2012
2013

FACULTY OF SCIENCES

Master of Statistics: Biostatistics

Masterproef

After 5 years of action: has patient safety culture improved in the Belgian hospitals? Results of a benchmark database

Promotor :
Mevrouw Liesbeth BRUCKERS

Quinta Nanga Momuluh

Master Thesis nominated to obtain the degree of Master of Statistics , specialization Biostatistics

Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

universiteit hasselt
KNOWLEDGE IN ACTION

Universiteit Hasselt | Campus Hasselt | Martelarenlaan 42 | BE-3500 Hasselt
Universiteit Hasselt | Campus Diepenbeek | Agoralaan Gebouw D | BE-3590 Diepenbeek

universiteit hasselt | Maastricht University

2012
2013

FACULTY OF SCIENCES

Master of Statistics: Biostatistics

Masterproef

After 5 years of action: has patient safety culture improved in the Belgian hospitals? Results of a benchmark database

Promotor :
Mevrouw Liesbeth BRUCKERS

Quinta Nanga Momuluh

Master Thesis nominated to obtain the degree of Master of Statistics , specialization Biostatistics

universiteit hasselt | Maastricht University

CERTIFICATION

This is to certify that this project was carried out by Momuluh, Quinta Nanga under the supervision of;

Mevrouw BRUCKERS Liesbeth	Dr. SCHROOTEN Ward
……	…….
Supervisor	External Supervisor

Momuluh, Quinta Nanga

…….

Student

ACKNOWLEDGEMENTS

Special thanks to God Almighty for His omnipresence in every situation.

Much gratitude to all my professors, group members and non-teaching staffs for their time and guardians in making the learning process a success.

I am especially thankful to my supervisor Mevrouw BRUCKERS Liesbeth for the time and help she put in to realize this project.

Finally much thanks to my family for being there in the good and bad times. I will remain grateful to you all.

LIST OF TABLES

Table of Contents

ABSTRACT

Introduction: *Reason (1998) highlights that safety culture 'is a concept whose time has come', stating that there is both a challenge and an opportunity to develop a clear theoretical understanding of organizational issues to create a principled basis for more effective culture enhancing practices. An organization requires safety culture as a product of combined effects of organizational culture, professional culture and often national culture. These organizations include hospitals where patient safety culture is a major area of concern.*

Objective: *This study seeks to know if there are differences in safety culture between Belgian hospitals. It also seeks to know if there were significant variations in safety culture over time (5 years) and also to investigate possible covariates that could have a significant effect on dimensional scores.12 Dimensional scores including 2 outcome dimensions were used as tools in investigating safety culture.*

Methodology: *The data set of interest consisted of 111 hospitals which had both first and second measurements. Of these 111 hospitals, there were 69 acute hospitals, 34 psychiatric hospitals and 8 long term care hospitals. These hospitals contributed a total 86,199 respondents (observations). Means, proportions and missingness were explored to have an overall picture of the study. Linear mixed models which captured important aspects of the data were fitted in an attempt to achieve the study objectives. Different assumptions regarding missingness were considered to see if missingness had a significant effect. Complete case and multiple imputation analysis were considered for this purpose.*

Results: *Similar results were obtained under the different assumptions of missingness. There was a low level of correlation between the dimensional scores. The mean dimensional score was >3 on a scale of 5 for all the dimensions. There were variations in safety dimensions within and between the hospitals. The intra cluster correlation indicated a small level of correlation between respondents in the same hospital. Meanwhile a likelihood ratio test confirms the need for the random effect.*

Conclusions: *Results indicates a difference in patient safety culture after 5 years. There was within and between variability with a corresponding intra cluster correlation for the hospitals. Highest correlations were observed within the long term hospitals while the acute hospitals had the least between variability. The number of significant covariates was dependent on the dimension in question. For all the dimensions, covariates patient safety grade (E1), Number of events reported (G1), staff position in the hospital (H4) and language spoken in the hospital (Taal) had a significant effect on the dimensional scores.*

1 INTRODUCTION

Safety culture could be defined as the way safety is perceived, valued and prioritized in an organization or institution. It reflects real commitment to safety at all levels in the organization. Safety culture is what people believe about importance of safety. If someone therefore believes that safety is not really important, unsafe decisions and judgments could be the result. Gledon et al, (2006) highlights that when defining safety culture, the premises of some researchers is to focus on attitude; while others emphasize safety culture being expressed through their behavior and work activities. Clarke (2006) states that safety culture is not only observed within the 'general state of the premises and working conditions but in the attitudes of employees towards safety'.

