both motor and cognition FIM difference scores, the change in motor and cognition FIM difference scores due to age is small and thus, practically not meaningful. There is little change in the cognition FIM scores for orthopedic patients. The factors that statistically affect the cognition FIM difference scores, probably do not have strong practical significance (possibly with exception of ethnicity). The variables admission class and ethnicity affect cognition FIM difference scores for stroke patients. Those readmitted have a lower cognition FIM difference scores, while Hispanic patients have a higher difference score.

Table of Contents

Chapter 1

Introduction

1.1 Background & Motivation

There are numerous ways that a person can be injured or become ill, and after a serious injury, illness, or surgery, patients often recover slowly. Patients need to regain strength, relearn skills, or find new ways of performing tasks. This process is called rehabilitation. Rehabilitation often focuses on different aspects. For example, it may utilize physical therapy to help gain strength, mobility and fitness. It may include occupational therapy to help with daily activities, or it may include speech-language therapy to help with speaking, understanding, reading, writing and swallowing. The type of therapy and the goals of therapy may be differ. For instance, an older person who has had a stroke may simply want rehabilitation to be able to dress or bathe without help, while a younger person who has had a heart attack may go through cardiac rehabilitation to return to work and normal activities. Whatever the case, it is necessary to measure the disability of the patient and then the recovery that the patient gained during the rehabilitation process. This is where the Functional Independent Measure (FIM) becomes important.

1.1.1 FIM

The FIM is the most widely accepted functional assessment measure used in the rehabilitation community. It is used by approximately 60% of rehabilitation hospitals in the US as an assessment measure (Stineman and Fiedler, 1998). The FIM provides a standardized system of measurement for disability based on the International Classification of

Impairment, Disabilities and Handicaps; it measures the level of a patient's disability and indicates how much assistance is required for the individual to carry out activities of daily living. Thus the admission and discharge FIM of a patient can be used as an assessment of the rehabilitation process. The FIM is an 18-item ordinal scale, used for all diagnoses within a rehabilitation population. The item scores range from one to seven: a FIM item score of seven is categorized as "complete independence," while a score of one is "total assist" (performs less than 25% of task). Scores falling below six require another person for supervision or assistance. The FIM measures independent performance in self-care, sphincter control, transfers, locomotion, communication, and social cognition. The 18-items of FIM can also be divided into 13 motor tasks and 5 cognitive tasks. By adding the points for each item, the possible total score ranges from 18 (lowest) to 126 (highest) level of independence.The 18 items are: eating, grooming, bathing, upper body dressing, lower body dressing, toileting, bladder management, bowel management, bed to chair transfer, toilet transfer, shower transfer, locomotion (ambulatory or wheelchair level), stairs, cognitive comprehension, expression, social interaction, problem solving and memory.

1.2 Significance of the Study

Generally, it may be assumed that rehabilitation hospitals focus on providing continuous improvement of services to their patients. In order to do so, they need to keep track of the improvements of their patients. One way to achieve this is to use FIM difference scores. FIM difference is calculated by subtracting the admission FIM from the discharge FIM. Whenever the FIM difference is high, this indicates an improvement in the patient. Another concern for rehabilitation hospitals is whether the place of discharge may be accurately predicted using admission FIM scores. Accurately predicting place of discharge early in care can improve care and prepare the patient for the likely discharge location.

1.2.1 Rehabilitation Hospitals of the Southwestern United States

A Southwestern rehabilitation hospital consistently reported large FIM difference scores. In order to maintain this high ranking, they wish to investigate the reasons behind these large difference scores so that they can further improve their facilities and identify factors that influence the FIM. As a result we had a chance to visit this hospital and were given access to patient data. Using the online software portal called Uniform Data System (UDS) central, a software portal that is maintained by Uniform data system for medical rehabilitation (UDSMR), the data were obtained. One of the concerns raised by the staff was how the admission FIM scores, scores from an index measuring motor and cognitive functioning, affects place of discharge. The staff also wanted to know how the patient characteristics affect FIM difference scores.

1.2.2 Research Hypothesis

This research intended to cover its analysis in the following areas.

1. How does admission FIM affect place of discharge for this rehabilitation hospital located in the southwestern United States?
2. What patient characteristics affect FIM difference scores?

Many studies have been done in the past relating to place of discharge, admission FIM scores, FIM difference scores and patient characteristics. However, most of these studies fail to account for the multi-layered dimensionality of the FIM and the measurement error associated with the FIM items. These issues will be discussed explicitly in Chapter 2. This study attempts to address this issue by using Generalized Linear Latent Mixed Models (GLLAMM) and Structural Equation Models (SEM) for modeling FIM data. More details about these models will be discussed in Chapter 3.

To investigate the above mentioned hypotheses, the orthopedic and stroke sub-populations of rehabilitation patients are considered. The stroke and orthopedic populations were se-

lected for the study because they comprise a large proportion of the hospital population and the FIM is believed to operate differently for these distinct populations. Whatever insights gained after the analysis will be shared with the hospital staff in order to improve services to patients.

Chapter 2

Use of the Functional Independent Measure

In rehabilitation, place of discharge plays an important role. Having an assessment of where a patient is likely to be at discharge may provide additional time to their families members to prepare both emotionally and financially for the long-term consequences of the patient's disability. For example, if the patient is going to be discharged to his/her home, some changes might be needed in the bathroom or in his/her room. Knowing early that the patient is coming home will give the family members some time to prepare. Knowing early that the patient is going to be discharged to a nursing home might give some opportunity to prepare financially.

By design, the FIM instrument is not intended to incorporate all the activities that could possibly be measured, or that might need to be measured, for clinical purposes. Rather, the FIM instrument is a basic indicator of severity of disability that can be administered comparatively quickly and can be used to generate data on large groups of people. As the severity of disability changes during rehabilitation, the data generated by the FIM instrument can be used to track such changes and analyze the outcomes of rehabilitation.

This chapter focuses on the studies that have been done regarding place of discharge and FIM, and the study implications. Improvements that can be done regarding those studies will also be discussed.

2.1 Past Research

2.1.1 Why Place of Discharge ?

According to Kaplan (2011), hospitals tend to discharge patients needing post-acute care to too high level of care setting at least 1/3 of the time. This implies that, one out of three discharged individuals who receive care after hospitalization would get a similar clinical result in a lower level of care, or with no post-acute care. As a result of this, some new changes were made to the Patient Protection and Affordable Care Act in 2011 (Kaplan, 2011). Therefore most of the hospitals and other Medicare providers will have a strong interest in delivering care that is of the highest quality and lowest cost.

2.1.2 FIM as a Measure of Place of Discharge

In the article, "The right person in the right place" it is described how difficult it is to find the right discharge place for the rehabilitation patients (Seiger, 2011). Even though the FIM can be used to make decisions about the discharge place, it is difficult to find the right cut off and the cut off may change based on the many factors such as patient, hospital, type of disability, etc (Seiger, 2011).

In their study, Mauthe et al. (1996) analyzed FIM scores to determine whether the initial FIM scores could be used to predict place of discharge. The sample contained 298 patients whose attending physician referred the patient to the rehabilitation team. Apart from the patients whose acute stroke was confirmed by the treating physician, patients with a secondary diagnosis of stroke, such as one that occurs during surgery, were also included in the study. Patients with brain metastases and the patients who died during the study were excluded from the analysis. Therefore, the analysis was based on the remaining 279 patients. For those patients, an 18-item FIM was administered within 3 days of admission and 24 hours of discharge. A chi-square test was performed to determine whether there was a significant association between place of discharge and initial categorized FIM score.

The 18 items of the initial FIM were analyzed using a discriminant analysis to identify which items were most highly correlated to discharge disposition. Chi square test results were significant, indicating a significant association between place of discharge and initial FIM. Discriminant analysis identified 6 items as being statistically significant in predicting discharge to home, rehabilitation facility, or nursing home. Then the initial FIM scores of these 6 items (bathing, bowel, toileting, social interaction, dressing lower body, and eating) were used to predict the place of discharge. The correct classification was 81% for home, 60% for rehab and 71% for nursing home.

A similar study was conducted by Sandstrom et al. (1998) which describes discharge destination using FIM items. However, this study was limited to FIM items related to motor function. The sample was obtained from stroke patient data of a retrospective review of 293 cases from the years 1993 to 1995. Patients who were discharged home had higher admission and discharge motor FIM scores than those discharged to a sub acute facility or long-term care facility, although the association between motor FIM score and discharge destination was low to moderate.

Stineman et al. (1997) intended to find impairment-specific dimensions beyond motor and cognitive dimensions of the FIM. A sample of 93,829 patients discharged in 1992 from 252 free standing rehabilitation hospitals were obtained from the Uniform Data System for Medical Rehabilitation. Initially a conceptual model was developed to characterize expected constructs in an instrument based on physiological or socio-biological principles. Then, factor analysis was used to test the adequacy of this conceptual model. The conceptual model had four dimensions. Total FIM at level 1 followed by motor and cognitive FIM as level 2 (two factors). The level 3 dimension consisted of three factors: activities of daily life, mobility and cognitive FIM. In the level 4 dimension, activities of daily life were divided into sphincter control and self-care to create four factors (for more information refer to Figure 2.1). There were 20 categories of impairment and results differed depending on the patients' impairment. For example, for stroke patients, the three-factor structure explained 71.3% of the variation compared with 65.8% for the two-factor stroke structure.

According to the results, four impairments had a 3 dimensional factor structure, and 14 had a 4 dimensional structure. The impairment-specific dimensions were always nested within the motor-FIM subscale. The reliability coefficients for subscales based on these dimensions ranged from .74 to .97. As a conclusion Stineman et al. (1997) suggested that the FIM can be viewed as a multi-layered multi-dimensional measure of human functions and be classified to cluster by the area of body involved, neurological level, or relative energy consumption. They also suggested that the total FIM score is not appropriate as a measure if the intention is to differentiate physical disabilities from other types.

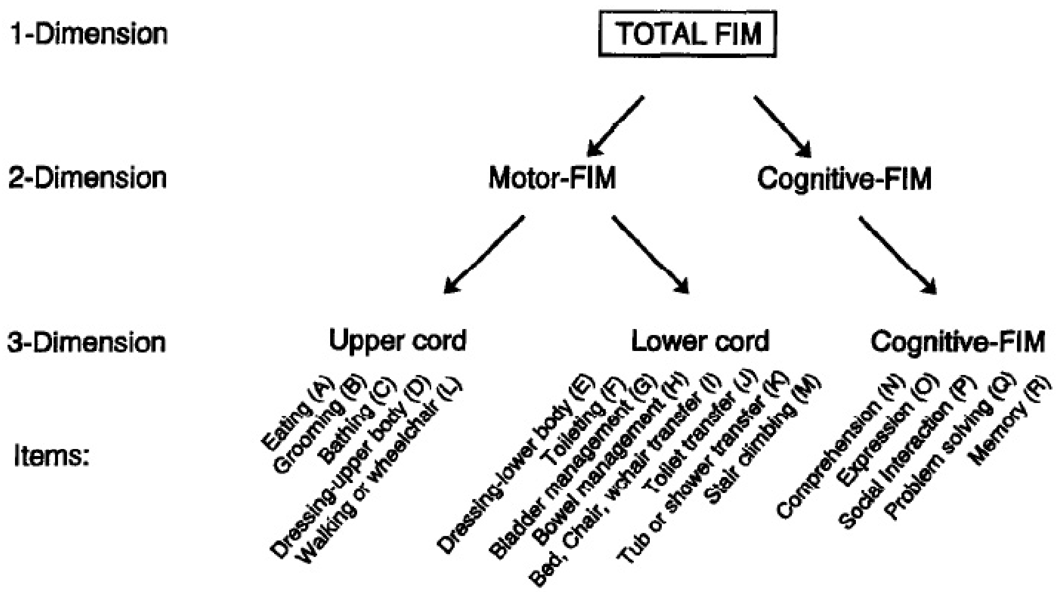

Figure 2.1: Dimensions within the FIM, (Stineman et al., 1997)

2.1.3 Improvements

In the study by Mauthe et al. (1996), discriminant analysis was used to identify the FIM items relating to place of discharge. Given that the assumptions are met, discriminant analysis is more powerful and efficient when compared to logistic or probit models that can also be used to identify the FIM items relating to discharge disposition. However, discriminant analysis requires the independent variables to be normally distributed, linearly related, and

have equal variance within each group (Tabachnick and Fidell, 1996). When dealing with rehabilitation data, these assumptions are rarely valid; thus using discriminant analysis on a sample that violates these assumptions might lead to incorrect insights. Mauthe et al. (1996) also failed to account for the measurement error that associated with FIM scores. Thus, the result of this study can be further improved using a model that can incorporate the measurement error while relaxing the assumptions of normality or equal variance.

