A LONGITUDINAL ANALYSIS FOR THE IDENTIFICATION OF THE FACTORS THAT AFFECT THE CASE MIX INDEX OF HOSPITALS IN THE U.S

Danai Chasioti

Submitted to the faculty of the University Graduate School

in partial fulfillment of the requirements

for the degree

Master of Science

in the Richard M. Fairbanks School of Public Health,

Indiana University

December 2017

Accepted by the Graduate Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Master of Science.

Master's Thesis Committee

Constantin T. Yiannoutsos, PhD _____

Ying Zhang, PhD _____

Barry P. Katz, PhD _____

Giorgos Bakoyannis, PhD _____

Danai Chasioti

A Longitudinal Analysis for the identification of the factors that affect the Case Mix Index of Hospitals in the U.S.

The present thesis is an analysis of longitudinal data collected through the years 2011-2013, from a complex of four hospitals located in Indiana, USA. The aim of the analysis was the detection of changes (especially a decline) in the disease related group (DRG) weights (and thus, the case mix index (CMI)), and the determination of the predictors that significantly affect these changes.

The document is divided in four major parts. In the first part it is described the statistical theory required for the the analysis, in the second part the reimbursement strategies for the hospitals in the USA, are briefly described and the concept of the DRG and CMI are explained. In the third part the actual analysis is presented while the last part contains a summary of the findings and some conclusions.

The correlation between the observations was taken into account by modeling the data using linear mixed models (LMM). Three major factors were studied for their effect on the DRG weight of the hospitals: the changes in the type of cases (i.e. the product lines), the changes in the number of the Surgical cases, and also the changes of the length of stay (LOS). The analysis did not indicate any significant DRG change in any of the hospitals except from the H4. The H4 hospital has a significant decline over time regarding the Cardio-vascular (CV) DRG weights. For the hospitals H1, H2 and H3 the only decline observed in the product lines was that for the Medical-Surgical DRG. Finally, no significant change was observed for the LOS, or the number of Surgical cases.

In addition to the three predictors studied, changes in the coding system, the documentation etc. may also affect the DRG and CMI. However, these changes are not possible to be detected through this analysis, since no available information was given in the present data.

Constantin T. Yiannoutsos

Contents

Chapter 1

Theoretical Background

1.1 Linear Models for Continuous Longitudinal Data

1.1.1 Overview

In longitudinal studies, the investigators collect information about specific characteristics of a cohort over time. The observations are collected on individual level and may differ in number and time point of collection.

Let Y_{ij} be the response variable for the individual i at the j^{th} time measurement . To accommodate unbalanced measurements, we assume that each individual i has n_i observed responses. Let N the total number of the individuals in the study. For each of these individuals there is an $n_i \times 1$ corresponding vector of responses

$$Y_i = \begin{bmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{in_i} \end{bmatrix}, i = 1, \ldots, N$$

where, $Y_{i1}, \ldots, Y_{in_i}$ are dependent to each other (within-individual association). At each measurement occasion j, the values of the (p) covariates X_{ijl} $(l = 1, \ldots, p)$ are measured. Thus, to each individual i corresponds a design matrix X_i with dimensions $n_{i \times p}$

$$X_i = \begin{bmatrix} X'_{i1} \\ X'_{i2} \\ \vdots \\ X'_{in_i} \end{bmatrix} \Leftrightarrow X_i = \begin{bmatrix} X_{i11} & \ldots & X_{i1p} \\ X_{i21} & \ldots & X_{i2p} \\ \vdots & \ldots & \vdots \\ X_{in_i1} & \ldots & X_{in_ip} \end{bmatrix}, i = 1, \ldots, N$$

Each of the Y_{ij} responses is described by a linear regression model :

$$Y_i = X_i\beta + e_i \text{ where,} \tag{1.1}$$

Y_i the $n_i \times 1$ matrix of responses, X_i the $n_i \times p$ matrix of covariates, β the $p \times 1$ vector of regression coefficients and e_i the $n_i \times 1$ vector of the random errors (that is, the e_{ij} is the deviation of the predicted $E(Y_{ij}|X_{ij})$ from the observed Y_{ij} response).

$$Y_i = \underbrace{X_i\beta}_{\text{Systematic part}} + \underbrace{e_i}_{\text{Random part}} \quad \text{where,}$$

the vector Y_i follows a Multivariate Normal distribution with mean $E(Y_i) = \mu_i = X_i\beta$ and covariance matrix $Cov(Y_i) = \Sigma_i$. Each of the elements $Y_{ij} \sim N(\mu_{ij}, \sigma_j^2)$. The within-subject association, between the repeated measurements of each individual, is expressed by the off-diagonal elements of the covariance matrix $\Sigma_i = Cov(Y_i)$.

In order to properly take into account the within-subject association of the repeated measurements, we need not only to model the mean response but also the covariance. Bellow it is presented a brief presentation of the alternative ways used to model the covariance (See Figure 1.1).

To model the repeated measurements' covariance, we can either assume specific pattern for the covariance matrix (covariance pattern models) or allow the covariance to vary randomly over time (unstructured covariance). In both of these cases, the mean and the variance are modeled separately (but not independently). Finally, we can account for the correlation by assuming a covariance pattern in the same model used to model the mean responses (the random effects in the mixed models induce association between the responses).

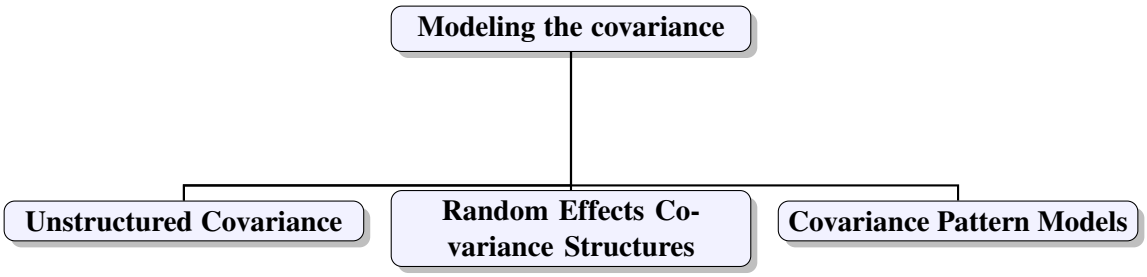


Figure 1.1: Modeling the Covariance

By correctly determine the within-subject covariance pattern and by assuming a model for the mean response, we can get accurate estimates of the regression parameters . When the investigator's goal is the interpretation of the findings on population's level then, the mean response can be modeled

using either the response profiles or the parametric/semi-parametric curves methods (See Figure 1.2).

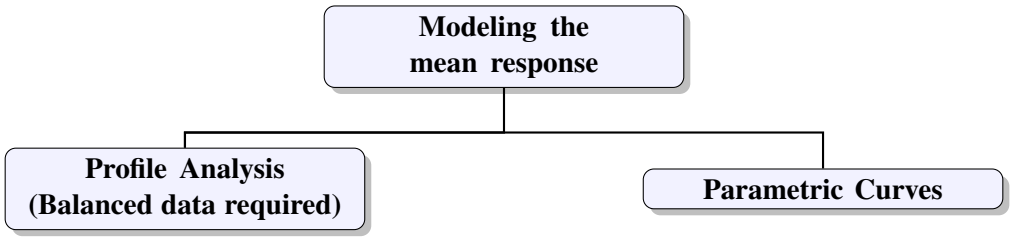


Figure 1.2: Modeling the Mean response

When the measurements have been taken at common time points for each individual in the study (the design is balanced), and the number of the measurements is small, the profile analysis can be used to describe the mean changes of the response over time. A general form of the response profile model is presented below:

$$Y_i = X_i\beta + e_i \quad \text{where,} \tag{1.2}$$

X_i is a $j \times k$ design matrix (for the subject i) with $j = 1,\ldots,m$ measurement occasions. The X_i can be written as

$$\begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_l \end{bmatrix} \text{ where,}$$

$G_1, G_2, \ldots, G_l$ are $m \times m$ square sub-matrices and $l = 1,\ldots,n$ is the total number of groups. So, the matrix X_i has in total $k = l \times m$ columns.

This type of analysis is indicated when the study design is balanced and the number of covariates is small. The response profile analysis incorporates the time variable in its categorical form resulting to a saturated model that over fits the data. Also, it uses more parameters for the mean change over time than it is required (usually the changes over time are not sharp) and has less power to detect departs from the null hypothesis. So, a more parsimonious model with higher power is desirable. An alternative to the response profile analysis would be the use of parametric/semi-parametric models. By assuming the correct mean response pattern over time, these models have greater power to test

changes over time compared to the response profiles. In addition, the parametric models can handle the unbalanced and incomplete data (which is common characteristic of the data in longitudinal studies). Linear, quadratic, higher order trends and linear splines are used to describe the change of mean response over time. In this case, time is treated as a continuous variable. Thus, a design matrix X_i for a quadratic model and an exposed individual, could be the following:

$$X_i = \begin{bmatrix} 1 & t_{i1} & t_{i1}^2 & 1 & 1 & t_{i1}^2 \\ 1 & t_{i2} & t_{i2}^2 & 1 & 1 & t_{i2}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_{in_i} & t_{in_i}^2 & 1 & 1 & t_{in_i}^2 \end{bmatrix}$$

where, the fourth column is the group variable (1 for treatment/exposure and 0 for the unexposed/controls) and the last two columns represent the $group \times time$ interaction and the $group \times time^2$ respectively. The number of the repeated measurements for the i^{th} individuals is n_i.
Given that the data on which the parametric/semi-parametric models are applied on, is usually unbalanced, an unstructured covariance pattern would lead to an extremely high number of parameters to estimate. That implies that, a more "conservative" pattern should be adopted in order to get valid estimates for the regression parameters.