Patient safety is a crucial point of health care quality. Achieving safety culture requires an understanding of values, beliefs and norms of what attitudes and behaviors related to patient safety (Gledon et al, 2006). Improving the culture of safety within health care is an essential component of providing overall health care quality. In the quest to improve the overall health care quality of her patients, the Belgian government has put much effort in the baseline assessment of safety culture within acute, psychiatric and long term care hospitals. Acute hospitals are hospitals intended for short term medical and surgical treatment and care (treatment of patients with emergency needs). On the other hand, a long term care hospital provides services to meet both medical and non-medical needs of people with chronic illnesses or disabilities. Such patients can not care for themselves over a long period of time. Also known as mental hospitals, psychiatric hospitals specializes in short term or out-patient therapy for low risk patients. Others may offer permanent care to patients who require assistance and treatment in specialized and controlled environment. Irrespective of the type of hospital, the task of providing care to the patients is tedious and delicate as even the slightest errors can put the patient's safety at risk. The Belgian federal program therefore promotes safety culture as a key component to improve patient safety (Vlayen A, et al 2011). This was a five year program between 2007 and 2012.

Safety culture is generally measured by surveys. Validated surveys include Agency for Health Research and Quality (AHRQ) safety culture survey. This was also used for the Belgian safety culture program. The survey asks respondents to rate the safety culture in their unit and in the hospital as a whole. This was done with regards to 12 dimensions including 2 outcome dimensions. The dimensions are considered as key features which will aid in identifying those areas of needs to improve patient safety.

Previous studies based on 90 acute hospitals in the Belgian safety culture survey found language, work area and profession to have significant effect on patient safety. This study also indicated that patient safety cultures were low. Hand offs and transitions, staffing and management support for patient safety were identified as major problem dimensions which were required to be an organizational wide priority (Vlayen A. et al, 2011)

Objective

After five years of patient safety culture sensitization and awareness, the objectives of this study were the following;

-To conduct exploratory analysis based on the benchmark database from the hospital survey patient safety culture (HSPSC).

-To examine variation in patient safety within and across hospitals.

-To investigate if there were significant variations in patient safety culture over time.

-Finally, interest was also to examine possible covariates.

2 THE DATA

The benchmark database for the patient safety dimension contained 176 hospitals with a total of 115,764 observations. In this study, emphasis was only based on those hospitals which had records for the first and second measurement. Hence the dataset of interest had 111 hospitals. Of these 111 hospitals, there were 69 acute hospitals, 34 psychiatric hospitals and 8 long term care hospitals. These hospitals contributed a total 86,199 respondents (observations). The database included 12 safety dimensions, two of which were outcome dimensions. These are shown on Table 1 below.

Table 1: Safety Dimensions and their Description.

Safety Dimensions	**Description**
D1score	Supervisor/manager expectations and actions promoting safety
D2score	Organizational learning–continuous improvement
D3score	Teamwork within units
D4score	Communication openness
D5score	Feedback and communication about error
D6score	Nonpunitive response to error
D7score	Staffing
D8score	Management support for patient safety
D9score	Teamwork across units
D10score	Handoffs and transitions
	Outcome dimensions
O1score	Overall perceptions of patient safety
O2score	Frequency of events reported

The dataset contained the following variables some of which were considered as possible covariates in investigating patient safety.

Table 2: Variables in Database and their Description.