Similarly, Sandstrom et al. (1998) made most of their conclusions using descriptive statistics for the FIM, like median. This study can be further improved using proper statistical models for inferential statistics. Even though Stineman et al. (1997) suggested FIM items have dimensions beyond motor and cognition, the results of their study indicate that not all impairment levels can be explained using higher dimensions. For example the impairment levels for, stroke, non-traumatic brain, traumatic brain or non-traumatic spinal cord patients can only be explained up to dimension 2 (motor and cognition) and the results related to dimensions 3 and 4 did not result in large improvements over dimension 2. The most important fact displayed in the results about dimension 2 (motor and cognition) is that it can be used to explain any impairment level with a minimum of 70% internal consistency.

Most of the research that has been done regarding the place of discharge focuses on either admission FIM or discharge FIM. Sometimes researchers use the total admission and discharge FIM difference (Amundson et al., 2000). Also some researchers included total FIM as a covariate, which will affect the accuracy of the model since FIM items can have different correlation structures. Wieslaw and Ban-ecu (1997) used a neural network model to predict the place of discharge using FIM. Yet Wieslaw and Ban-ecu (1997) reported a R^2 of 0.337 and 0.247 for their models and stated that the models have difficulties in accurately predicting the place of discharge.

In the following chapter, the method of analysis for examining the relationship between FIM and place of discharge are described. These models will take into account the multi-

dimensional nature of the FIM, account for the measurement error inherent in FIM scores, and will relax normality assumptions for the FIM scores.

Chapter 3

Methodology and the Techniques Applied

3.1 Generalized Linear Models (GLM)

"Generalized linear models (GLMs) represent a class of fixed effects regression models for several types of dependent variables (i.e., continuous, dichotomous, counts)" (Everitt and Howell, 2005). In other words, GLMs are a generalization of ordinary linear regression that are flexible enough to include a wide range of common situations, but at the same time allow most of the familiar ideas of normal linear regression to carry over. Common GLMs include linear regression, logistic regression, and Poisson regression. Any of these GLMs are identifiable by random component, systematic component and link function.

Random Component

The random component of a GLM identifies the response variable $\mathbf{Y}$ and assumes a probability distribution for the response. Denote the response $\mathbf{Y}$ by $(Y_1, ..., Y_n)$. Standard GLMs treat $Y_1, ..., Y_n$ as independent. Depending on the application, the observations on $\mathbf{Y}$ can be continuous, binary, or count. If it is binary then the preferred probability distribution for $\mathbf{Y}$ will be binomial. If it is a count, then either Poisson or negative binomial distributions will be used depending on the context. If each observation is continuous, the normal distribution will be assumed for $\mathbf{Y}$.

Systematic Component

The systematic component of a GLM specifies the explanatory variables. These enter linearly as predictors on the right-hand side of the model equation. This linear combination of the explanatory variables is called the linear predictor. Denote the linear predictor as η_i. Then,

$$\eta_i = \boldsymbol{x_i'\beta} \tag{3.1}$$

where $\boldsymbol{x}_i$ is the vector of regressors for unit i with fixed effects $\boldsymbol{\beta}$.

Link Function

Denote the expected value of Y_i , the mean of its probability distribution, by $\mu_i = E(Y_i|x_i)$. The third component of a GLM, the link function, specifies a function g(.) that relates μ_i to the linear predictor as

$$g(\mu_i) = \eta_i = \boldsymbol{x_i'\beta}. \tag{3.2}$$

The simplest link function is $g(\mu) = \mu$. This is called as the identity link. It specifies a linear model for the mean response,

$$\mu_i = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k \tag{3.3}$$

This is the form of regression models for a continuous response. If the response is binary, then common link functions are the logit or probit. Suppose $\mathbf{Y}$ is binary and the link function is logit. Then the corresponding GLM is,

$$\text{logit}[P(Y_i = 1)] = \log\left(\frac{P(Y_i = 1|x_i)}{1 - P(Y_i = 1|x_i)}\right) = \boldsymbol{x_i'\beta} \tag{3.4}$$

The above logistic regression formula implies the following formula for the response probability,

$$P(Y_i = 1) = \frac{\exp(\boldsymbol{x_i'\beta})}{1 - \exp(\boldsymbol{x_i'\beta})}. \tag{3.5}$$

In this study, place of discharge is the response. By coding the successful binary response as 'Home' (Y=1) and 0 otherwise, probability of the place of discharge can become a binary response. Then the above logistic regression can be used to model the place of discharge. Suppose sex, age, length of stay and ethnicity are the covariates of interest. Then the model will be,

$$\text{logit}[P(Y_i = 1|x_i)] = \beta_1 Age + \beta_2 Sex + \beta_3 Length\,of stay + \beta_4 Ethnicity \tag{3.6}$$

The most popular method for parameter estimation for GLM is Maximum likelihood estimation (MLE) method. Apart from MLE, a Bayesian approach or weighted least squares approach can also be used to estimate parameters.

3.2 Generalized Linear Mixed Models

One of the main assumptions of fixed effects models is that all observations are independent of each other. However, this is not appropriate when analyzing correlated data structures, in particular clustered and longitudinal data. In clustered data, the observed units are often nested within larger units. For example, students are nested within schools and patients are nested within hospitals. In longitudinal designs, repeated measures are nested within subjects. Such data can also be referred to as multilevel or hierarchical data, random cluster and/or subject effects can be added in to the regression model to account for the correlation structure. The resulting model is referred to as a mixed model including the usual fixed effects for the regressors plus the random effects. This combined model with fixed effects and random effects within the GLM framework is called a Generalized Linear Mixed Model (GLMM).

The linear predictor of a GLMM is,

$$\eta_i = \boldsymbol{x_i'\beta} + \boldsymbol{z}_i'\boldsymbol{v}_i \ , \tag{3.7}$$

with fixed effect covariate x_i, random effect covariates $\boldsymbol{z}_i$, fixed effect parameters β, and random effect parameters $\boldsymbol{v}_i$ for each subject i. Inclusion of the random effects allows for complex correlation structures, such as those resulting from clustered data.

3.3 Latent Variable

A latent variable is a variable that cannot be observed directly. For example, quality of life, business confidence, or intelligence are examples of unobservable or latent variables. These variables can only be measured indirectly by means of other variables that are directly observed. Various types of mathematical models have been developed in order to explain the observed variables in terms of latent variables and these latent variable models are used in many areas, including economics, biostatistics, bioinformatics, natural language processing, psychology, artificial intelligence, and the social sciences. One of the main advantages of using latent variables is that it reduces the dimensionality of data.

3.4 Covariate Measurement Error Models

Latent variable modeling is often utilized when observed variables exhibit significant covariate measurement error (CME). Covariates are often measured with error, their true value being unobservable or latent. For instance, FIM items can measure how well a patient completes some tasks but it cannot measure the true cognitive or motor ability of the patient. So the true ability can be treated as a latent covariate that is measured with error using FIM items. The covariate measurement error model consists of three submodels, outcome model, measurement model and true covariate model (Rabe-Hesketh et al., 2003). Suppose v_i is the latent covariate and other covariates $\boldsymbol{x_i}$ are perfectly measured. Then the outcome model can be written as,

$$g(\mu_i) = \boldsymbol{x_i'\beta} + v_i\beta_v \tag{3.8}$$

where $\boldsymbol{\beta}$ is the fixed effect parameter vector, β_v is the parameter for the latent variable and $\mu_i = E[Y_i|\boldsymbol{x}_i, v_i]$. Consider the situation where the true covariate v_i has been measured by n_i items, giving fallible measures $\boldsymbol{w}_{ij}, j = 1, ..., n_i$. Then the measurement error model can be stated as follows,

$$\boldsymbol{w_{ij}} = v_i + \boldsymbol{\epsilon_{ij}} \ , \qquad \boldsymbol{\epsilon_{ij}} \sim N(0, \sigma^2). \tag{3.9}$$

In the measurement error model, $\boldsymbol{\epsilon}_{ij}$ is the measurement error distributed normally with mean zero and constant measurement error variance σ^2. Thus, measurements $\boldsymbol{w}_{ij}(j = 1, ..., n_i)$ are conditionally independent given the true covariate v_i, and item responses for a unit having the mean equal to the true covariate : $E[w_{ij}|v_i] = v_i$.

The true covariate model is used to model the relationship between the true covariate and the other observed covariates in the outcome model (Rabe-Hesketh et al., 2003). The true covariate model is given by,

$$v_i = \boldsymbol{x_i'\lambda} + \boldsymbol{\zeta_i} \ , \qquad \boldsymbol{\zeta_i} \sim N(0, \tau^2) \tag{3.10}$$

where $\boldsymbol{\lambda}$ are regression parameters and $\boldsymbol{\zeta}_i$ are residuals, assumed to be independent of the fixed covariate $\boldsymbol{x_i}$. The total effect can be obtained by simply combining (3.10) into (3.8),

$$g(\mu_i) = \boldsymbol{x}_i'(\boldsymbol{\beta} + \boldsymbol{\lambda}\beta_v) + \boldsymbol{\zeta}_i\beta_v \tag{3.11}$$

The reduced form of the measurement model can be obtained by substituting the true covariate model (3.10) into the measurement model (3.9),

$$\boldsymbol{w}_{ij} = \boldsymbol{x_i'\lambda} + \boldsymbol{\zeta}_i + \boldsymbol{\epsilon}_{ij} \tag{3.12}$$

.

The combined model, comprising the three sub models (3.8), (3.9), (3.10) is illustrated in Figure 3.1 for the case of $j = 2$ measurements. Circles represent latent variables and rectangles represent observed variables. Long arrows represent linear relations involving the linear predictor, and short arrows represent residual variability.