1.1.2 Linear Mixed Models

Whenever the goal of the investigator is to predict the individual trajectories, the use of mixed models is indicated. In addition, the mixed effects models combine the advantages of the parametric curve models, in terms of their ability to handle unbalanced data and be parsimonious (small number of covariance parameters to be estimated).

A linear mixed effects model applies to longitudinal data where the response variable is continuous and the residuals e_{ij} follow a Normal distribution. The inference about an individual is becoming possible since the model "breaks" the inherited variability into two parts: the population variability and the individual variability (i.e. the between-subject and within-subject variability). In addition to the ability to make individual-based predictions/inferences, the linear mixed effects models have

the ability to model the covariance of the repeated measurements by incorporating one (or more) random effects. A compound symmetry covariance pattern is introduced by including a random intercept in the mixed model. A random intercept term is translated as an deviation of an individual's mean response from that of the population's which, is constant over time. Such a model is expressed by the following function:

$$Y_i = X_i\beta + Z_i b_i + e_i \text{ where,} \tag{1.3}$$

X_i is an $n_i \times p$ matrix of variables, β is a $p \times 1$ vector of fixed-effect coefficients, Z_i is the $n_i \times q$ design matrix of random effects ($q \leq p$) and b_i the $q \times 1$ vector of random-effect coefficients while, e_i is an $n_i \times 1$ vector of measurement/sample errors. Assuming independency between the random-effect coefficients (b_i) and the errors (that is, $corr(b_i, e_i) = 0$), we derive the following relation:

$$\begin{aligned} Cov(Y_{ij}, Y_{ik}) &= Cov(X'_{ij}\beta + Z'_{ij}b_i + e_{ij}, X'_{ik}\beta + Z'_{ik}b_i + e_{ik}) = \\ &= Cov(Z'_{ij}b_i + e_{ij}, Z'_{ik}b_i + e_{ik}) = Z_{ij}Cov(b_i, b_i)Z'_{ik} \Rightarrow \\ &\Rightarrow Cov(Y_{ij}, Y_{ik}) = Z_{ij}GZ'_{ik} \text{ where,} \end{aligned} \tag{1.4}$$

G is an $q \times q$ square covariance matrix for the random-effect coefficients vector b_i and, Z_{ij} an $n_i \times 1$ vector from the random effects design matrix for the individual i at the j^{th} measurement occasion. Specifically for a random intercept model, Z_i is an $n_i \times 1$ vector, b_i is an 1×1 element and, G is an 1×1 element as well. Thus,

$$Cov(Y_{ij}, Y_{ik}) = \sigma_b^2, \ \forall j \neq k \tag{1.5}$$

for the off-diagonal elements of the covariance matrix $Cov(Y_i)$ while, the diagonal elements of the same matrix are:

$$\begin{aligned} Cov(Y_{ij}, Y_{ij}) &= Cov(X'_{ij}\beta + Z'_{ij}b_i + e_{ij}, X'_{ij}\beta + Z'_{ij}b_i + e_{ij}) = \\ &= Z_{ij}Cov(b_i, b_i)Z'_{ij} + Cov(e_{ij}, e_{ij}) \Rightarrow \\ &\Rightarrow Var(Y_{ij}) = \sigma_b^2 + \sigma^2, \ j = 1, \ldots, n_i \end{aligned} \tag{1.6}$$

That is,

$$Var(Y_i) = \begin{bmatrix} \sigma_b^2 + \sigma^2 & \sigma_b^2 & \dots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 + \sigma^2 & \sigma_b^2 & \sigma_b^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_b^2 & \sigma_b^2 & \dots & \sigma_b^2 + \sigma^2 \end{bmatrix} \tag{1.7}$$

Notice that, the random intercept model introduces a compound symmetry covariance pattern for the data. A more general covariance pattern can be adopted by incorporating more than one random effects in the model. In that case, the marginal covariance is:

$$Cov(Y_i) = Cov(X_i\beta + Z_i b_i + e_i) = Z_i Cov(b_i) Z_i' + Cov(e_i)$$

For $Cov(b_i) = G$ and $Cov(e_i) = R_i = \sigma^2 I$ we get,

$$Cov(Y_i) = Z_i G Z_i' + R_i \tag{1.8}$$

So, $Y_i \sim N(X_i\beta,\ Z_i G Z_i' + R_i)$, $e_i \sim N_{n_i}(0, R_i)$, $b_i \sim N_q(0, G)$ while the subject-specific response $Y_i|b_i \sim N(X_i'\beta + Z_i' b_i, R_i)$, $Cov(Y_i|b_i) = Cov(e_i) = R_i$ the conditional covariance and

$$E(Y_i|b_i) = E(X_i\beta + Z_i b_i + e_i) = X_i\beta + Z_i b_i \tag{1.9}$$

which has an individual-based interpretation. In contrast, the population based inference uses the marginal mean response

$$E(Y_i) = X_i\beta \tag{1.10}$$

1.2 Generalized Linear Models for Discrete Longitudinal Data

1.2.1 Marginal Models

Frequently, the response of interest is a binary, count or even a continuous variable which, does not fulfill the distribution assumption of normality. In order to handle such situations, the generalized linear models (adequately modified to handle longitudinal data) are available. Again, there are

different models depending on whether it is of interest a population or an individual-based inference. For the former case of population-based inferences, models which contain no random effects, known as marginal models, can be used. The within-subject association of the responses is taken into account by determining the correlation pattern $Corr(Y_{ij}, Y_{ik})$ among the repeated measurements in the case of continuous responses/counts and by determining the $log(OR)[Y_{ij}, Y_{ik}]$ for the case of binary responses.

In contrast to the continuous response, the discrete responses (e.g., counts, binary outcome) have a restrictive range which might not always coincide with that of the predictions' $X_i\beta$. To overcome this issue, the marginal models, similarly to the GLMs, use link functions:

$$g(E(Y_{ij})) = g(\mu_{ij}) = \eta_{ij} = X'_{ij}\beta \tag{1.11}$$

A link function maps the elements of the mean response's range to the domain of the covariates.

1.2.2 Generalized Linear Mixed Models

While in the marginal models the association between the measurements is taken into account by separately modeling the correlation (or the $log(OR)$; depending on the nature of the response variable), with the Generalized Linear Mixed Models (GLMMs), the introduction of random effects induces the association.

In contrast to the marginal models which, do not require any distributional assumptions for estimating the parameters, for the GLMMs we need to specify the distribution of the outcome Y_{ij} conditioning on the random effects vector b_i. The distribution for the vector of the random effects is necessary as well. Similarly to the Generalized Linear Models (GLMs), we can use the same idea to model continuous, binary or count data and similarly to Linear Mixed Models (LMMs), the within-subject correlation is taking into account by the incorporation of the random effects in the model. A general form of a GLMM is:

$$\eta_{ij} = E(Y_{ij}|b_i) = X'_{ij}\beta + Z'_{ij}b_i \text{ where,} \tag{1.12}$$

η_{ij} is the link function. Since the estimation of the mean response is conditional on the subject-

specific random effects, the interpretation is also subject-specific as opposed to the population-based interpretation of the marginal models. Thus, the estimated coefficient of a fixed effect indicates the average individual's change of its transformed mean, given the values of the random variables at the specific time point and for a unit change of the (time varying) variable X_{ill}.

An exception to the above interpretation is the case of an identity link function $\eta_i = 1$ where, the conditional mean response coincides with the interpretation of the population mean.

$$g(E(Y_{ij}|X_{ij}, b_i)) = \eta_i = E(Y_{ij}|X_{ij}, b_i) = E(X'_{ij}\beta + Z'_{ij}b_i) = X'_{ij}\beta \tag{1.13}$$

Chapter 2

Health Services in U.S.

Back in 1970, Yale university introduced a system for the classification of similar inpatient cases in hospitals, known as the Disease Related Groups (DRGs) [1]. Each case is represented by one DRG code based on the patient's demographic information such as age and sex, as well as the primary and secondary diagnosis, the possible co-morbidities, complications, surgery, other procedures performed and discharge status [2], [6]. There are total 474 DRGs [3]. The DRG classification was initially used by the private insurance systems while, in the early '80s the U.S health system adopted the idea of DRGs by, incorporating it in a DRG-based payment system, primarily by the Medicare program (1983) [1], [3], [4]. An accurate use of the DRG system will help control the cost and increase the resources' efficiency.

Known as "case-based" or "case-mixed-case" [5], the DRG system is based on classification algorithms in order to map each hospital case to one of the DRGs. There are totally 6 algorithms (called groupers) for DRG classification two of which are not used any more (the CMS-DRG and the AP-DRG) [3]. Cases belonging in the same DRG, are not only medically similar but also financially [5]. Inter alia, the role of the DRG classification in the payment system is to provide an average level of hospital resources required to treat a patient within a group. That is obtained by assigning a relative weight to each of the DRGs. Consecutively, Medicare's payment to the hospital is an amount proportional to this relative weight. By that, it is implied that the DRG coding system can affect the hospital's payment through the DRG's impact on the Case Mix Index (CMI) values (Table 2.1).