Covariates	**Description**
A0	Respondent's primary work area or unit in the hospital
E1	Patient safety grade
G1	Number of events reported
H1	Duration of service in hospital
H2	Duration of service in work area/unit
H3	Number of hours typically worked per week
H4	Staff position in the hospital
H5	Staff position with respect to direct interaction or contact with patients
H6	Duration of service in current specialty or profession
Other Variables	
BenchmarkID	Unique identity code for each respondent
Year	Year of measurement
Measurement	First and second measurement
Hospcode	Anonymous and unique hospital code
Type	Type of hospital; AZ (acute hospitals), PZ(psychiatric hospitals), SP(long-term care hospitals)
Taal	Language; N=Dutch speaking ; F=French speaking; B=both Ductch and French

3 METODOLOGY

3.1.0 Exploratory Data Analysis

In other to get to know the data; means of the dimensional scores were computed. Given that the needs and hence the care pertaining to the safety of patients in the different type of hospitals (acute, psychiatric and long term care) are different, the means for these hospitals were computed and compared to the overall mean. Box plots were also used to check possible trends in the submission of the dimensional scores over the years and to see the level of participation of the hospitals based on the language they speak. Some missingness was also encountered in this study and was explored by computing the frequencies and proportions of missing observations for the different dimensions and variables. PROC MI in SAS 9.2 was also used to explore the missingness pattern in this dataset.

3.1.1 Exploring Associations between Dimensions

Pairwise Pearson correlation was used to check if possible associations existed between the dimensional scores. A situation which would warrant the need reduce the dimensions since the given pair would have the same meaning. This was done under the following null and alternative hypothesis.

H_o :$r = 0$ (no correlation).

H_a :$r \neq 0$ (correlation different from zero)

The test statistics for this hypothesis follow a t-distribution as follows;

$$t = r\sqrt{\frac{n-2}{1-r^2}} \quad \text{.....(1)}$$

Where: r^2 is the coefficient of determination which measures the strength of linear relationship between any two variables (dimensions).

n is the sample size and

r is the correlation coefficient which follows a t-distribution with n-2 degrees of freedom.

3.1.2 Missing Data

Missing data is a common occurrence in the area of research. Data may be missing for variety of reasons. If missingness is related to the sensitivity of questions in a survey for instance, such missingness is informative. In such situations, it may therefore not be wise to exclude those with missing data from the analysis as this might lead to biased results.

Rubin (1976) defines three main missing data mechanisms; missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). An observation is said to be MCAR if the missingness is independent of all observed and unobserved data. Observations can also be MCAR if missingness only depends on values of fixed covariates (Covariate dependent dropout).

When observations are MAR, missingness is dependent on the observed values. A valid analysis that ignores the missing value mechanism can be obtained using this assumption (MAR). This is termed 'ignorable' by Rubin (1976) and Little and Rubin (2002). MCAR is a special case of MAR and are both referred to as 'ignorable' mechanisms. Also, methods based on MAR are valid if data are MCAR or MAR. On the contrary, MCAR are only valid if data are MCAR (Mallinckrodt et al, 2003).

Finally when missingness depends not only on the observed data but also on the unobserved (missing) data, this is termed MNAR (Little and Rubin, 1987). MNAR mechanism is often reffered to as non-ignorable missingness because the missing data mechanism cannot be ignored when the goal is to make inference about the distribution of the complete survey analysis. Most standard methods become invalid when data are MNAR. In order to make valid estimators, joint models for response vectors and missing data mechanisms are required. The three mechanisms therefore differ in terms of assumptions concerning missingness with respect to whether missingness is related to observe or unobserved responses (Fritzmaurice et al, 2004).

When there is missingness, it might be challenging to attribute the missingness to any of the three mechanisms (MCAR, MAR and MNAR) since missing information is not observed. Hence it is not possible to know how the complete sample looks like. Molenberghs and Verbeke (2005) advice that sensitivity analysis would be the most sensible action to perform whenever assumptions are made about missing data mechanisms.

3.2.1 Complete Case (CC) Analysis

When the missing data mechanism is MCAR; all methods that yield valid inference in the absence of missing data will also yield valid inference if the analysis is based on all available data. This is also true when the analysis is restricted to subjects with complete observations. When analysis is restricted to respondents with complete observations, this is known as complete case analysis. The analysis is achieved by deleting all subjects with incomplete responses. This method is computationally simple but leads to substantial loss of information. Consistent estimates can therefore be obtained if only the missingness process is MCAR.