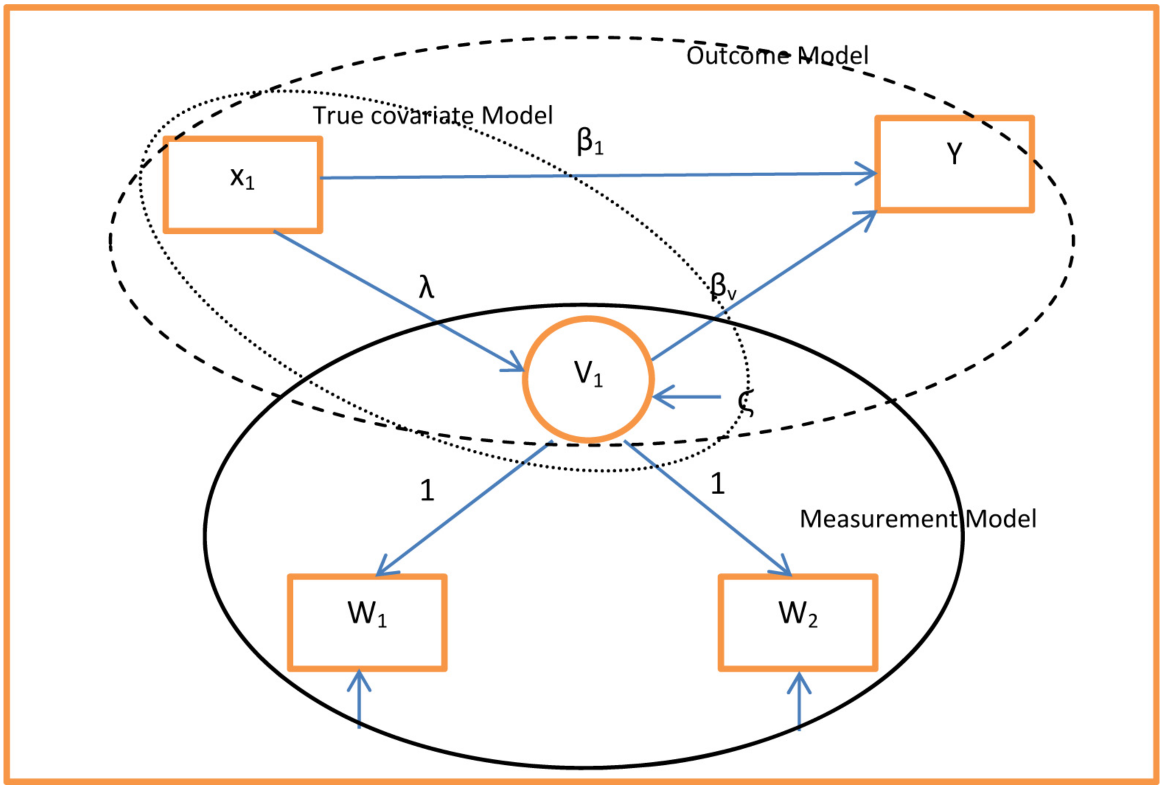

Figure 3.1: CME with one latent variable

In section 3.5 and 3.6, two major types of CME models are discussed that have application to FIM modeling.

3.5 Structural Equation Modeling (SEM)

Structural equation modeling is a class of statistical models that allow complex relations between one or more independent variables and one or more dependent variables. There are many ways to describe SEM, among them the most common is, as a hybrid between

some form of analysis of variance (ANOVA)/regression and some form of factor analysis. SEM encompasses a broad array of models from linear regression to measurement models to simultaneous equations, including confirmatory factor analysis (CFA) models, correlated uniqueness models, latent growth models, and multiple indicators and multiple causes (MIMIC) models.

In SEM, models are often illustrated in a path diagram (See Figure 3.2).

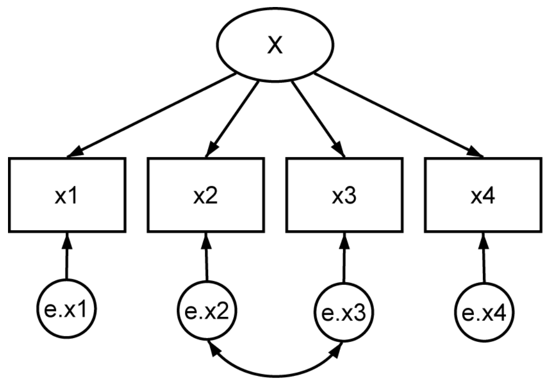

Figure 3.2: SEM Example

This diagram is composed of the following: Boxes represent variables that are directly observed. Circles contain variables that are unobserved, known as latent variables. Arrows, called paths, that connect some of the boxes and circles. When a path points from one variable to another, it means that the first variable affects the second. Double headed arrows indicate association between two variables.

3.5.1 Limitations, Assumptions and Parameter Estimates Regarding SEM

Since SEM is a confirmatory method, the full model should be specified before the analysis and the number of parameters, relationships, and covariance should also be specified. SEM also requires large samples to model complex relationships between multivariate data. Like other multivariate statistical methodologies, most of the estimation techniques used in

SEM require multivariate normality and the data need to be examined for univariate and multivariate outliers. Suppose our model is the one in Figure 3.3.

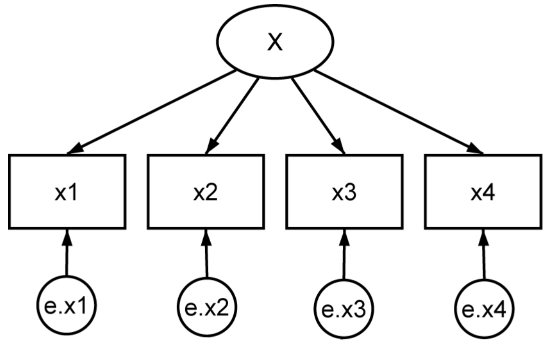

Figure 3.3: SEM Example

Then,

$$x_1 = \alpha_1 + X\beta_1 + e_{x_1} \tag{3.13}$$

$$x_2 = \alpha_2 + X\beta_2 + e_{x_2} \tag{3.14}$$

$$x_3 = \alpha_3 + X\beta_3 + e_{x_3} \tag{3.15}$$

$$x_4 = \alpha_4 + X\beta_4 + e_{x_4} \tag{3.16}$$

where each covariate and error term are independently normally distributed with mean μ and variance Σ. The model will estimate the parameter set $\theta = (\alpha, \beta, \mu, \Sigma)$. The default estimation method used for SEMs is the maximum likelihood method. MLE's are consistent and asymptotically normal assuming joint normality of all the variables, but usually one can derive most of the desired properties from conditional normality (StataCorp, 2011).

QML, quasi maximum likelihood, uses ML to fit the model but relaxes the normality assumptions when estimating the standard errors. Thus, concerning the parameter estimates, everything just said about ML applies to QML. Concerning standard errors, theoretically, one can expect consistent standard errors with QML. However, there are some estimation methods that do not require normality. For example, asymptotic distribution free (ADF)

method makes no assumption of joint normality or even symmetry (StataCorp, 2011). ADF assumes the observed variables and errors are iid with mean μ and variance Σ.

ADF utilizes weighted least squares (WLS) and is a generalized method of moments (GMM) estimator. The fitting function of the ADF can be written,

$$F_{ADF} = (\boldsymbol{s} - \boldsymbol{\sigma})' \boldsymbol{W}^{-1} (\boldsymbol{s} - \boldsymbol{\sigma}) \tag{3.17}$$

where $\mathbf{s}$ is the $p(p+1)/2$ vector of distinct elements of $\mathbf{S}$, sample covariance matrix. $\boldsymbol{\sigma}$ contains the corresponding elements of $\mathbf{\Sigma}$ and $\mathbf{W}$ is a consistent estimator of the asymptotic covariance matrix of $\mathbf{s}$. For ADF, $\mathbf{W}$ is given by following equation.

$$(N-1)cov(s_{ij}, s_{kl}) = \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk} + \frac{N-1}{N}\kappa_{ijkl} \tag{3.18}$$

where κ_{ijkl} is a fourth order cumulant. (Muthen and Kaplan, 1985).

In simulations of the measurement model, ADF produces excellent results, even for the standard error of the variance of $\mathbf{X}$ (StataCorp, 2011). However, ADF is less efficient than ML when ML's assumptions hold, whatever those minimal assumptions are. On the other hand, under the less restrictive ADF assumptions, the ADF estimator is more efficient than QML, although the QML estimator will still be consistent and have correct coverage.

SEM techniques only look at first-order (linear) relationships between variables. Linear relationships can be explored by creating bivariate scatter plots for all of your variables. Power transformations can be made if a relationship between two variables seems quadratic. Multicollinearity among the independent variables for directly observed variables can be an issue too.

3.6 Generalized Linear Latent Mixed Models (GLLAMM)

The GLLAMM generalizes the GLMM further by allowing latent variable as a response and/or explanatory variable. Suppose there are L levels, and M_l latent variables at level l, then the linear predictor has the form

$$\eta_i = \boldsymbol{x_i'}\boldsymbol{\beta} + \sum_{l=2}^{L}\sum_{m=1}^{M_l} \boldsymbol{v}_m^{(l)}\boldsymbol{\lambda}_m^{(l)}\boldsymbol{z}_m^{(l)} \tag{3.19}$$

with the 1^{st} element of $\boldsymbol{\lambda}_m^{(l)}$, set to 1, i.e. $\lambda_{m1}^{(l)} = 1$.
The elements of $\boldsymbol{x_i}$ are explanatory variables associated with the fixed effects $\boldsymbol{\beta}$, $\boldsymbol{v}_m^{(l)}$ is the m^{th} latent variable at level l, and each latent variable is multiplied by a linear combination of explanatory variables given by $\boldsymbol{\lambda_m^{(l)\prime} z_m^{(l)}}$. Here the superscript of $\boldsymbol{z}_m^{(l)}$ denotes how the corresponding latent variable varies at level l. The latent variables at the same level are generally mutually correlated whereas latent variables at different levels are independent (Rabe-Hesketh et al., 2004).
For example, in this study, there are two levels, the item level (level 1) and the patient level (level 2). Patient level has two latent variables, true motor ability of the patient (v_{mo}) and the true cognition ability (v_{co}) of the patient. Thus, the equation (3.20) can be reduced into,

$$\eta_i = \boldsymbol{x_i'}\boldsymbol{\beta} + v_{mo}^{(2)}\boldsymbol{z}_{mo}^{(2)} + v_{co}^{(2)}\boldsymbol{z}_{co}^{(2)} \tag{3.20}$$

where, $\boldsymbol{z}_{mo}^{(2)}$ represent the variables related to the latent covariate $v_{mo}^{(2)}$ and $\boldsymbol{z}_{co}^{(2)}$ represent the variables related to the latent covariate $v_{co}^{(2)}$. The variable $\lambda^{(2)}$ act as an indicator that allocates variables to each latent variable at level 2.

3.6.1 GLLAMM approach to CME

This section illustrate how the CME model mentioned above using equation (3.8), (3.9) and (3.10) is connected to the GLLAMM model (3.20)

By reorganizing the GLLAMM ,

$$g_{ij}(\mu_{ij}) = \boldsymbol{t'_{ij}}\boldsymbol{\beta} + v_i\boldsymbol{u'_{ij}}\boldsymbol{\lambda} \ , \qquad \lambda_1 = 1 \tag{3.21}$$

where $\boldsymbol{t}'_{ij}$ and $\boldsymbol{u}_{ij}$ are variables, $\boldsymbol{\beta}$ and $\boldsymbol{\lambda}$ parameters and v_i is a latent variable. The ij subscript for the link implies that different link functions can apply to different responses. For simplicity, the model is restricted to only two levels and a single latent variable (the general framework is described in Rabe-Hesketh, Skrondal, and Pickles (2004) and Skrondal and Rabe-Hesketh (2004)). All response variables, the outcome Y_i , and the item responses $\boldsymbol{w}_{ij}$ of the true covariate are stacked in a single response vector $\mathbf{Y}$ with elements indexed ij, where $j = 1$ if the element corresponds to the outcome for unit i and $j = 2, ..., k+1$ for the fallible measurements. For this study, $j = 1$ corresponds to the value of the response variable 'place of discharge' for patient i and $j = 2, ..., 19$ correspond to values of the 18 FIM items for patient i. A dummy variable is used to include the observed covariates $\mathbf{x}$ only in the outcome model,

$$d_{1ij} = \begin{cases} 1 & if\ j = 1 \\ 0 & if\ j > 1 \end{cases} \tag{3.22}$$

The variables $\boldsymbol{u}_{ij}$ are typically dummy variables, allowing the regressions for different response variables on the latent variable v_i to have different regression coefficients $\boldsymbol{\lambda}$. In the present example, using the dummy variables $u_{1ij} = d_{mij}$,

$$d_{mij} = 1 - d_{1ij} = \begin{cases} 0 & if\ j = 1 \\ 1 & if\ j > 1 \end{cases} \tag{3.23}$$

and $u_{2ij} = d_{1ij}$ results in the following model:

$$g_{ij}(\mu_{ij}) = d_{ij}\boldsymbol{x'}\boldsymbol{\beta} + v_i(\lambda_1 d_{mij} + \lambda_2 d_{1ij}) \ , \qquad \lambda_1 = 1 \tag{3.24}$$

where $d_{ij}\boldsymbol{x'} = \boldsymbol{t}'_{ij}$. The equation (3.24) can be further reduced using the index j,

$$g_{ij}(\mu_{ij}) = \begin{cases} \boldsymbol{x'\beta} + v_i\lambda_2 & if\ j = 1 \\ v_i & if\ j > 1 \end{cases} \tag{3.25}$$

The Figure 3.4 illustrates this scenario for the response, place of discharge, with two latent covariates.