Table 2.1: Changes in DRG Case-Mix Indexes and in average DRG weight due to documentation/code system improvements, 1981-1987 [2]

	Actual change					Percentage change			
	1981	1984	1985	1986	1987	1981-84	1984-85	1985-86	1986-87
Average CMI	1.00	1.06	1.09	1.11	1.13	5.93%	3.09%	2.04%	1.66%
Average case weight	1.05	1.13	1.18	1.21	1.24	7.73%	4.37%	2.88%	2.40%
Adjusted average case weight	1.05	1.13	1.19	1.23	1.26	7.75%	5.52%	2.88%	2.39%

The CMI, which is directly related to the DRG weights, is a measure of the complexity of a hospital's cases. Developed in 1981, the CMI is the basis for payment methods, and it is calculated

by averaging the DRG relative weights of a hospital, for each time period (e.g. month, quarter, year etc.). There are five categories of CMI in total (Overall, Medical, Surgical, Adjusted, Medical/Surgical mix and volume-adjusted CMI) [8]. Based on the definition of the Case Mix Index, we can see that changes of its levels over time, affect the revenue stream of the hospital. Specifically, the higher the CMI level, the higher the reimbursements/revenue for that hospital since, a high CMI level suggests a higher-weight DRG cases and thus, higher amount to receive per patient. The DRG-based payment system depends on several factors that affect its accuracy and efficiency which are not always easy to detect and control. Hence, the coding accuracy and documentation can have a significant impact on the hospital's revenue by declaring an over or under-estimated CMI level (relative DRG-weight).

The intentionally false coding that indicates a higher than the real case complexity (and thus, a higher reimbursement amount), is a weak point in the relation between the hospitals and the government. The improvement in the documentation system and the detection of other factors that could possibly be responsible for the observed CMI (or DRG-weight) changes, could soothe the hospital-government relation by revealing the truth behind the changes.

In addition to the technical reasons that could affect the CMI levels, such as coding and documentation, the change in mix of weights (that is change in the DRG weights), the decrease of the Length Of Stay (LOS) as well as the technological and treatment improvements (less inpatient cases) can also have an significant effect on the fiscal measures of a hospital. The emerge for the study of the CMI changes emanate from the need to set fair payments as well as from to further discover how the health system works.

The substitutions of the retrospective cost-based payment system by the DRG-based payment system (PPS), gave rise to several incorrect coding incidences. A study held by David C. Hsia in 1988 [9], [10], shows that for the time period of October 1984 to March 1985, the chance that a hospital will inaccurately classify a case was 20.8%. It is noteworthy that, while the case-based system would most probably lead to a hospital's underpayment in case of coding errors, under the PPS, 61.7% of the coding errors would favor the hospital towards an over-payment. The changes in CMI (especially when they are upward) need to be controlled and restricted when, they are the result of intentional (or unintentional) errors. The source of false CMI change, based on a sample of the

1986-1987 Medicare discharged data, is 33.3% [11] of the total change for that time period. From that 33.3%, 85% [11] was due to hospitals' incorrect coding, 0.7% [11] of the CMI's divergence can be attributed to the sensitivity differences between the several Groupers. A direct consequence of the over-coding is the effect on the level of hospital's CMI and thus, the amount of reimbursement. The importance of controlling such incidence can be explained by the fact that, the expenditure resulted for 1% increase of the CMI level is equivalent to $40 million for the government [11].

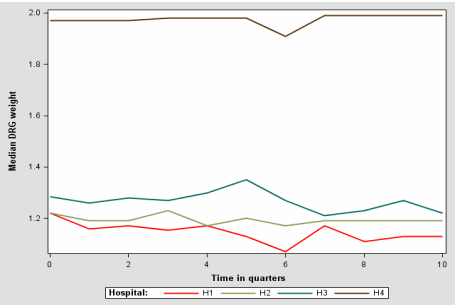
Based on the results of a study that compared the pre-PPS to the post-PPS period, it has been observed a decrease of the LOS by 24% (95% CI: (3.1, 3.8)) [12], which did not affect the quality of services though [13]. The PPS it also does not seem to affect [14] the rising quality of the in-hospital services. On the other hand, data of the period July 1985 to June 1986 (post-PPS period) indicates an increase of the number of patients which had been discharged at home in a non-stable condition [15]. A follow up study on data selected during the years 1983-1992 [16], indicates that people of age 65 or older have three times higher risk of re-admission, 2 months after their discharge. A different study, showed more specifically that, 180 days after the admission, the number of readmission does not differ for the periods 1981-1982 (pre-PPS) and 1985-1986 (post-PPS) (there was 1% decrease, p-value > 0.05). In contrast, the respective number, one year after the admission, had been significantly decreased (p-value< 0.05).

Chapter 3

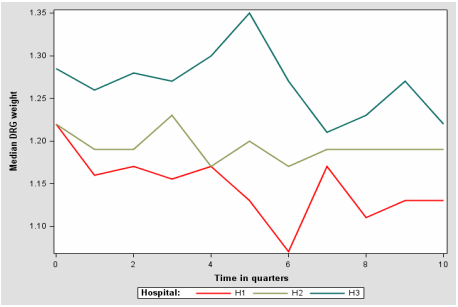
Analysis

The goal of this work is the investigation of the CMI changes of a network of hospitals in the state of Indiana, and the factors that could have possibly affect the CMI levels. The data contains information collected during the period 2011-2013.

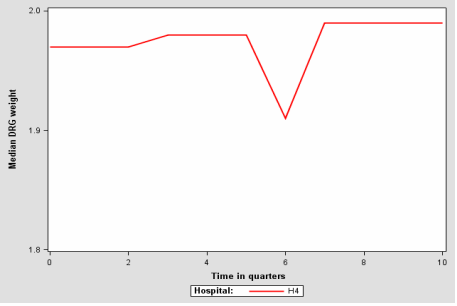
An overview of the data is given in Figure 3.1, which describes the levels of DRG-weights over time for each of the hospitals. Notice the big difference of the DRG-weight levels between the hospital H4 and the rest of the hospitals. That was expected since H4 is a hospital mainly specialized to a specific disease (90.3% of its cases) with high surgery needs (70.54% among the specific disease's cases). The rest three hospitals though, seem to be very similar in terms of their DRG-weight levels. The variation over time looks constant for all four hospitals. However, these are just observations based on the graph given below (Figure 3.1: subfigure 3.1.1).



Subfigure 3.1.1: Median DRG weight over time for all hospitals



Subfigure 3.1.2: Median DRG weight over time for H1, H2 and H3 hospitals



Subfigure 3.1.3: Median DRG weight over time for the H4 hospital

Figure 3.1: Median DRG weight over time for the hospitals.

By taking a closer look on the DRG-weight plots (Figure 3.1: subfigure 3.1.2, subfigure 3.1.3) it

seems that, the H4 hospital has a slightly rising DRG-weight over time while, the DRG-weight for the rest of the hospitals is declining.

Since the nature of the data requires the application of Linear Mixed Models (LMM), we have to check some of the assumptions that need to be met. Because the normality assumption of the the outcome variable is violated (Figure 3.2.1), we apply the Box-Cox transformation ($t(y) = (y^{\lambda} - 1)/\lambda; \lambda = -0.5$; see Figure 3.2.2). Thus, instead of modeling the DRG weight, we are modeling a function of this quantity, and as we will se latter, we use the inverse quadratic root to model the length of hospitalization (LOS). Using a random intercept and slope LMM, we get that the marginal mean of the transformed DRG weight for the H3 hospital is the second largest, following the H4 which is used as reference hospital for comparisons, followed by the H1 and H2 hospitals (See Table 3.1, Model M1).

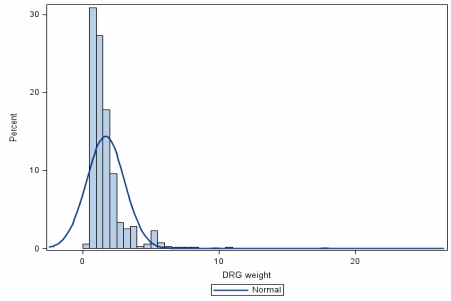


Figure 3.2.1: DRG weight

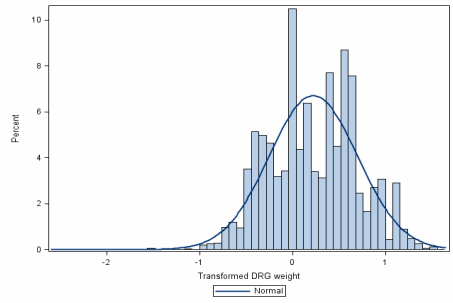


Figure 3.2.2: Box-Cox transformation of DRG weight

Table 3.1: Regression of the hospital factor on the transformed DRG weight, using LMM.

Model M1 [1]: $\mathrm{E}[g(DRG)_{ij}] = b_{0i} + \beta_0 + b_{1i}Hospital + + \beta_1 H1_{ij} + \beta_2 H2_{ij} + \beta_3 H3_{ij} + \beta_4 Time_{ij}$

Model M1				
Hospital	Estimate	Std.error	95%CI	p-value
Intercept	0.345	0.026	(0.293, 0.398)	< 0.0001
H1	-0.238	0.027	(-0.292, -0.185)	< 0.0001
H2	-0.259	0.026	(-0.311, -0.208)	< 0.0001
H3	-0.169	0.028	(-0.225,-0.113)	< 0.0001

From the same model (Model M1) we also get that, the levels of the DRG weight for each of the H1, H2, H3 hospitals are significantly different from these of the baseline category (i.e. the H4 hospital) (p-value< 0.0001 for all of the hospitals, see Table 3.1, Model M1). A statistically significant

[1] Box-Cox transformation: $g(y) = \frac{y^{\lambda}-1}{\lambda}$, where $\lambda = -0.5$

rate difference is also observed (interaction of the each of the three hospitals with the Time; See Table 3.2, Model M2) between each of the hospitals and the hospital H4. Specifically, the mean DRG weight for the general hospitals (H1, H2, H3) significantly differs, over time, from that of the H4 hospital (see Table 3.2, Model M2; interaction terms). These hospitals have a declining DRG weight over time (see Table 3.2, additional estimations) however, this reduction is not big enough to be considered as statistically significant (see Table 3.2; additional estimations: p-value>0.05 for all three hospitals).

Finally, the negative coefficient of the time variable expresses a statistically significant declining rate for the DRG weight of the H4 hospital (See Table 3.2, Model M2; p-value$<$ 0.0001). That is, among the cases in the H4 hospital, the mean DRG weight is declining over time and that decline is statistically significant (See Table 3.2, Model M2, time coefficient; See Figure 3.3).

Table 3.2: Regression of the hospital factor, the time variable and their interaction on the transformed DRG weight, using LMM.

Model M2: $\mathrm{E}[g(DRG)_{ij}] = b_{0i} + b_{1i}Hospital + \beta_0 + \beta_1 H1_{ij} + \beta_2 H2_{ij} + \\ + \beta_3 H3_{ij} + \beta_4 Time_{ij} + \beta_5 (TimeH1)_{ij} + \beta_6 (TimeH2)_{ij} + \beta_7 (TimeH3)_{ij}$

Model M2

Variable	Estimate	Std.error	95%CI	p-value
Intercept	0.3725	0.0276	(0.3182, 0.4268)	$<$ 0.0001
H1	-0.258	0.015	(-0.3139, -0.2021)	$<$ 0.0001
H2	-0.2855	0.0146	(-0.3398, -0.2312)	$<$ 0.0001
H3	-0.1963	0.0161	(-0.2551, -0.1374)	$<$ 0.0001
Time(Quarters)	-0.0058	0.00157	(-0.0089, -0.0028)	0.0002
Time*H1	0.00422	0.00189	(0.00049, 0.0079)	0.0264
Time*H2	0.00561	0.0018	(0.002, 0.0092)	0.0021
Time*H3	0.00575	0.002	(0.0016, 0.0098)	0.0057

Additional estimations

Varible	Estimate	Std.error	95%CI	p-value
$\beta 4 + \beta 5$	-0.00166	0.001	(-0.00376, 0.00044)	0.12
$\beta 4 + \beta 6$	-0.00027	0.0009	(-0.0021, 0.00155)	0.76
$\beta 4 + \beta 7$	-0.00013	0.0013	(-0.0028, 0.00254)	0.92

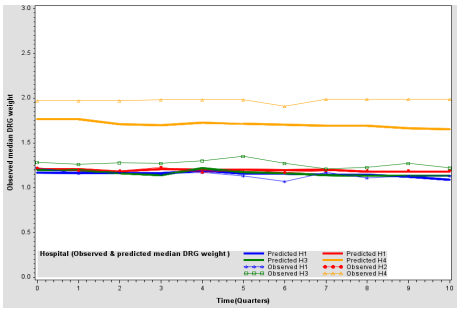


Figure 3.3: Observed vs. estimated DRG weight over time (Model M2)

After applying a linear mixed model to analyze the data that excludes H4 (See Table 3.3; Model M2.1; See Figure 3.4) we observe that, there is no significant reduction for the H3 hospital (H3 used as reference category) over time, in terms of its mean DRG weight (See Table 3.3, time coefficient; see Figure 3.4). Also, both H2 and H1 hospitals have significantly lower DRG weight levels (at the baseline) compared to the H3 hospital (See Table 3.3- Model M2.1; p-value=0.0006 and p-value=0.022 respectively) while, their DRG-weight rate is not significantly different (see Table 3.3, interaction terms are not statistically significant) from that of the H3 hospital.

Table 3.3: Regression of the hospital factor, the time variable and their interaction on the transformed DRG weight, using LMM (H4 excluded).

Model M2.1: $\mathrm{E}[g(DRG)_{ij}] = b_{0i} + b_{1i}Hospital + \beta_0 + \beta_1 H1_{ij} + \\ + \beta_2 H2_{ij} + \beta_3 Time_{ij} + \beta_4 (TimeH1)_{ij} + \beta_5 (TimeH2)_{ij}$

Model M2.1				
Variable	Estimate	Std.error	95%CI	p-value
Intercept	0.1692	0.018	(0.132,0.206)	< 0.0001
H1	-0.0542	0.0235	(-0.1, -0.008)	0.022
H2	-0.0795	0.0227	(-0.1244, -0.0346)	0.0006
Time(Quarters)	-0.00019	0.00133	(-0.0028, 0.0024)	0.889
Time*H1	-0.0015	0.00169	(-0.0048, 0.00186)	0.3743
Time*H2	-0.00011	0.0016	(-0.0032, 0.003)	0.9479

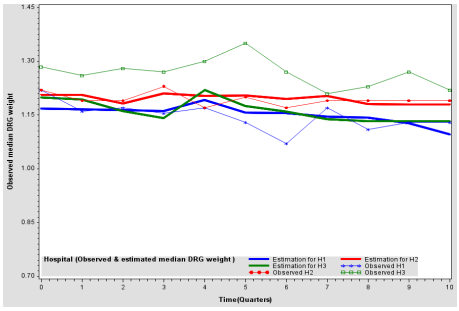


Figure 3.4: Observed vs. estimated DRG weight over time (Model M2.1; H4 excluded)

Below are presented some candidate factors that could have affected the trend of the hospitals' DRG weight over time:

1. **Change of the type of cases (i.e. change of the product lines)**

   Given that the different product lines have different DRG weights over time (Figure 3.5), a change of the type of cases administered in the hospitals could cause a change on the overall DRG weight. Based on the graph (Figure 3.5), although it looks that there are significant differences between the mean DRG weight of the product lines, it is not clear whether there

is significant change over time (i.e. whether there is an effect of time on the product lines). Using the product line called "Women and Children" as the reference category we get that, the "Spine" product line is the one with the highest DRG weight followed by the "Orthopedics" and the "Cardiovascular" (see Figure 3.5). All of the product lines have significantly higher DRG weight levels than the baseline (Women and Children product line) at time zero (See Table 3.4,-Model M.3; see Figure 3.5), which seems to remain stable over time (the estimated coefficient of the time variable has p-value=0.0825, see Table 3.4,-Model M3).

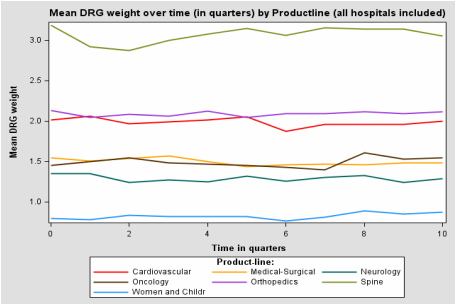


Figure 3.5: Mean DRG weight over time per product line

Table 3.4: Regression of the Product-line factor, the time variable and their interaction on the transformed DRG weight, using LMM.

Model M3: $E[g(DRG)_{ij}]) = b_{0i} + b_{1i}Hospital + \beta_0 + \beta_1 Cardio_{ij} + \\ + \beta_2 MedSur_{ij} + \beta_3 Neuro_{ij} + \beta_4 Oncology_{ij} + \beta_5 Ortho_{ij} + \beta_6 Spine_{ij} + \\ \beta_7 Time_{ij} + \beta_8 (Cardio * Time)_{ij} + \beta_9 (MedSur * Time)_{ij} + \beta_{10} (Neuro * Time)_{ij} + \\ + \beta_{11} (Oncology * Time)_{ij} + \beta_{12} (Ortho * Time)_{ij} + \beta_{13} (Spine * Time)_{ij}$

Model M3

Variable	Estimate	Std.error	95%CI	p-value
Intercept	-0.27	0.023	(-0.318, -0.226)	< 0.0001
Cardiovascular	0.479	0.023	(0.433, 0.526)	< 0.0001
Medical-Surgical	0.434	0.0229	(0.389, 0.478)	< 0.0001
Neurology	0.489	0.0255	(0.439, 0.538)	< 0.0001
Oncology	0.504	0.0279	(0.45, 0.55)	< 0.0001
Orthopedics	0.637	0.026	(0.586, 0.688)	< 0.0001
Spine	0.6906	0.029	(0.634, 0.747)	< 0.0001
Time(Quarters)	0.0052	0.003	(-0.00067, 0.0111)	0.0825
Cardiovascular*Time	-0.0095	0.0032	(-0.0158, -0.0032)	0.0032
Medical-Surgical*Time	-0.00714	0.0031	(-0.013, -0.001)	0.0219
Neurology*Time	-0.00616	0.0035	(-0.013, 0.0008)	0.0825
Oncology*Time	-0.00163	0.0042	(-0.009, 0.0065)	0.6973
Orthopedics*Time	-0.0032	0.0034	(-0.01, 0.0036)	0.359
Spine*Time	-0.00355	0.004	(-0.011, 0.004)	0.383

Additional estimations

Varible	Estimate	Std.error	95%CI	p-value
$\beta 7 + \beta 8$	-0.00427	0.0011	(-0.00647, -0.00207)	0.0001
$\beta 7 + \beta 9$	-0.00191	0.0008	(-0.00351, -0.00031)	0.019
$\beta 7 + \beta 10$	-0.00093	0.0018	(-0.0046, 0.0027)	0.06203
$\beta 7 + \beta 12$	0.00167	0.0027	(-0.0037, 0.007)	0.5414

Type 3 test of Fixed effects

Varible	X^2	p-value
Productline	782.78	< 0.0001
Time*Productline	19.96	0.0028

A significantly different rate of change, compared to Women and Children, is indicated for the "Cardiovascular" and the "Medical-Surgical" product lines, over time (See Table 3.4-Model M3, the estimated coefficient of the corresponding interaction effects have p-value=0.0032 and p-value=0.0219 respectively). Specifically (See Table 3.4, Model M3; additional estimations), the Cardiovascular product line of the network has a significant declining trend over time and that also holds for the Medical-Surgical product line, although the slope is not as steep as for the Cardiovascular.