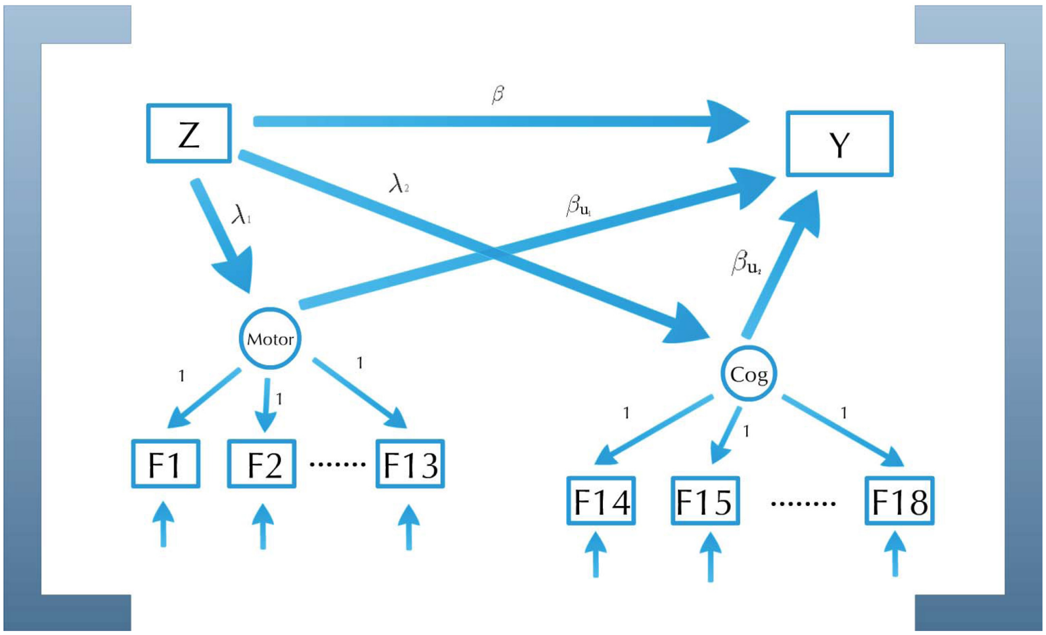

Figure 3.4: CME with two latent variables

3.6.2 Parameter Estimates

3.6.2.1 Maximum Likelihood Estimates (MLE)

The maximum likelihood estimate of a parameter is the value for which the probability of the observed data takes its greatest value. Suppose there is a sample $x_1, x_2, ..., x_n$ of n independent and identically distributed observations with pdf $f(.|\theta)$ and θ is unknown. Then the likelihood is given by,

$$L(\theta|x_1, x_2, ..., x_n) = f(x_1, x_2, ..., x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta) \tag{3.26}$$

The estimate for the parameter θ can be obtained by MLE,

$$\hat{\theta}_{MLE} \subset \{\text{arg max}_{\theta \in \Theta}\ \hat{l}(\theta | x_1, x_2, ..., x_n)\} \tag{3.27}$$

where, $\hat{l} = \frac{1}{n}$lnL. An MLE is the same regardless of whether the maximization done using the likelihood or the log-likelihood function, since log is a monotone transformation. Estimators based on the method of maximum likelihood are popular because they have good large-sample behavior. Most importantly, it is not possible to find good estimators that are more precise, in terms of having smaller large-sample standard errors. Also, large-sample distributions of ML estimators are usually approximately normal (Agresti, 2007).

3.6.2.2 Newton Raphson Algorithm

The Newton Raphson algorithm is an iterative procedure that can be used to calculate MLE's (Agresti, 2007). It utilizes the 2^{nd} order Taylor series expansion. Suppose $f(\mathbf{x})$ is the function that is to be maximize. First, expand $f(\mathbf{x})$ around an arbitrary point $\mathbf{a}$ using Taylor series expansion,

$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + (\boldsymbol{x} - \boldsymbol{a})^T f^{'}(\boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})^T f^{''}(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a}) \tag{3.28}$$

where $f^{'}(.)$ is the gradient vector and $f^{''}(.)$ is the Hessian matrix. Since 3.28 is a quadratic equation maximization can be achieved by differentiating and equating to zero.

$$\frac{d}{d\boldsymbol{x}}\left[f(\boldsymbol{a}) + (\boldsymbol{x} - \boldsymbol{a})^T f^{'}(\boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})^T f^{''}(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a})\right] = f^{'}(\boldsymbol{a}) + (\boldsymbol{x} - \boldsymbol{a})^T f^{''}(\boldsymbol{a}) = 0 \tag{3.29}$$

Rearranging the terms,

$$\boldsymbol{x} = \boldsymbol{a} - [f^{'}(\boldsymbol{a})]^T [f^{''}(\boldsymbol{a})]^{-1} \tag{3.30}$$

Then the Newton Raphson process can be used to update the current estimate until convergence. Let denote x_0 the initial point. It is assumed to be arbitrary. Then x_1 will

obtained by plugging the value of x_0 into equation 3.31. This iteration continues until the convergence achieved.

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i - [f^{'}(\boldsymbol{x}_i)]^T[f^{''}(\boldsymbol{x}_i)]^{-1} \tag{3.31}$$

If $f(\mathbf{x})$ is a concave function, then the Newton Raphson algorithm is assured to converge to the correct answer. If $f(\mathbf{x})$ is a convex function, for some values of $\mathbf{x}$, then the algorithm may or may not converge (Viele, 2008).

3.6.2.3 Gaussian Quadrature

The quadrature rule is used to approximate the definite integral of a function by using weighted sum of function values at specified points within the domain of integration. Gaussian quadrature is a quadrature rule that gives an exact result for polynomials of degree 2n-1 or less by a suitable choice of the points x_i and weights w_i for $i = 1, ..., n$. Gaussian quadrature is optimal because it fits all polynomials up to degree 2n-1 exactly (Rabe-Hesketh et al., 2002). To use the Gaussian quadrature for integral approximation, first it must be changed into an integral over [-1, 1]. Suppose $f(x)$ is an integral over [a, b]. Then it can be converted in to [-1, 1] using the following change of variables,

$$\int_a^b f(x)dx = \frac{b-a}{2}\int_{-1}^{1} f\left(\frac{b-a}{2}x + \frac{a+b}{2}\right) dx \tag{3.32}$$

Then,

$$\int_a^b f(x)dx \approx \frac{b-a}{2}\sum_{i=1}^{n} w_i f\left(\frac{b-a}{2}x_i + \frac{a+b}{2}\right) \tag{3.33}$$

where w_i's are the weights and x_i's are the chosen points.

3.6.2.4 Adaptive Quadrature

Most of the quadrature rules continually subdivide the integration domain until sufficient accuracy is achieved. However, the drawback of this approach is that some portions of

the curve may have large functional variations that require more attention than other portions of the curve. Adaptive quadrature, is a quadrature rule that addresses this problem (Rabe-Hesketh et al., 2002). It is also subdivide the integration domain into sub-intervals. Suppose f(x) is the function integrated over the domain [a, b]. Then [a, b] is divided into n subintervals $[a_j, b_j]$, for $j = 0, 1, ..., n-1$, and the trapezoidal rule is used on each subinterval to compute:

$$I_j(f) = \int_{a_j}^{b_j} f(x)dx \tag{3.34}$$

by any quadrature rule. The difference introduced when using adaptive quadrature is that the $I_j(f)$ has not computed with sufficient accuracy then the subinterval $[a_j, b_j]$ is subdivided. This also avoids unnecessary computation and obtains the same accuracy as with other quadrature procedures, yet with much less computational effort. To check the accuracy, the difference of I_j and I_{j+1} will be used. If the difference is greater than some threshold value the method will proceed to subdivide, otherwise, return the value of I_{j+1} as the answer. There are different methods for choosing the threshold and it is beyond the scope of this study.

3.6.2.5 Parameter Estimation of GLLAMM

The Stata GLLAMM package uses the Maximum Likelihood method and modified Newton Raphson algorithms to calculate the parameter estimates. For a two level GLLAMM model likelihood is given by

$$\prod_j \int \prod_i (f(y_{ij}|x_{ij}, \eta_j)g(\eta_j))d\eta_j \tag{3.35}$$

where $f(y_{ij}|x_{ij}, \eta_j)$ is the conditional density of the response variable given the latent and explanatory variables and $g(\eta_j)$ is the prior density of the latent variables. When the latent variables are discrete, the integral becomes a sum of the form

$$\prod_j \sum_r \pi_r \prod_i f(y_{ij}|x_{ij}, \eta_j = e_r) \tag{3.36}$$

where the locations e_r and masses π_r are freely estimated. For a single normally distributed latent variable, the same expression is used to approximate the likelihood, where locations and masses are given by Gaussian quadrature (Rabe-Hesketh et al., 2004). Ordinary Gaussian quadrature sometimes performs poorly because there are insufficient locations under the peak of the integrand in (3.35). According to Rabe-Hesketh et al. (2002) adaptive quadrature perform better in these situations.

Chapter 4

Data Analysis

4.1 Obtaining the Data

The data set used for the analysis was obtained from the UDSMR software portal using SQL queries. Discharge date was used to obtain the sample for the time period 01/01/2008 to 01/01/2010. Due to hospital policies, some information like treating physician was not available to us. However the retrospective data set contained 107 variables and 1715 observations. Only patients classified as stroke or orthopedic patients are utilized in the analysis.

4.2 Data Coding and Descriptive Statistics

Since primary interest is to investigate a relationship between place of discharge and the FIM scores and also to find the variables affecting the admission and the discharge FIM difference, some of the variables were excluded from the study. New variables were also created using the original variables. For example, there was no variable called AGE, thus the date of birth and the admission year variables were used to recreate the variable AGE. The variable length of stay was created using admission and discharge date variables.

Primary Pay Source

Variable PRIMARY PAY SOURCE contains the information about patient's primary payment method for inpatient rehabilitation services. There were a total of 18 payment methods coded from 1-16 and 51, 52.

According to the Table 4.1, code 2 (Medicare non-MCO) contains 77.64% of orthopedic and 72.73% of stroke patients of the data, so most of the patients used ‘Medicare non-MCO’ as their primary payment method. Also, there were no stroke patients belonging to the primary payment category ‘workers compensation’.