Due to the high percentage of Cardiovascular cases[2] in the H4 hospital (H4 had 6,701 cases while, the combined number of cases of other hospitals was 7,628) and because it has the highest average Cardio-DRG weight, we analyze the product lines' DRG weight for the rest of the hospitals separately, excluding the cases belonging to the H4, in order to see how that affects the DRG weights change of the other hospitals (See Table 3.5-Model M3.1.1).
The Medical-Surgical DRG weight still shows a significantly different rate of change over time (Table 3.5- Model M3.1.1; p-value=0.0136) compared to the baseline rate of change whereas, there is a marginally significant difference of the rate of change for the Cardiovascular DRG weight, when it is compared to the baseline rate (See Table 3.5- Model M3.1.1; p-value=0.0726). Specifically, the DRG weight for Cardiovascular product-line, shows a slight reduction over time which is not statistically significant (See Table 3.5-Model M3.1.1; additional estimations: p-value=0.693).

That is, the significant estimation of the Cardiovascular product-line by time interaction, presented by the Model 3 (See Table 3.4), must be either due to a significant rate of change of the H4 DRG weight or due to an increased power of the Model 3 (compared to the M3.1.1 model where, the H4 data has been excluded). On the other hand the Medical-Surgical product line indicates a statistically significant reduction, in terms of its DRG weight, over time (See Table 3.5, Model M3.1.1; additional estimations (coefficient $\beta 7 + \beta 9$)).

Based on the results of the model M3.1.2 (See Table 3.6; see Figure 3.6), the Cardiovascular DRG weight of the H4 hospital, shows a statistically significant decline over time com-

[2] Also, the H4 hospital has the highest Medical-Surgical DRG weight, although the highest percentage of Medical-Surgical cases is obtained for the H2 hospital.

Table 3.5: Regression of the Product-line factor, the time variable and their interaction on the transformed DRG weight, using LMM (H4 excluded).

Model M3.1.1: $\mathrm{E}[g(DRG)_{ij}] = b_{0i} + b_{1i}Hospital + \beta_0 + \beta_1 Cardio_{ij} + \beta_2 MedSur_{ij} + \\ + \beta_3 Neuro_{ij} + \beta_4 Oncology_{ij} + \beta_5 Ortho_{ij} + \beta_6 Spine_{ij} + \beta_7 Time_{ij} + \\ + \beta_8 (Cardio * Time)_{ij} + \beta_9 (MedSur * Time)_{ij} + \beta_{10} (Neuro * Time)_{ij} + \\ + \beta_{11} (Oncology * Time)_{ij} + \beta_{12} (Ortho * Time)_{ij} + \beta_{13} (Spine * Time)_{ij}$

Model M3.1.1

Variable	Estimate	Std.error	95%CI	p-value
Intercept	-0.28	0.02	(-0.32, -0.24)	< 0.0001
Cardiovascular	0.427	0.024	(0.38, 0.47)	< 0.0001
Medical-Surgical	0.4325	0.022	(0.38, 0.47)	< 0.0001
Neurology	0.487	0.025	(0.43, 0.53)	< 0.0001
Oncology	0.511	0.027	(0.45, 0.56)	< 0.0001
Orthopedics	0.631	0.0256	(0.58, 0.68)	< 0.0001
Spine	0.6935	0.028	(0.63, 0.75)	< 0.0001
Time(Quarters)	0.0052	0.0029	(-0.0007, 0.011)	0.077
Cardiovascular*Time	-0.0058	0.00331	(-0.012, 0.0005)	0.0726
Medical-Surgical*Time	-0.0075	0.003	(-0.013, -0.005)	0.0136
Neurology*Time	-0.0065	0.00347	(-0.013, 0.0003)	0.0604
Oncology*Time	-0.0025	0.004	(-0.01, 0.005)	0.5449
Orthopedics*Time	-0.0032	0.0034	(-0.009, 0.003)	0.3457
Spine*Time	-0.0037	0.004	(-0.011, 0.0042)	0.365

Additional estimations

Varible	Estimate	Std.error	95%CI	p-value
$\beta 7 + \beta 8$	-0.0006	0.0015	(-0.00356, 0.00237)	0.693
$\beta 7 + \beta 9$	-0.00214	0.0008	(-0.00388, -0.00073)	0.0075

Type 3 test of Fixed effects

Varible	X^2	p-value
Productline	803.94	< 0.0001
Time*Productline	13.59	0.0346

pared to the rest of the hospitals (see Table 3.6-Model M3.1.2; additional estimations; p-value< 0.0001), while a significant increase of the Medical-surgical DRG weight is observed for the H4 (See Table 3.7-Model M3.1.3; additional estimation, and Figure 3.7). Furthermore, one can notice that there is an opposite slope indicated by the observed and predicted Cardiovascular-DRG weight for the general hospitals (see Figure 3.6). That difference can be explained by the fact that the observed median presented in the Figure 3.6 doesn't take into account the correlation between the observations but in contrary, the predictions take the correlation into account by introducing the random effects in the mixed model.

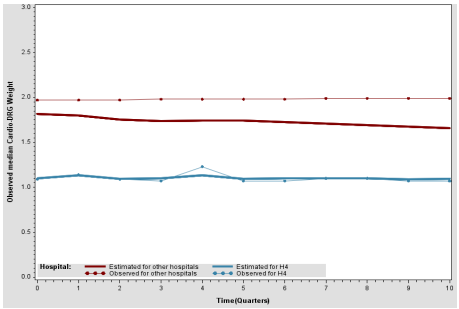


Figure 3.6: Observed vs. estimated Cardio-DRG weight of the hospital H4 vs. other hospitals (Model M3.1.2)

Table 3.6: Regression of the time variable on the transformed Cardiovascular-DRG weight, using LMM.

Model M3.1.2: $\mathrm{E}[g(DRG)_{ij}] = b_{0i} + b_{1i}Hospital + \beta_0 + \\ + \beta_1 Cardio_{ij} + \beta_2 Time_{ij} + \beta_3 (CardioTime)_{ij} + \beta_4 H4_{ij} + \\ + \beta_5 (TimeH4)_{ij} + \beta_6 (CardioH4)_{ij} + \beta_7 (CardioTimeH4)_{ij}$

Model M3.1.2				
Variable	Estimate	Std.error	95%CI	p-value
Intercept	0.132	0.0136	(0.106, 0.159)	< 0.0001
Cardio	-0.019	0.0104	(-0.0403, 0.00048)	0.0556
Time(Quarters)	-0.00086	0.0006	(-0.002, 0.00044)	0.197
Cardio*Time	0.00078	0.0016	(-0.002, 0.004)	0.638
H4	0.019	0.036	(-0.0504, 0.0904)	0.578
Time*H4	0.0102	0.005	(0.0003, 0.02)	0.0427
Cardio*H4	0.267	0.029	(0.2095,0.3252)	< 0.0001
Cardio*Time*H4	-0.0177	0.0055	(-0.028,-0.007)	0.0012

Additional estimations				
Varible	Estimate	Std.error	95%CI	p-value
$\beta 2 + \beta 3 + \beta 5 + \beta 7$	-0.00758	0.0016	(-0.0108, -0.00436)	< 0.0001

Summarizing the observations made up to this point, it might be the decline of the Cardiovascular product line of the H4 hospital that significantly affects the overall (H1, H2, H3 and H4) DRG weight. That could be justified by the fact that, the significantly higher Cardiovascular DRG weight (compared to the rest of the hospitals -see Table 3.6 Model 3.1.2; Cardio*H4 coefficient) of the H4 hospital, shows a significant decline over time (see Table 3.6, Model M3.1.2; additional estimation). In contrary, the Medical-Surgical DRG weight of the H4 is significantly lower compared to the other hospitals (see Table 3.7, Model 3.1.3; Medical-Surgical*H4 coefficient).