Table 4.1: Frequency table for Primary Pay Source by Impairment category

Orthopedic				Stroke			
Primary Pay Source	Freq.	Percent	Cum.	Primary Pay Source	Freq.	Percent	Cum.
1	48	5.96	5.96	1	22	7.69	7.69
2	625	77.64	83.6	2	208	72.73	80.42
3	10	1.24	84.84	3	4	1.4	81.82
4	19	2.36	87.2	4	11	3.85	85.66
5	58	7.2	94.41	5	26	9.09	94.76
6	4	0.5	94.91	10	1	0.35	95.1
10	5	0.62	95.53	14	1	0.35	95.45
14	14	1.74	97.27	15	2	0.7	96.15
15	1	0.12	97.39	51	5	1.75	97.9
51	9	1.12	98.51	52	6	2.1	100
52	12	1.49	100	Total	286	100	
Total	805	100					

1-Blue Cross, 2-Medicare non-MCO, 3-Medicaid non-MCO, 4-Commercial Insurance, 5-MCO HMO, 6-Workers Compensation, 10-Private Pay, 14-Other, 15-None, 16-No-Fault Auto Insurance, 51-Medicare MCO, 52-Medicaid MCO

Most of the categories other than category Blue Cross, Medicare non-MCO and MCO HMO

contain fewer observations. Thus a decision was taken to recode the variable PRIMARY PAY SOURCE as,

$$PrimaryPaySource = \begin{cases} 1 & if\ PrimaryPaySource = 01, 04, 05, 10, 16 \\ 0 & otherwise \end{cases} \tag{4.1}$$

AGE

The variable AGE contains the age in years of the patient up to the admission date. Table 4.2 shows that 75% of the patients in either impairment category are above the age of 67. It also shows that only two stroke patients are less than 30 years old and the rest of the patients are above the age of 46. The mean and the standard deviation of age for both impairment categories are similar.

Table 4.2: Descriptive analysis for AGE by Impairment category

Orthopedic					**Stroke**		
	Percentiles	Smallest			Percentiles	Smallest	
1%	39	25			46	28	
5%	54	25			54	29	
10%	60	26	Obs	805	60	46	286
25%	68	28	Sum of Wgt.	805	67	46	286
50%	76		Mean	74.8720	77		74.7132
		Largest	Std. Dev.	11.6388		Largest	11.4475
75%	83	98			83	95	
90%	88	100	Variance	135.4625	88	97	131.0473
95%	92	101	Skewness	-0.8180	90	99	-0.7732
99%	95	103	Kurtosis	4.3428	97	100	4.1014

Figures 4.1 and 4.2 show histograms for the variable age for orthopedic and stroke sub

populations. Age values of stroke patients are less skewed than orthopedic patients. Most of the stroke patients are older than 45. Age values of the orthopedic patients clearly skewed to the left.

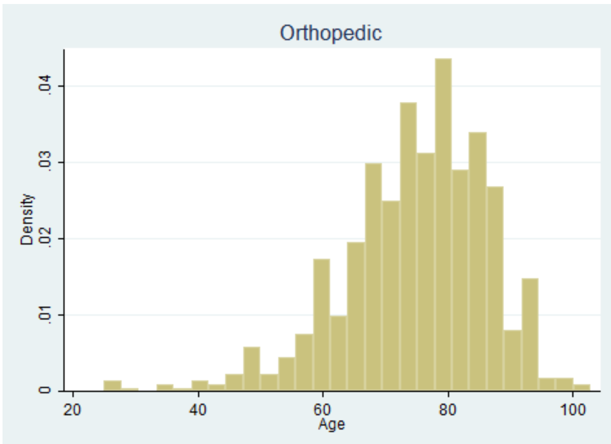

Figure 4.1: Histogram-AGE(orthopedic)

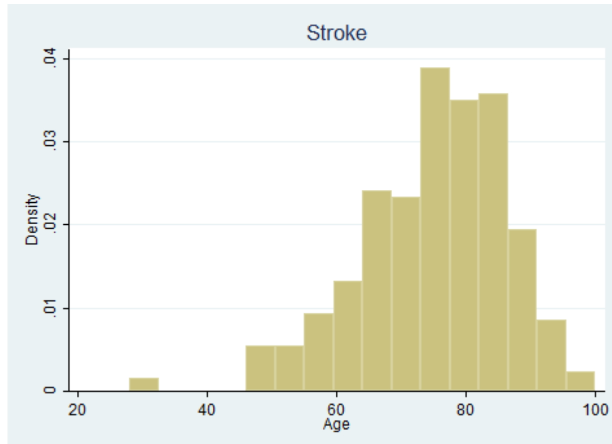

Figure 4.2: Histogram for AGE (stroke)

Length of Stay

The variable length of stay contains the duration (number of days) that the patient spent on the hospital. According to the Table 4.3, the mean number of days an orthopedic patient spent in the hospital is 12 and the mean number of days a stroke patient spent on the hospital is 18. Hence, on average, stroke patients have a longer length of stay. However the standard deviation of orthopedic patients is 5 and standard deviation of the stroke patient is 9. The largest value for the orthopedic patient is 36 and for a stroke patient is 77.

Table 4.3: Descriptive analysis for length of stay by Impairment category

Orthopedic					**Stroke**		
	Percentiles	Smallest			Percentiles	Smallest	
1%	1	0			1	0	
5%	5	0			5	1	
10%	6	1	Obs	805	8	1	286
25%	9	1	Sum of Wgt.	805	12	1	286
50%	12		Mean	12.5354	18		17.9615
		Largest	Std. Dev.	5.0002		Largest	8.8869
75%	16	27			23	42	
90%	19	29	Variance	25.0027	29	43	78.9774
95%	21	31	Skewness	0.4568	31	47	1.2302
99%	25	36	Kurtosis	3.5793	43	77	9.1445

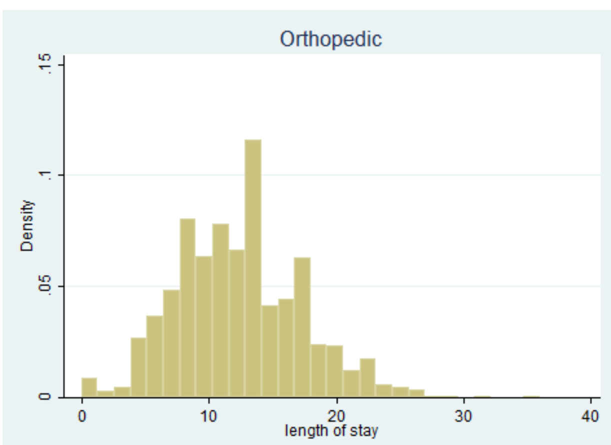

Figure 4.3: Histogram length of stay

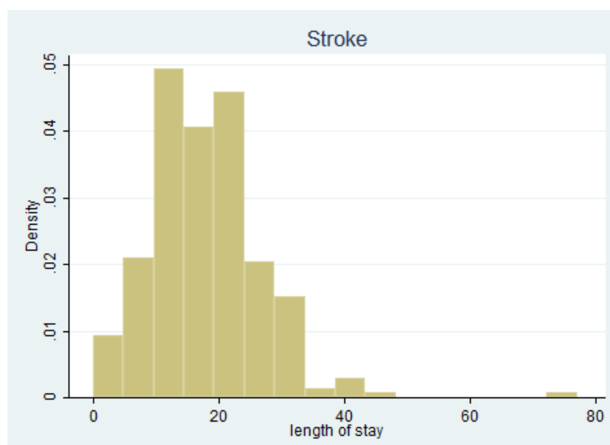

Figure 4.4: Histogram length of stay

According to Figure 4.3, length of stay values for orthopedic patients have a nearly symmetric distribution. The histogram for the stroke patients (Figure 4.4) is slightly skewed to the right. There are few observations to the far right around the value 80 causing this

skewness in the length of stay data of stroke patients.

Gender

Table 4.4 shows that 2/3 of the orthopedic patients are females, but for stroke patients, the ratio between male and female is near 0.5.

Table 4.4: Frequency table for Gender

Orthopedic				Stroke				
Gender	Freq.	Percent	Cum.	Gender	Freq.		Percent	Cum.
1-Male	254	31.55	31.55	1		150	52.45	52.45
2-Female	551	68.45	100	2		136	47.55	100
Total	805	100		Total		286	100	

Ethnicity

The hospital used six binary variables to record the variable ethnicity. Each of these binary variables represent an ethnicity type. For the population of orthopedic patients, 66.83% were white, 17.02% were black, 8.32% were Hispanic, 3.48% were islanders and 0.5% were Asians. For the population of stroke patients, 56.99% were white, 22.03% were black, 11.19% were Hispanic, 4.9% were islanders 0.35% were native americans and 0.35% were Asians

The majority of the patients were white. However, there are considerable number of black and Hispanic patients, too. Being a hospital that close to the Mexican border, the staff of the rehabilitation facility was concerned about any differences in Hispanic patients. Therefore, the variable Ethnicity was recoded as Hispanic and not Hispanic.

Marital Status

According to Table 4.5, approximately half of the orthopedic patients are married. For the population of stroke patients, more than half of the patients were married.

Table 4.5: Frequency Table for Marital Status

Orthopedic				Stroke			
Marital Status	Freq.	Percent	Cum.	Marital Status	Freq.	Percent	Cum.
1	112	14.18	14.18	1	38	13.67	13.67
2	374	47.34	61.52	2	154	55.4	69.06
3	251	31.77	93.29	3	68	24.46	93.53
4	4	0.51	93.8	5	18	6.47	100
5	49	6.2	100	Total	278	100	
Total	790	100					

1-Never Married, 2-Married, 3-Widowed, 4-Separated, 5-Divorced

The variable was recoded as follows before the analysis.

$$MaritalStatus = \begin{cases} 1 & if \quad MaritalStatus = 2 \\ 0 & Otherwise \end{cases} \tag{4.2}$$

Admission Class

Table 4.6: Frequency Table for Admission Class

Orthopedic				Stroke			
Admission Class	Freq.	Percent	Cum.	Admission Class	Freq.	Percent	Cum.
1	785	97.64	97.64	1	278	97.2	97.2
3	17	2.11	99.75	3	7	2.45	99.65
4	1	0.12	99.88	4	1	0.35	100
5	1	0.12	100	Total	286	100	
Total	804	100					

1-Initial Rehabilitation, 2-Evaluation, 3-Readmission, 4-Unplanned Discharge, 5-Continuing Rehabilitation

The variable Admission Class has 5 levels. The Table 4.6 shows that, 97% of the stroke patients or orthopedic patients were admitted for the first time to rehabilitation. In order to assess whether readmissions has an effect on place of discharge, the variable Admission Class is recoded as follows,

$$AdmissionClass = \begin{cases} 1 & if \quad AdmissionClass = 3 \\ 0 & Otherwise \end{cases} \tag{4.3}$$

.

Pre-Hospital Living setting

The variable pre-hospital living setting contains information about the setting where patient was living prior to being hospitalized. There were a total of 13 categories representing the living setting coded from 1-10 and 12, 13, 14.

Table 4.7: Frequency Table for Pre-Hospital Living Setting

Orthopedic				Stroke			
Pre-Hospital Living Setting	Freq.	Percent	Cum.	Pre-Hospital Living Setting	Freq.	Percent	Cum.
1	745	95.64	95.64	1	260	94.89	94.89
2	1	0.13	95.76	5	2	0.73	95.62
5	3	0.39	96.15	7	1	0.36	95.99
7	10	1.28	97.43	8	1	0.36	96.35
9	8	1.03	98.46	9	4	1.46	97.81
10	3	0.39	98.84	14	6	2.19	100
14	9	1.16	100	Total	274	100	
Total	779	100					

1-Home, 2-Board and Care, 5-Skilled Nursing Facility, 6-Acute unit of your own facility, 7-Acute unit of another facility, 8-Chronic Hospital, 9-Rehabilitation Facility, 10-Other, 14-Assisted Living Residence

According to Table 4.7, 96% of the orthopedic patients and 95% of the stroke patients were living at home prior to the stroke or injury. However, there were no orthopedic patients or stroke patients admitted from an 'Acute unit of your own facility'. 'Chronic Hospital', and there were no stroke patients admitted from 'Board and Care facilities' or from the category 'other'. Most of the categories other than category 'Home' contain fewer patients. Thus, the variable pre-hospital living setting is recorded as

$$Pre-HospitalLivingsetting = \begin{cases} 1 \quad if \quad PreHospital\ Livingsetting = 1 \\ 0 \qquad\qquad\qquad\qquad\qquad otherwise \end{cases} \tag{4.4}$$

.