Table 3.7: Regression of the time variable on the transformed Medical-Surgical-DRG weight, using LMM (H4 excluded)

Model 3.1.3: $\mathrm{E}[g(DRG)_{ij}] = b_{0i} + b_{1i}Hospital + \beta_0 + \beta_1 MedSur_{ij} + \\ + \beta_2 Time_{ij} + \beta_3 (MedSurTime)_{ij} + \beta_4 H4_{ij} + \beta_5 (TimeH4)_{ij} + \\ + \beta_6 (MedSurH4)_{ij} + \beta_7 (MedSurTimeH4)_{ij}$

Model M3.1.3				
Variable	Estimate	Std.error	95%CI	p-value
Intercept	0.15	0.014	(0.12, 0.17)	< 0.0001
Medical-Surgical	-0.036	0.0074	(-0.051, -0.022)	< 0.0001
Time(Quarters)	0.0008	0.0008	(-0.0009, 0.0025)	0.36
Medical-Surgical*Time	-0.0029	0.0011	(-0.005, -0.00068)	0.011
H4	0.27	0.026	(0.22,0.32)	< 0.0001
Time*H4	-0.0081	0.0018	(-0.0117, -0.00451)	< 0.0001
Medical-Surgical *H4	-0.24	0.0317	(-0.303,-0.18)	< 0.0001
Medical-Surgical *Time*H4	0.0216	0.006	(0.009,0.0333)	0.0003

Additional estimations				
Varible	Estimate	Std.error	95%CI	p-value
$\beta 2 + \beta 3 + \beta 5 + \beta 7$	0.0114	0.0056	(0.000315, 0.02247)	0.0438

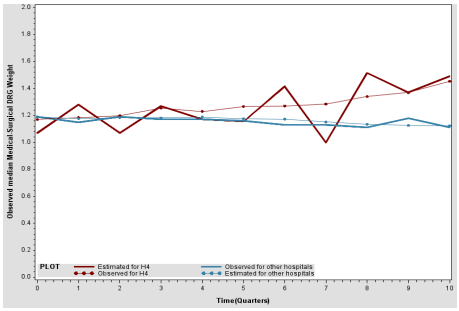


Figure 3.7: Observed vs. estimated Med-DRG weight of the hospital H4 vs. other hospitals (Model M3.1.3)

2. **Change in the number of Surgical cases**

   Another hypothesis that we need to study when trying to explain the hospitals' CMI change (and as a result, DRG weight change), is that of the different DRG weights between the Surgical and Medical type of cases.

   From the Figure 3.8 we observe a big difference between the DRG weight of these two types of cases. Indeed, by applying a LMM for the overall mean DRG weight (i.e. the mean DRG weight without distinguishing among hospitals) we observe that the mean DRG weight of the Medical cases is significantly lower compared to the Surgical cases (See Table 3.8-Model M4; p-value$<$ 0.0001). Also, the rate of change over time is significantly different for the two DRG types (see Table 3.8, Model M4, interaction term with p-value=0.035).

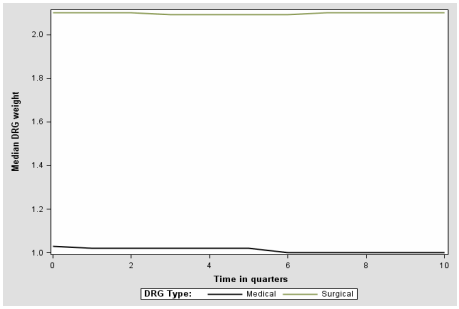


Figure 3.8: Comparison of Medical vs. Surgical cases' median DRG weight

Taking a closer look on the change of the surgical cases within each of the hospitals (see Figure 3.9), looks like all hospitals except from the H1, retain a steady percentage of Surgical type of DRG over time. Compared to the H4, which has the highest percentage of surgical

Table 3.8: Regression of the DRG type (Surgical vs. Medical) on the transformed DRG weight, for all hospitals, using LMM.

Model 4: $\mathrm{E}[g(DRG)_{ij}] = b_{0i} + b_{1i}Hospital + \beta_0 + \\ + \beta_1 Medical_{ij} + \beta_2 Time_{ij} + \beta_3 (MedicalTime)_{ij}$

Model M4				
Variable	Estimate	Std.error	95%CI	p-value
Intercept	0.6	0.012	(0.57, 0.62)	< 0.0001
Medical	-0.73	0.006	(-0.745, -0.721)	< 0.0001
Time (Quarters)	0.00024	0.0007	(-0.0012, 0.0017)	0.75
Medical*Time	-0.00196	0.0009	(-0.0038, -0.00014)	0.035
Additional estimations				
Varible	Estimate	Std.error	95%CI	p-value
$\beta 2 + \beta 3$	-0.00172	0.0058	(-0.00287, -0.00057)	0.0034

DRG type among its cases (that is, 68.25% while the H3 and H2 have 35.33% and 33.12% respectively), the hospitals H1, H2 and H3 have significantly lower percentage of Surgical cases at the baseline (See Table 3.9, Model M.5, p-value$<$ 0.0001 for the H1, H2 and H3). In addition, the odds of Surgical type of DRG for the H1 hospital differ significantly (see Table 3.9; interaction term time by H1, p-value=0.0419) from these of the H4 hospital (specifically there is a decline- see Table 3.9, Model M5; additional estimations) while, its percentage of surgical cases is the lowest compared to the other hospitals and differs significantly form that of the H4 (see Table 3.9).

Overall, there have not been observed significant changes, in terms of the DRG type (i.e. Medical or Surgical), that could have affected the mean DRG weight of the network, over time (only the H1 hospital shows a change but this hospital does not have significant changes of its DRG weight).

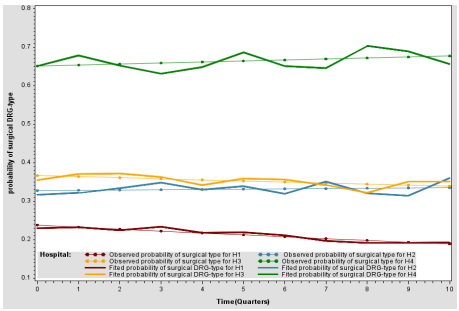


Figure 3.9: Observed vs. fitted probabilities of surgical type of cases, for each hospital

Table 3.9: Logistic regression (GLMM) of the DRG type (Surgical vs. Medical) over the time and hospital variables. Independent correlation matrix.

Model M5: $log[\frac{Pr[g(DRGtype)ij=1]}{Pr[g(DRGtype)ij=0]}] = \beta_0 + \beta_1 H1_{ij} +$
$+\beta_2 H2_{ij} + +\beta_3 (H3)_{ij} + \beta_4 Time_{ij} +$
$+\beta_5 (TimeH1)_{ij} + \beta_6 (TimeH2)_{ij} + \beta_7 (TimeH3)_{ij}$

Model M5				
Variable	Estimate	Std.error	95%CI	p-value
Intercept	0.622	0.1517	(0.325, 0.9196)	< 0.0001
H1	-1.785	0.2773	(-2.33, -1.24)	< 0.0001
H2	-1.342	0.264	(-1.86, -0.82)	< 0.0001
H3	-1.169	0.2534	(-1.66, -0.67)	< 0.0001
Time (Quarters)	0.0116	0.0128	(-0.013, 0.036)	0.362
Time*H1	-0.041	0.02	(-0.08, -0.0015)	0.0419
Time*H2	-0.0081	0.0194	(-0.046, 0.03)	0.674
Time*H3	-0.0237	0.0342	(-0.091, 0.043)	0.4897
Additional estimations				
Variable	Estimate	Std.error	95%CI	p-value
$\beta 4 + \beta 5$	-0.0294	0.015	(-0.059,0.000)	0.0504
$\beta 4 + \beta 6$	-0.0035	0.0144	(-0.032, 0.025)	0.8085
$\beta 4 + \beta 7$	-0.012	0.0319	(-0.074, 0.051)	0.7058

3. **Change of the LOS**

Another factor that could potentially affect the CMI of a hospital is the change in the LOS. An increase on the number of short-stay cases (defined as $\leq$2 days hospitalization) in the hospital can be a reason for the CMI's decline. Based on the given data, when the length of hospitalization is increased by one day, the mean DRG weight is increased as well and, this effect is statistically significant (see Table 3.10, model M.6, p-value$<$ 0.0001).

Table 3.10: Regression of the LOS(in days) on the transformed DRG weight for all hospitals, using LMM

Model M6: $E[g(DRG)_{ij}] = b_{0i} + b_{1i} Hospital + \beta_0 + \beta_1 LOS_{ij}$

Model M6				
Variable	Estimate	Std.error	95%CI	p-value
Intercept	-0.0168	0.013	(-0.042, 0.009)	0.202
LOS (Days)	0.0394	0.00037	(0.038,0.04)	< 0.0001

The LOS depends significantly on the time period we refer to (over time, the LOS is in general increased; See Table 3.11-Model M7, p-value$<$ 0.0001, see Figure 3.10), the hospital (See Table 3.12-Model M8; see Figure 3.12) and the specific product-line [3] (see Figure 3.11 and Table 3.15).

[3] The lowest is LOS for Women and Children and the highest LOS is for oncology; mean=3.47 days, median=2 days and mean=5.13 days, median=4 days respectively

Table 3.11: Regression of the Time variable over the transformed LOS (in days),using LMM.