Place of Discharge

The variable place of discharge contains information about the setting where patient was discharged. There were a total of 14 categories representing the discharge setting coded from 1-14.

Table 4.8: Frequency Table for Place of Discharge

Orthopedic				Stroke			
Place of Discharge	Freq.	Percent	Cum.	Place of Discharge	Freq.	Percent	Cum.
1	644	80	80	1	198	69.23	69.23
2	3	0.37	80.37	4	1	0.35	69.58
4	1	0.12	80.5	5	20	6.99	76.57
5	29	3.6	84.1	7	38	13.29	89.86
6	2	0.25	84.35	8	1	0.35	90.21
7	69	8.57	92.92	9	1	0.35	90.56
8	5	0.62	93.54	11	1	0.35	90.91
11	1	0.12	93.66	13	13	4.55	95.45
13	27	3.35	97.02	14	13	4.55	100
14	24	2.98	100	Total	286	100	
Total	805	100					

1-Home, 2-Board and Care, 4-Nursing home, 5-Skilled Nursing Facility, 6-Acute unit of your own facility, 7-Acute unit of another facility, 8-Chronic Hospital, 9-Rehabilitation Facility, 10-Other, 11-Died, 13-Subacute Setting, 14-Assisted Living Residence

According to the Table 4.8, 80% of orthopedic patients were discharged to home, yet only 68% of stroke patients were discharged to home. By comparing orthopedic and stroke

patients, it is clear that stroke patients have a higher chance to be discharged into a Skilled Nursing Facility than orthopedic patients.

First, place of discharge was recoded as a binary variable for the GLLAMM model. Since the results were sucessfull, it was recoded again as a 3-level nominal variable. 1-home, 2- nursing facility, 3- hospital. There were two patients who died during the study. Those cases were excluded from the analysis.

4.3 Results

The GLLAMM model is utilized to assess the relationship between place of discharge and admission FIM. Since FIM items measure functional independence with error, two latent variables, motor and cognition are utilized. The package GLLAMM of the statistical software STATA was used to build the models (Rabe-Hesketh et al., 2004). First, a model is attempted allowing for both latent covariates, motor FIM and cognition FIM, to be included. However, the models did not converge with these two latent variables. The next option was to try the model with one latent variable measuring true FIM score. The internal consistency values for the GLLAMM models utilizing the latent covariate for the motor dimension of FIM was 5% for the orthopedic patients and 21% for stroke patients. Since the models lack validity and internal consistency in reporting true motor FIM scores, they were excluded from the result section. However, the models that utilized the latent covariate cognition were sound and the results are given in the following sections. In the following, covariates that affect place of discharge through the outcome model are called direct effects and covariates that affect place of discharge through the true covariate model are called indirect effects. Total effects are combined direct and indirect effects of these covariates.

4.3.1 Predicting Place of Discharge with Cognition for an Orthopedic patients

The variables, age, gender, prehospital living setting and the latent covariate cognition have a significant direct effect on the place of discharge (see Table 4.9). The odds ratio for the latent covariate cognition is 1.68. Hence the odds of discharge to home is 68% higher for a patient with a one unit increase in the cognition FIM sum score.

Table 4.9: Outcome Model for Orthopedic data

Place of Discharge	Coef.	Std. Err.	z	P>z	95% Conf.Interval	
primary pay source	0.1390	0.3197	0.43	0.664	-0.4877	0.7658
length of stay	-0.0077	0.0196	-0.4	0.692	-0.0462	0.0307
age	-0.0246	0.0101	-2.42	0.015	-0.0445	-0.0047
gender	0.4225	0.2124	1.99	0.047	0.0061	0.8389
ethnicity	0.5080	0.4219	1.2	0.229	-0.3189	1.3350
marital status	0.3664	0.2083	1.76	0.079	-0.0418	0.7747
admission class	0.5406	0.7299	0.74	0.459	-0.8899	1.9713
prehospital living	1.1745	0.4232	2.78	0.006	0.3450	2.0040
cog	0.5225	0.0802	6.51	0	0.3651	0.6799
intercept	-1.0909	1.1126	-0.98	0.327	-3.2718	1.0898

Table 4.10: True Covariate Model for Orthopedic data

Cog	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
primary pay source	0.3510	0.1293	2.72	0.007	0.0976	0.6045
length of stay	-0.0554	0.0096	-5.74	0	-0.0744	-0.0365
age	-0.0304	0.0043	-7.08	0	-0.0389	-0.0220
gender	0.1418	0.1019	1.39	0.164	-0.0578	0.3416
ethnicity	0.0578	0.1687	0.34	0.732	-0.2728	0.3885
marital status	0.1072	0.0956	1.12	0.262	-0.0802	0.2947
admission class	-0.0126	0.3504	-0.04	0.971	-0.6995	0.6742
pre-hospital living	0.7146	0.2360	3.03	0.002	0.2519	1.1772
intercept	6.9164	0.4502	15.36	0	6.0340	7.7989
res. var.	1.5337	0.0843			1.3728	1.7035

The variables, primary pay source, length of stay, age and pre-hospital living setting have a significant indirect effect on the place of discharge via the latent covariate cognition FIM (see Table 4.10). The total effect odds ratio of a directly observed covariate can be assessed by summing and then exponentiating the slope coefficients from the outcome and true covariate models (Rabe-Hesketh et al., 2003). The total effect odds ratio for primary pay source is 1.63. Thus, the odds of discharge to home is 63% higher for insured patient than self pay patient. Total effect odds ratio for length of stay is 0.938. Therefore, the odds of discharge into home is 7% higher for a patient with one day decrease in length of stay. If the unit of measurement for length of stay changes to 5 days, the estimated total odds ratio is 0.7290. Thus, the odds of discharge into home is 37% higher for a patient with a five day decrease in length of stay. The total effect odds ratio for age is 0.7591. Thus, the odds of discharge into a home is 32% higher for every 5 years decrease in age. Total effect odds ratio for gender is 1.75. Thus, the odds of discharge into a home is 75% higher for a female patient than for a male patient. The total effect odds ratio for prehospital living setting

is 6.6. Thus, the odds of discharge to home is 6.6 times for a patient admitted from home than from anywhere else. For the measurement model the internal consistency is 73%.

Table 4.11: Measurement Model for Orthopedic data

	Coef.	Std. Err.	95% Conf.Interval	
error var.	0.5651	0.0144	0.5374	0.5942
internal consistency	0.7307	0.0120	0.7065	0.7539

4.3.2 Predicting Place of Discharge with Cognition for Stroke Patients

The variables, ethnicity, pre-hospital living setting and the latent covariate cognition have significant direct effect on the place of discharge (see Table 4.12). The odds ratio for the latent covariates cognition is 1.514. Hence the odds of discharge into home is 51% higher for a stroke patient with one unit increase in the cognition FIM items.

Table 4.12: Outcome Model for Stroke data

Place of Discharge	Coef.	Std. Err.	z	P>z	95% Conf.Interval	
primary pay source	0.0974	0.3715	0.26	0.793	-0.6308	0.8257
length of stay	0.0224	0.0165	1.36	0.174	-0.0099	0.0549
age	-0.0049	0.0129	-0.38	0.703	-0.0303	0.0204
gender	-0.2835	0.3049	-0.93	0.352	-0.8811	0.3141
ethnicity	1.6565	0.6455	2.57	0.01	0.3912	2.9217
marital status	0.0592	0.3066	0.19	0.847	-0.5417	0.6602
admission class	-1.2310	0.9061	-1.36	0.174	-3.0070	0.5449
pre-hospital living	1.3340	0.6600	2.02	0.043	0.0403	2.6277
cog	0.4150	0.1070	3.88	0	0.2051	0.6248
intercept	-1.7840	1.4095	-1.27	0.206	-4.5467	0.9786

Table 4.13: True Covariate Model for Stroke data

Cog	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
primary pay source	-0.0618	0.2312	-0.27	0.789	-0.5150	0.3912
length of stay	-0.0352	0.0101	-3.47	0.001	-0.0550	-0.0153
age	-0.0224	0.0081	-2.76	0.006	-0.0383	-0.0064
gender	0.0097	0.1907	0.05	0.959	-0.3641	0.3836
ethnicity	-0.2231	0.2842	-0.78	0.432	-0.7803	0.3340
marital status	0.3584	0.1908	1.88	0.06	-0.0154	0.7324
admission class	-0.6156	0.5723	-1.08	0.282	-1.73745	0.5061
pre-hospital living	0.5148	0.4096	1.26	0.209	-0.2880	1.3177
intercept	5.4332	0.8383	6.48	0	3.790127	7.0764
res. var.	1.9903	0.1820			1.6495	2.3632

Additionally, the variables, age, length of stay and marital status have a significant

indirect impact on the place of discharge via the latent covariate cognition FIM (see Table 4.13). The total effect odds ratio for ethnicity is 4.19, thus, the odds of discharge into a home is 4.2 times higher for the Hispanic patients than non Hispanic patients. The total effect odds ratio for length of stay is 0.9873. Therefore, the odds of discharge into a home is 1% higher for a patient with one day decrease in length of stay. The total effect odds ratio for age is 0.8721, thus, the odds of discharge into a home is 14% higher for every 5 years decrease in age. The total effect odds ratio for marital status is 1.51, thus, the odds of discharge into a home is 51% higher for a married patient than for other patients. Total effect odds ratio for prehospital living setting is 6.35, thus, the odds of discharge into a home is 6.3 times that for a patient admitted from home than from anywhere else. For the measurement model the internal consistency is 78%.

Table 4.14: Measurement Model for Stroke data

error var.	0.5621	0.0243	0.5164	0.6119
internal consistency	0.7797	0.0175	0.7438	0.8126

4.3.3 Improved Type I Error Control for GLLAMM Models

This section utilized a method introduced by Sun et al. (2000) to improve the familywise type I error of the set of total effect odds ratios. The method produced two critical values for the orthopedic and stroke subpopulations. For orthopedic data critical value was 2.3827 and for stroke data critical value was 2.4112. Table 4.15 illustrates the modified confidence intervals for the odds ratios.

Table 4.15: Adjusted confidence intervals for type I error

Variable	Orthopedic	Stroke
Total Effects	estimate(adjusted C.I.)	estimate(adjusted C.I.)
primary pay source	0.0356 (0.4574, 2.3474)	0.4901 (0.5720, 4.6596)
length of stay[2]	-0.0635 (0.7244, 1.2157)	-0.3160 (0.5777, 0.9199)
age[1,2]	-0.2735 (0.5857, 0.9878)	-0.5509 (0.3994, 0.8317)
gender	-0.2737 (0.4354, 1.3282)	0.5643 (0.7426, 4.1634)
ethnicity[1]	1.4333 (1.4275, 12.3141)	0.5658 (0.3233, 9.5905)
marital status	0.4177 (0.8824, 2.6130)	0.4737 (0.6758, 3.8160)
admission class	-1.8467 (0.0232, 1.0720)	0.5279 (0.1300, 22.1124)
pre-hospital living[1,2]	1.8488 (2.0220, 19.9594)	1.8891 (1.0277, 42.5619)

1- variable is statistically significant for orthopedic population, 2- variable is statistically significant for stroke population

After multiplicity adjustments, variables age, ethnicity and pre-hospital setting are significant for the orthopedic patients and variables length of stay, age and pre-hospital setting are significant for stroke patients.

4.3.4 Place of Discharge as a Multinomial Variable

Recall the recoding of the response variable 'place of discharge', where 1 is home, 2 is nursing home, and 3 is hospital. This section illustrates the results of the GLLAMM model for orthopedic and cognition patients with the nominal response 'place of discharge'. In all cases, hospital was selected as the baseline category. The models didn't successfully converge for the stroke data, thus the results were limited to orthopedic data.