Model M7[4]: $E[g(LOS)_{ij}] = b_{0i} + b_{1i}Hospital + \beta_0 + \beta_1 Time_{ij}$

Model M7				
Variable	Estimate	Std.error	95%CI	p-value
Intercept	0.856	0.011	(0.835, 0.878)	< 0.0001
Time (Quarters)	0.0078	0.00074	(0.006, 0.009)	< 0.0001

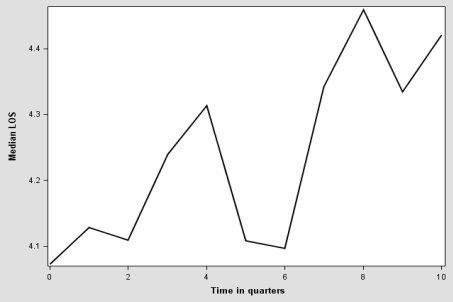


Figure 3.10: Overall LOS (in days). All hospitals included.

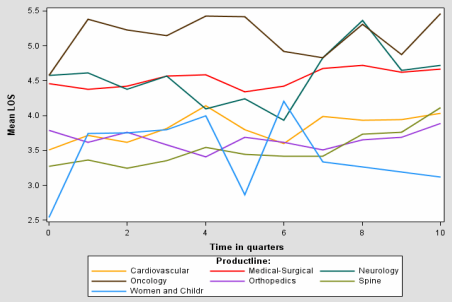


Figure 3.11: Overall LOS (in days). All hospitals included.

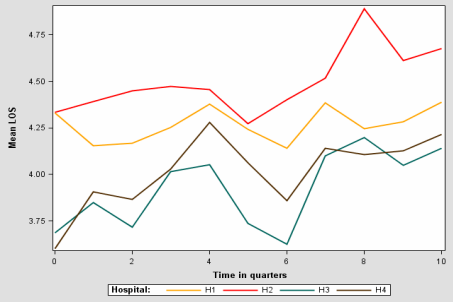


Figure 3.12: Median LOS (in days) by hospital

The hospitals H3 and H4 have the lowest mean LOS (3.91 and 3.99 days respectively while, the hospital H2 has the highest mean LOS: 4.49 days). The mean LOS is not significantly different among the H3 and H4 hospitals (see Table 3.12-Model M8, p-value=0.075). For the rest of the hospitals (i.e. the H1 and H2), the LOS is significantly higher at the baseline (see Table 3.12-Model M8, both of the p-values< 0.05).

Over time, all of the hospitals indicate a statistically very significant increase of the LOS (see Table 3.14, Model M10). In terms of CMI (and thus DRG weight), we would expect that this implies an increase of the hospitals' DRG weight as well. Indeed, for the short stay cases

[4] Box-Cox transformation: $g(y) = \frac{y^{\lambda}-1}{\lambda}$, where $\lambda = -0.25$

(defined as hospitalization for at most two days), the mean DRG weight is significantly lower compared to cases who needed longer hospitalization (see Table 3.13-Model M9).

Table 3.12: Regression of the Hospital and Time variable on the transformed LOS (in days), using LMM.
Model M8: $\mathrm{E}[g(DRG)_{ij}] = b_{0i} + b_{1i}Hospital + \beta_0 + \beta_1 H1_{ij} + \\ + \beta_2 H2 + \beta_3 H4 + \beta 4 Time_{ij}$

Model M8				
Variable	Estimate	Std.error	95%CI	p-value
Intercept	0.83	0.0146	(0.8, 0.86)	< 0.0001
H1	0.046	0.022	(0.033,0.09)	0.034
H2	0.049	0.02	(0.0089, 0.09)	0.017
H4	-0.0535	0.029	(-0.1125, 0.0055)	0.075
Time (Quarters)	0.0077	0.00074	(0.006, 0.009)	< 0.0001

Type 3 test of Fixed effects		
Varible	X^2	p-value
Hospitals	5.34	0.0001

Table 3.13: Regression of the short stay cases (i.e LOS($\leq$ 2days)) on the transformed DRG weight, using LMM.
Model M9: $\mathrm{E}[g(DRG)_{ij}] = b_{0i} + b_{1i}Hospital + \beta_0 + \beta_1 LOS_{ij}$

Model M9				
Variable	Estimate	Std.error	95%CI	p-value
Intercept	0.246	0.0127	(0.22, 0.271)	< 0.0001
LOS($\leq$ 2days)	-0.245	0.0033	(-0.252, -0.239)	< 0.0001

Table 3.14: Regression of the Hospital and Time variable on the transformed LOS (in days), using LMM (interaction term included).
Model M10: $\mathrm{E}[g(LOS_{ij})] = b_{0i} + b_{1i}Hospital + \beta_0 + \beta_1 H1_{ij} + \beta_2 H2_{ij} + \\ + \beta_3 H3_{ij} + \beta_4 Time_{ij} + \beta_5 (TimeH1)_{ij} + \beta_6 (TimeH2)_{ij} + \beta_7 (TimeH3)_{ij}$

Model M10				
Variable	Estimate	Std.error	95%CI	p-value
Intercept	0.76	0.028	(0.7, 0.817)	< 0.0001
H1	0.128	0.032	(0.066, 0.19)	< 0.0001
H2	0.124	0.0307	(0.064, 0.185)	< 0.0001
H3	0.0707	0.032	(0.007, 0.134)	< 0.0001
Time (Quarters)	0.012	0.002	(0.0082, 0.0161)	< 0.0001
Time*H1	-0.006	0.0024	(0.011, -0.014)	0.011
Time*H2	-0.0047	0.0023	(-0.009, -0.00018)	0.042
Time*H3	-0.00423	0.0026	(-0.009, 0.0014)	0.15

Additional estimations				
Varible	Estimate	Std.error	95%CI	p-value
$\beta 4 + \beta 5$	0.00596	0.0013	(0.0032, 0.00866)	< 0.0001
$\beta 4 + \beta 6$	0.0074	0.00118	(0.005, 0.0097)	< 0.0001
$\beta 4 + \beta 7$	0.00831	0.0017	(0.00492, 0.0117)	< 0.0001

Type 3 test of Fixed effects		
Varible	X^2	p-value
Hospitals	20.76	0.0001
Time*Hospitals	0.00831	0.0848

Studying the change of the LOS for the product lines (see Table 3.15-Model M11), we see that the Orthopedics as well as the Neurology have a significantly different rate of LOS change, compared to the reference product line (p-values=0.003 and 0.038 respectively) while, no significant difference on the LOS rate of change is observed for the rest of the product lines compared to the baseline category. The DRG weight of the product lines over time has been estimated, in order to check wether it is observed a decline of the DRG weight for these two product lines (the Neurology and the Orthopedics). However, it seems that these two product lines retain a steady level of DRG weight over time in contrast to the DRG weight of the Cardiovascular and Medical-Surgical product lines which is declined over time (See Table 3.4, model M3). So, it looks that the reduction of the LOS for the Neurology and Orthopedics has not affected the DRG weight of the hospitals. Finally, it has not been mentioned any significant increase of the number of the short-stay cases (See Table 3.16, model M12).

Table 3.15: Regression of the Product-line and Time variables on the transformed LOS, using LMM .

Model M11: $\mathrm{E}[g(LOS_{ij})] = b_{0i} + b_{1i}Hospital + \beta_0 + \beta_1 Cardio_{ij} + \beta_2 MedSur_{ij} + \\ + \beta_3 Neuro_{ij} + \beta_4 Oncology_{ij} + \beta_5 Ortho_{ij} + \beta_6 Spine_{ij} + \beta_7 Time_{ij} + \\ + \beta_8 (Cardio * Time)_{ij} + \beta_9 (MedSur * Time)_{ij} + \beta_{10} (Neuro * Time)_{ij} + \\ + \beta_{11} (Oncology * Time)_{ij} + \beta_{12} (Ortho * Time)_{ij} + \beta_{13} (Spine * Time)_{ij}$

Model M11

Variable	Estimate	Std.error	95%CI	p-value
Intercept	0.689	0.027	(0.634, 0.743)	< 0.0001
Cardiovascular	0.101	0.03	(0.044, 0.16)	0.0005
Medical-Surgical	0.195	0.029	(0.14, 0.25)	< 0.0001
Neurology	0.2	0.031	(0.139, 0.263)	< 0.0001
Oncology	0.19	0.035	(0.13, 0.26)	< 0.0001
Orthopedics	0.344	0.032	(0.28, 0.41)	< 0.0001
Spine	0.138	0.036	(0.067, 0.209)	0.0005
Time(Quarters)	0.0135	0.0038	(0.006, 0.021)	0.0005
Cardiovascular*Time	-0.002	0.004	(-0.001, 0.006)	0.618
Medical-Surgical*Time	-0.0062	0.0039	(-0.014, 0.0016)	0.1119
Neurology*Time	-0.0094	0.0045	(-0.018, -0.0005)	0.038
Oncology*Time	-0.0018	0.0054	(-0.012, 0.008)	0.74
Orthopedics*Time	-0.0133	0.0044	(-0.02,-0.004)	0.003
Spine*Time	-0.004	0.0052	(-0.014, 0.006)	0.432

Additional estimations

Varible	Estimate	Std.error	95%CI	p-value
$\beta 7 + \beta 8$	0.0115	0.0014	(0.00865, 0.01434)	< 0.0001
$\beta 7 + \beta 9$	0.0073	0.0011	(0.00526, 0.00938)	< 0.0001

Type 3 test of Fixed effects

Varible	X^2	p-value
Productline	216.14	< 0.0001
Time*Productline	23.97	0.0005

Table 3.16: Logistic regression of the LOS (dichotomous)[5] on time and hospitals (interaction term included), using GLMM. Correlation structure AR(1).