Orthopedic Data

The covariates length of stay, age, gender, pre-hospital living setting and cognition are significant for the model $\log(\frac{\Pi_1}{\Pi_3})$, where Π_1 is the probability of a patient being discharged home and Π_3 is the probability of a patient being discharge into a hospital. Covariates length of stay and cognition are significant for the model $\log(\frac{\Pi_2}{\Pi_3})$, where Π_2 is the probability of a patient being discharge into a nursing home. The relative log odds ratio of cognition for the first model is 1.67 and 1.96 for the second model. Thus, relative log odds of being discharge into a home vs. hospital is 67% higher for a one unit increase in cognition FIM items and relative log odds of being discharge into a nursing home vs. hospital is 96% higher for a one unit increase in cognition FIM items.

Table 4.16: Outcome Model for orthopedic data

placeofdis	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
c1						
primary pay source	0.2450	0.3773	0.65	0.516	-0.4946	0.9846
length of stay	0.0411	0.0243	1.69	0.091	-0.0065	0.0888
age	-0.0731	0.0212	-3.44	0.001	-0.1148	-0.0314
gender	0.5108	0.2443	2.09	0.037	0.0319	0.9896
ethnicity	0.3456	0.4526	0.76	0.445	-0.5415	1.2328
marital status	0.2834	0.8667	0.33	0.744	-1.4153	1.9822
pre-hospital living	1.0987	0.3493	3.1450	0.0008	0.4139	1.7834
cog	0.5134	0.1477	3.474	0	0.2238	0.8030
intercept	0.9704	1.1470	0.85	0.398	-1.2777	3.2186
c2						
primary pay source	-0.0074	0.6539	-0.01	0.991	-1.2890	1.2742
length of stay	0.1556	0.0357	4.36	0.000	0.0855	0.2256
age	0.0303	0.0191	1.58	0.113	-0.0071	0.0678
gender	0.2598	0.3875	0.67	0.503	-0.4997	1.0194
ethnicity	-1.2444	1.1054	-1.13	0.260	-3.4110	0.9221
marital status	-0.2165	0.3918	-0.55	0.581	-0.9846	0.5515
admission class	-0.4790	1.2614	-0.38	0.704	-2.9514	1.9934
pre-hospital living	-0.3666	0.6047	-0.61	0.544	-1.5518	0.8185
cog	0.67342	0.17236	3.9070	0	0.3355	1.0112
intercept	-5.2600	1.8728	-2.81	0.005	-8.9307	-1.5892

Table 4.17: True covariate Model for orthopedic data

Variable	Coef. (Std. Err.)
length of stay	0.0279 (0.0216)
age	-0.0098 (0.0074)
gender	0.0373 (0.0837)
pre-hospital living	0.3612 (0.3227)

According to these results, the variables age, gender, pre-hospital living setting and length of stay have significant indirect impact on the place of discharge via the latent covariate cognition FIM. The total effect relative log odds ratio for age is 0.6602. Thus, relative log odds of being discharge into a home vs. hospital is 51.4% higher for every 5 years decrease in age. The total effect relative log odds ratio for gender is 1.72. Therefore, the relative log odds of being of discharged home vs. hospital is 72% higher for a female patient than for a male patient. The total effect relative log odds ratio for prehospital living setting is 4.3, thus, the relative log odds of being discharged home vs. hospital is 4.3 times for a patient admitted from home than from anywhere else. The total effect relative log odds ratio for length of stay is 1.2014; the relative log odds of being discharged into a nursing home vs. hospital is 20% higher for a patient with one day increase in length of stay.

4.3.5 SEM for Difference of FIM

The SEM model was used to find the patient characteristics that affect the FIM difference scores. FIM difference was calculated using admission and discharge FIM items. Since FIM items measuring functional independence with error, two latent variables, motor and cognition are utilized. The package called SEM of the statistical software STATA was used to build the models (StataCorp, 2011). However, the MLE estimation method of SEM requires multivariate normality. Since the FIM difference score violate the normality

assumption, the estimation method ADF was used.

Orthopedic Data

Table 4.18: SEM for orthopedic data

Structural	Coef.	Std. Err.	z	P>z	[95% Conf.Interval]	
Motor						
primary pay source	0.3435	0.0731	4.7	0	0.2002	0.4868
length of stay	0.0106	0.0269	0.4	0.691	-0.0420	0.0634
age	-0.00716	0.0025	-2.76	0.006	-0.0122	-0.0020
gender	0.3339	0.1076	3.1	0.002	0.1228	0.5449
marital status	0.4115	0.0751	5.48	0	0.2642	0.5588
admission class	0.5266	0.1410	3.73	0	0.2501	0.8031
pre-hospital living	1.3128	0.1422	9.23	0	1.0341	1.5916
ethnicity	-1.0061	0.1411	-7.13	0	-1.2826	-0.7295
Cog						
primary pay source	0.1346	0.0785	1.71	0.087	-0.0193	0.2885
length of stay	0.0251	0.0082	3.06	0.002	0.0090	0.0413
age	0.0074	0.0030	2.46	0.014	0.0015	0.0133
gender	-0.16976	0.0781	-2.17	0.03	-0.3228	-0.0166
marital status	-0.0124	0.0807	-0.15	0.878	-0.1706	0.1458
admission class	-0.2295	0.1006	-2.28	0.023	-0.4267	-0.0323
pre-hospital living	-0.0118	0.0922	-0.13	0.898	-0.1925	0.1689
ethnicity	-0.4346	0.1187	-3.66	0	-0.6673	-0.2018

The variables primary pay source, age, gender, marital status, admission class, pre-hospital living setting and ethnicity have significant effect on the latent covariate motor (see Table

4.18). The results indicate, an insured patient has a 0.3435 higher motor difference FIM score than a self pay patient and a 10 years increase in age will reduce the motor FIM difference score by 0.0716. It also shows female patients have a 0.3340 higher motor FIM difference score than male patients and married patients have a 0.4116 higher motor FIM difference score than other patients. If the patient is readmitted then the motor FIM difference score is increased by 0.5266. A patient admitted from home has a 1.3129 higher motor FIM score than a patient admitted from other places and being a Hispanic patient will reduce motor FIM difference score by 1.0061.

Similarly, variables length of stay, age, gender, admission class,ethnicity have significant effect on the latent covariate cognition (see Table 4.18). The results show for each 10 day increase in length of stay, the cognition FIM difference score will increased by 0.2518 and every 10 years increase in age will increase the cognition FIM difference score by 0.0742. Being a female patient will reduce cognition FIM difference score by 0.1697 as compared with a male patient and if a patient is readmitted then the cognition FIM difference score is decreases by 0.2295. Also, being a Hispanic patient will reduce the cognition FIM difference score by 0.4346. The covariance coefficient between the latent variables cognition and motor was 0.1952 and it was significant with a p value < 0.001.

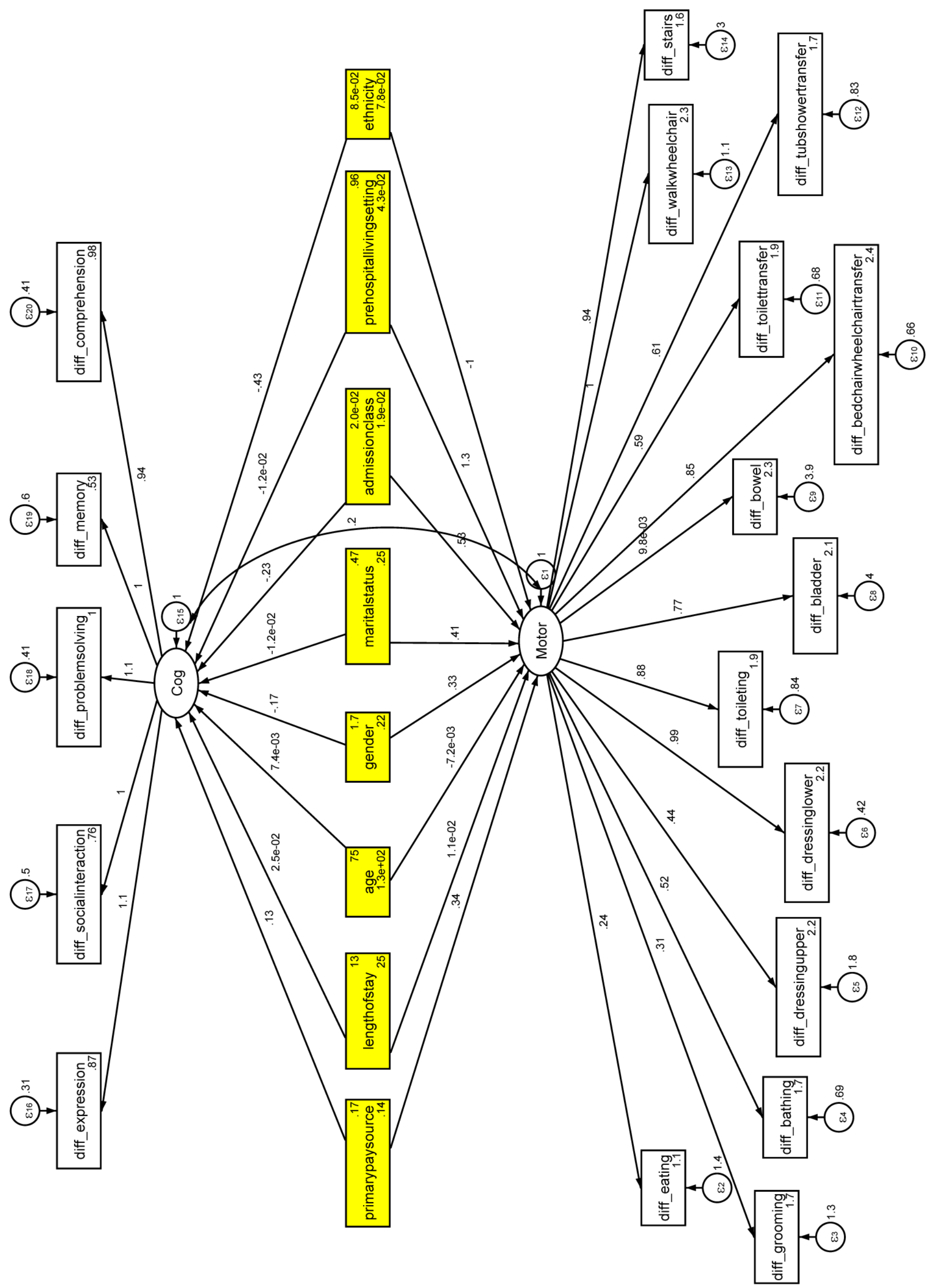

Figure 4.5: SEM for orthopedic data

Stroke Data

Recall that the motor category of FIM is likely to be multi-layered and multi-dimensional. Stineman et al. (1997) concluded that the motor scores had poor internal consistency , and did not align to a unidimensional construct. Thus, the SEM will only predict the cognition items of FIM for stroke data.

Table 4.19: SEM for Stroke Data

Structural	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
Cog						
primary pay source	-0.0341	0.1255	-0.27	0.786	-0.2802	0.2119
length of stay	0.0539	0.0067	8.03	0	0.0408	0.0671
age	-0.0224	0.0051	-4.38	0	-0.0324	-0.0124
gender	-0.1765	0.1264	-1.4	0.163	-0.4244	0.0714
marital status	0.2501	0.1195	2.09	0.036	0.0159	0.4844
admission class	-1.6713	0.2946	-5.67	0	-2.2488	-1.0938
pre-hospital living	-0.0940	0.2086	-0.45	0.652	-0.5029	0.3148
ethnicity	0.7559	0.1617	4.67	0	0.4389	1.0729

Table 4.19 displays the variables length of stay, age, marital status, admission class and ethnicity have significant effect on the latent covariate cognition. The results show for each 10 day increase in length of stay, the cognition FIM difference score increase by 0.0540 and for every 10 years increase in age decrease the cognition FIM difference score by 0.2244. A married patient has 0.2502 higher cognition FIM difference score than other patients. When a patient is readmitted, the cognition FIM difference score decreases by 1.6713. Finally, being a Hispanic will increase the cognition FIM difference score by 0.7560.