Model M12[6] : $log[\frac{Pr[g(LOS)ij=1]}{Pr[g(LOS)ij=0]}] = \beta_0 + \beta_1 H1_{ij} + \beta_2 H2_{ij} +$
$+\beta_3 H3_{ij} + \beta_4 Time_{ij} + \beta_5 (TimeH1)_{ij} + \beta_6 (TimeH2)_{ij} + \beta_7 (TimeH3)_{ij}$

Model M12

Variable	Estimate	Std.error	95%CI	p-value
Intercept	-0.55	0.11	(-0.765, -0.3412)	< 0.0001
H2	-0.0227	0.0128	(-0.275, 0.229)	0.86
H3	0.1628	0.148	(-0.128, 0.454)	0.273
H4	0.549	0.15	(0.241, 0.858)	0.0005
Time (Quarters)	-0.0143	0.009	(-0.0315, 0.0029)	0.1032
Time*H2	-0.011	0.012	(-0.0345, 0.0128)	0.368
Time*H3	-0.0035	0.017	(-0.0383, 0.0312)	0.841
Time*H4	-0.022	0.012	(-0.046, 0.0018)	0.069

Additional estimations

Variable	Estimate	Std.error	95%CI	p-value
$\beta 4 + \beta 5$	-0.025	0.0082	(-0.0414, -0.0089)	0.0025
$\beta 4 + \beta 6$	-0.0178	0.0154	(-0.048, 0.0124)	0.247
$\beta 4 + \beta 7$	-0.0367	0.0086	(-0.0536, -0.0198)	< 0.0001

Score statistics for Type 3 test (GEE Analysis)

Varible	X^2	p-value
Hospitals	12.18	0.0068
Time*Hospitals	3.55	0.3143

Chapter 4

Discussion

DRG is a case-classification system developed in 1970 in Yale. The goal of this system is to group together medically similar cases, which estimated to require, on average, similar hospital resources. The higher the DRG the higher the CMI and thus, the higher the hospital's reimbursements.

The changes in the CMI could be the result of several factors. For this thesis, three factors have been checked for their relatedness to the DRG weights (and as a result to the CMI). These three factors are the changes in the number of the several productlines (7 product lines in total), the surgical cases and the length of stay.

The data was collected for each quartile through the years 2011-2013, from 4 different hospitals in the state of Indiana. The results were obtained by applying linear mixed models to account for the correlation between the observations.

Overall, a DRG reduction of the Cardiovascular and Medical-Surgical product lines was obtained for the data combining all four hospitals. However, it seems that the decline in the Cardiovascular DRG could be the result of the significant decline of that product line in the H4 hospital specifically (which has the larger number of Cardiovascular cases compered to the rest of the hospitals). No significant change of the DRG over time was observed for the rest of the productlines. After studying the changes of the number of the Surgical cases and the length of stay (we were interested in the decline of the LOS and in the increase of the short-stay cases), no significant results were found.

As a conclusion, the three hospitals H1, H2, H3 show a decline in the DRG over time, which is not statistically significant. The only significant DRG decline is for the H4 hospital.The overall CV DRG also shows a decline, probably due to the the DRG decline of CV productline observed in the H4 hospital. There was no obvious effect of the Medical/Surgial volume, nor of the LOS. Further study of the CV-productline in the H4 hospital is needed by specialists. In addition, changes in the DRG that have been detected by the specific analysis could also be the result of other factors that affect the CMI/DRG, like the changes in the coding system, the documentation etc. Unfortunately these factors cannot be tested based on the given data.

Bibliography

[1] http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1_047260.hcsp?dDocName=bok1_047260

[2] B Steinwald and L A Dummit, "*Hospital case-mix change: sicker patients or DRG creep?*", Health Affairs, 8, no.2 (1989):35-47, doi: 10.1377/hlthaff.8.2.35

[3] http://californiahia.org/sites/californiahia.org/files/docs/CDQarticles/2013-04-current-drg-grouping-algorithms.pdf

[4] Mirella Cacace, Achim Schmid, " *The role of diagnosis related groups (DRGs) in health-care system convergence*", BMC Health Services Research 2009, 9(Suppl 1): A5, doi: 10.1186/1472-6963-9-S1-A5

[5] Inke Mathauera, Friedrich Wittenbecherb, "*Hospital payment systems based on diagnosis-related groups: experiences in low- and middle-income countries*", Bull World Health Organization 2013; 91:746-756A

[6] Thomas B. Bradley, Gerald F. Kominsk, "*Contributions of case mix and intensity change to hospital cost increases*", Health Care Financing Review, Winter 1992, volume 14, Number 2

[7] http://www.jblearning.com/samples/0763750344/45561_ch01.pdf

[8] http://multimedia.3m.com/mws/mediawebserver?mwsId=SSSSSuH8gc7nZxtUPxtvox_eevUqe17zHvTSevTSeSSSSSS--

[9] David C. Hsia, Cathaleen A. Ahern, et al., *Medicare Reimbursement Accuracy Under the Prospective Payment System, 1985 to 1988*, Jama, August 19, 1992-Vol 268, No.7

[10] David C. Hsia, W. Mark Krushat, et al., *Accuracy of diagnostic coding for Medicare patients under the Prospective-Payment System*, The New England Journal of Medicine, Feb. 11, 1988-Vol. 318, No.6

[11] Grace M. Carter, Joseph P. Newhouse, Daniel A. Relles,*How much change in the case mix index is DRG creep?*, J Health Econ,1990; 9(4): 411-28.

[12] Katherine L. Kahn, Emmet B. Keeler, et al., *Comparing outcomes of care before and after implementation of the DRG-based Prospective Payment System*, Reprinted from JAMA, The Journal of the American Medical Association, October 17, 1990, Vol. 264, No. 15, pp.1984-1988.

[13] Norbert Goldfield, *The evolution of diagnosis-related groups(DRGs): from its beginnings in Case-Mix and resource use theory, to its implementation for payment and now for its current utilization for quality within outside the hospital*, Q Manage Health Care, 2010, Vol. 19, No.1, pp.3-16

[14] William H. Rogers, David Draper, the al., *Quality of care before and after implementation of the DRG- based prospective payment system*, JAMA, October 17, 1990, Vol. 264, No. 15, pp. 1989-1994.

[15] Jacqueline Kosecoff, Katherine L. Kahn, et al., *Prospective payment system and impairment at discharge*, Reprinted from JAMA, The Journal of the American Medical Association, October 17, 1990, Vol. 264, No. 15, pp. 1980-1983.

[16] Qian Xufeng, Russell Luise B., et al., *Medicare's Prospective Payment System for Hospitals: new evidence on transitions among health care settings*, Working papers, Department of economics, Rutgers, the State University of New Jersey, No. 2007, 07.

Curriculum Vitae

Danai Chasioti

Education

Indiana University-Purdue University Indianapolis, IN, USA

Ph.D in Bioinformatics 2015-

Indiana University-Purdue University Indianapolis, IN, USA

MS in Biostatistics 2011-2017

Dissertation: A Longitudinal analysis for the identification of the factors that affect the Case Mix Index of hospitals in the U.S.

National & Kapodistrian University of Athens, Athens, Greece

MSc in Biostatistics & Epidemipology 2008-2011

Dissertation: Review of statistical methods in Genome wide Association Studies

University of Aegean, Samos, Greece

BSc Statistics & Actuarial Science 2001-2007

Dissertation: Game theory & Nash equilibrium

Professional experience

Biostatistician in network of local hospitals (Indianapolis, 08/2014 - 06/2015)

Statistical consultant for the Greek food company "Papadopoulou" (2009)

Publications

1. Chasioti D., Yao X., Zhang P., Ning X., Li L., Shen L. (2017) Mining directional drug interaction effects on myopathy using the FAERS database. PSB'17: Pac Symp Biocomput., poster #71, Big Island of Hawaii, January 3-7, 2017 (Poster)
2. Chasioti D., Yao X., Zhang P., Quinney S., Ning X., Li L., Shen L. (2017) Mining and visualizing the network of directional drug interaction effects. NetSci'17: Int. School and Conf. on Network Science, Indianapolis, IN, June 21-23, 2017 (Poster)

3. Kefei L. et al., Transcriptome-Guided Imaging Genetic Analysis via a Novel Sparse CCA Algorithm, Graphs in Biomedical Image Analysis, Computational Anatomy and Imaging Genetics, Cardoso, J. et al., Springer, 2017, pp.220-229 (Book Chapter)

4. Chasioti, D., Yao, X. et al., (2017). Mining directional drug interaction effects on myopathy using the FAERS database. CTSI Symp. on Disease and Therapeutic Response Modeling and Simulation, Indianapolis, USA. (Oral presentation)

5. Chasioti, D. (2011). Review of statistical methods in Genome wide Association Studies. MSc Thesis, National & Kapodistrian University of Athens, Greece.

6. Chasioti, D. (2007). Game theory & Nash equilibrium. BSc Thesis, University of the Aegean, Greece.

Teaching

1. Research assistant, Department of Mathematics IUPUI (2011-2014)

2. Tutoring in Statistics, Technological Educational institute of Athens (2011)

3. Tutoring in Statistics, National and Kapodistrian university of Athens (2011)

4. Teaching assistant, Biostatistics I laboratory, Harokopion University of Athens (2009)