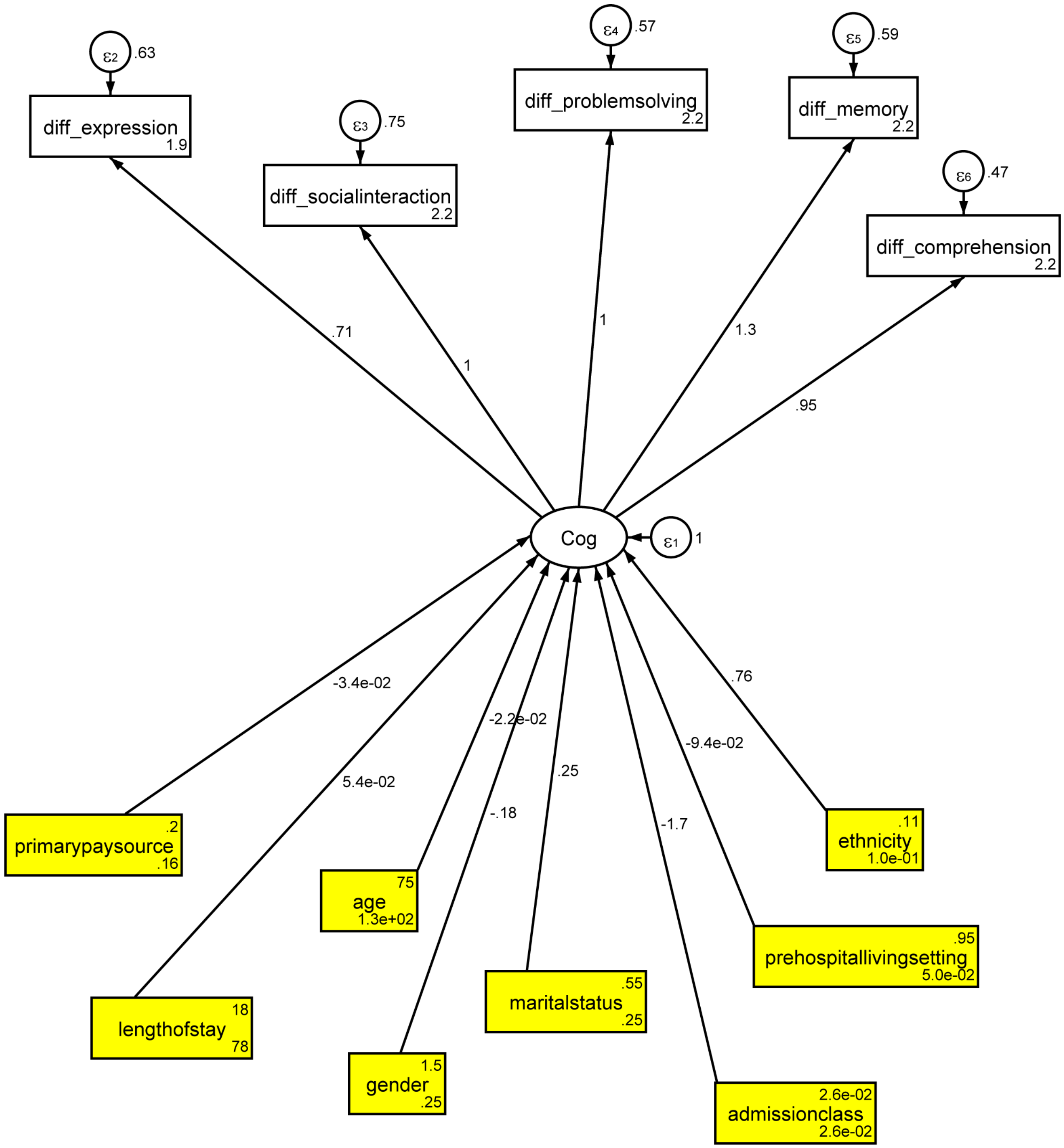

Figure 4.6: SEM for Stroke data

Chapter 5

Discussion

In the following, the overall results of the statistical analysis of FIM scores for orthopedic and stroke rehabilitation patients are discussed. Implications for further research are also provided.

5.1 Connection to the past research

Past studies questioned the uni-dimensionality of the motor subset of FIM in various populations (Stineman et al., 1997). In this study, the motor FIM items exhibit poor internal consistency and do not predict place of discharge well. On the contrary, cognition FIM items predict place of discharge well with a higher internal consistency . Stineman et al. (1997) also explained that the different impairment conditions can have varying effects on the FIM items. Our results indicate a similar trend. Orthopedic patients have higher odds of being discharged to home for the same unit of increase in cognition FIM items when compared to stroke patients.

5.2 Implications of the results

The main objective of this study was to help the hospital staff to improve their services to the patients by addressing their questions. GLLAMM models were built to find how the admission FIM affects place of discharge. According to the models, an orthopedic or a stroke patient with a 10 unit age increase has lower odds of being discharged home. Also orthopedic Hispanic patients have higher odds of being discharged home than any other

patient group. For a stroke patient with a 5 unit increase in length of stay, the odds of being discharged home is decreased. For both stroke and orthopedic patients, there is higher odds of being discharged home if they were admitted from home than anywhere else.

Table 5.1 shows the probability of being discharged home for different levels of mean cognition scores. Orthopedic patients have 50% or higher odds of being discharged home if their mean cognition score is higher than 2.5. Similarly, stroke patients have 50% or higher odds of being discharged home if their mean cognition score is higher than 4.

Table 5.1: Probability of place of discharge

	Probability of being discharged home	
Mean cognition score	Orthopedic	Stroke
very low (1)	0.3616	0.2028
low of stay (2.5)	0.5536	0.3216
medium (4)	0.7309	0.4690
high (5.5)	0.8561	0.6221
very high (7)	0.9287	0.7542

One other concern of the hospital staff was to identify the patient characteristics that affects the FIM difference scores. SEM models were used to assess this. According to the models, patients entering the hospital from home significantly improve motor FIM difference scores. On average, Hispanics tend to decrease the motor dimension of FIM by a full point. However, there may be factors that more directly create this effect, rather than ethnicity. Interestingly patients that were re-admitted tend to increase the motor dimension of FIM by over half a point. Even though the variable age is statistically significant for both motor and cognition FIM difference scores, the change in motor and cognition FIM difference scores due to age is small and thus, practically not meaningful. There is little change in the cognition FIM scores for orthopedic patients. The factors that statistically affect the cognition FIM difference scores, probably do not have strong practical significance

(possibly with exception of ethnicity). The variables admission class and ethnicity affect cognition FIM difference scores for stroke patients. Those readmitted lower the cognition FIM difference scores, while Hispanic patients increase the difference score.

5.3 Future research

Before the site visit, it was assumed that the data contains more than one hospital. However due to various reasons the data was limited to one hospital. Thus the result was limited to that particular hospital. This limitation can be avoided in future studies by including data on patients from more than one hospital.

In this study, FIM items were divided into two levels as motor and cognition. However, the models related to motor FIM items had convergence issues and low internal consistency. Lack of uni-dimensionality of motor FIM items might be the reason behind this issue. Stineman et al. (1997) suggested that motor FIM item can be further divided into two sub-levels. Due to time constraints we were unable to try the models with these sub-levels.

References

Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. John Wiley & Sons, Ltd.

Amundson, J., Brunner, A., and Ewers, M. (2000). Fim scores as an indicator of length of stay and discharge destination in cva patients: A retroactive outcomes study. *Fim scores as an indicator of length of stay*, pages 263–270.

Everitt, B. S. and Howell, D. C. (2005). *Generalized Linear Models (GLM)*. John Wiley & Sons, Ltd.

Kaplan, R. (2011). Acute hospital discharge: Why the right place is the best place. Technical report, SeniorMetrix, Boston, MA.

Mauthe, R., Haaf, D., Hayn, P., and Krall, J. (1996). Predicting discharge destination of stroke patients using a mathematical model based on six items from the functional independence measure. *Arch Phys Med Rehabil*, pages 10–13.

Muthen, B. and Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal likert variables. *British Journal of Mathematical and Statistical Psychology*, pages 171–189.

Rabe-Hesketh, S., Skrondaly, A., and Picklesz, A. (2002). Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal*, pages 1–21.

Rabe-Hesketh, S., Skrondaly, A., and Picklesz, A. (2003). Maximum likelihood estimation of generalized linear models with covariate measurement error. *The Stata Journal*, pages 386–411.

Rabe-Hesketh, S., Skrondaly, A., and Picklesz, A. (2004). Gllamm manual. *U.C. Berkeley Division of Biostatistics Working Paper Series*.

Sandstrom, R., J., P., Mokler, K., and Hoppe, M. (1998). Discharge destination and motor function outcome in severe stroke as measured by the functional independence measure/function-related group classification system. *Arch Phys Med Rehabil*, pages 762–765.

Seiger, C. (2011). The right person in the right place : Rehabilitation criteria for discharge placement recommendations. Technical report, Department of Physical and Occupational Therapy, Idaho State University, Pocatello, ID.

StataCorp (2011). *Stata Structural Equation Modeling.*

Stineman, M. and Fiedler, R. (1998). Functional task benchmarks for stroke rehabilitation. *Arch Phys Med Rehabil*, pages 497–504.

Stineman, M., Jette, A., Fiedler, R., and Granger, C. (1997). Impairment-specific dimensions within the functional independence measure. *Arch Phys Med Rehabil*, pages 636–646.

Sun, J., Loader, S., and McCormick, D. (2000). Confidence bands in generalized linear models. *The Annals of Statistics*, 28(2):429–460.

Tabachnick, B. and Fidell, L. (1996). Using multivariate statistics. *NY: HarperCollins*, page 579.

Viele, K. (2008). Newton raphson algorithm @ONLINE.

Wieslaw, J. and Ban-ecu, S. (1997). Neural network modeling accurately predicts the functional outcome of stroke survivors with moderate disabilities. *Arch Phys Med Rehabil*, pages 340–345.

Curriculum Vitae

Mudiyanselage Maduranga Kasun Dassanayake was born on July 03, 1984 in Colombo, Sri Lanka. He obtained his high school education from Royal College, Colombo 7 and graduated in the Spring of 2003. He entered to the University of Colombo, Sri Lanka in Fall 2005 and obtained his Bachelors degree in Statistics with honors in the Summer of 2009. He also did another bachelor in Information technology at University of Colombo School of Computing and graduated with second class upper in 2008. He joined to the Department of Statistics of University of Colombo as an instructor on record and worked there for a period of one year. During this time he publish a paper based on final year research project, named Predicting Trading Signals of Sri Lankan Stock Market using Genetic Algorithms and Neural Networks. The published paper is available in Springer, in the book 'Technological Developments in Networking, Education and Automation'. Maduranga moved to El Paso in the Fall of 2010 for his Masters degree in Statistics at UTEP. While studying, he worked as a Teaching Assistant to the Department of Mathematical Sciences. He worked at the Center for Institutional Evaluation, Research and Planning of UTEP (CIERP) during the Summer of 2011 as a research assistant. He also awarded the Academic Excellence award for Statistics at the Spring 2012 Pre-Commencement, University of Texas at El Paso. He is willing to start his doctoral studies in Statistics next year.

Permanent address: 236 Moratuwa Road,
Suwarapola, Piliyandala, Sri Lanka.