3.2.2 Multiple Imputation (MI) Analysis

When missing data are MAR but not MCAR, complete case analysis and other methods based on all available observations yield biased estimates of mean responses. A valid analysis that ignores the missing value mechanism can be obtained given the MAR assumption. Under the MAR assumption, multiple imputation (MI) is a possible approach to account for sampling variability and model uncertainty. This is possible by replacing each missing value by more than one imputed value before analysis is conducted. Under this approach, complete data sets (generated under m imputations) are separately analyzed and subsequently pooled for final inference via the MIANALYZE procedure (Molenberghs and Verbeke, 2005).

The MCMC (Markov Chain Monte Carlo) method was used with 5 imputations. The imputation model was done using the entire dataset imputing all outcome dimensions as well as covariates with missing information. These variables included ; A0, E1, G1, H1, H3, H4, H6, H5, D1score, D2score, D3score, D4score, D5score, D6score, D7score, D8score, D9score, D10score, o1score, o2score. (See description of variables in the variable section). D1score to o2score are the outcome variables. The covariates contained both continues and categorical variable. The categorical variables included A0, E1, H4 and H5. Each categorical variable in the dataset was replaced by a corresponding set of binary dummy variable with 0-1 value given the number of levels in the variable and then treated as an individual normal variable in the imputation step. The levels of the categorical variables were 14, 5, 12 and 2 for the A0, E1, H4 and H5 variables respectively. The imputed values were rounded off using a ROUND=1 with the Proc MI statement. A '1' is then assigned to which ever levels dummy with the highest imputed value and zero to the other levels. This leads to an imputed data that conforms to the nature of the actual data. The assumptions about the probability model underlying the imputed data included multivariate normality for the variables. The 5 imputed data sets were then pooled together via MIANALYZE procedure to obtained inference of the model parameters. This procedure works by first extracting the point estimates and estimated standard errors form the 5 separate analysis. These are then combined to arrive at a single point estimate, its estimated standard error and significant test for the model parameters. The MI and MIANALYZE procedures also assume that the parameters θ of the data model and parameters φ of the model for the missing data indicators are distinct. ie, knowing the value of θ does not provide additional information about φ and vice versa.

3.3.0 Linear Mixed Models (LMM)

Given that respondents within the same hospital are likely to respond to questions of this survey in a similar manner a good choice for a model will be a random effect model (linear mixed model). Linear mixed model involves the incorporation of subject specific effects in a model with a continuous response, giving a hierarchical interpretation to the model. The general linear mixed model is given by:

$$\boldsymbol{Y_i}=\boldsymbol{X_i\beta} + \boldsymbol{Z_i b_i} + \boldsymbol{\varepsilon_i}\ ,\ i=1,\ldots,m \qquad (2)$$

where m is the number of hospitals.

$\boldsymbol{b_i}$~N($\mathbf{0}$,D), $\varepsilon_i \sim N(0,\Sigma_i)$, Y_i~N($\boldsymbol{X_i\beta}$, Ω_i)

Where $E(\boldsymbol{Y}_i) = \boldsymbol{X_i\beta}$, $\text{cov}(\boldsymbol{Y_i}) = \boldsymbol{Z_i}D\boldsymbol{Z}_i' + \sigma^2\mathbf{I}_{ni} = \boldsymbol{Z}_i D\boldsymbol{Z}_i' + \Sigma_i = \Omega_i$ and $\mathbf{b}_1,\ldots,\mathbf{b}_m$, $\boldsymbol{\varepsilon}_1,\ldots,\boldsymbol{\varepsilon}_m$ are independent.

The vector $\boldsymbol{Y}_i = (Y_{i1}, Y_{i2},\ldots Y_{ini})'$ is an n_i-dimensional vector of all repeated measurements for the ith hospital (subject). The matrix $\boldsymbol{Z}_i$ is the within subject design matrix (matrix for the random effects in subject i) of order $n_i\times q$. The vector $\boldsymbol{\beta}$ is a p×1 vector of fixed population parameters. $\boldsymbol{b}_i$ is the q×1 vector of random effects for subject i. $\boldsymbol{\varepsilon}_i$ is the n×1vector of errors for observations in subject i. $\boldsymbol{D}$= q×q covariance matrix of the random effects.

$\Sigma_i = n_i \times n_i$ covariance matrix for the errors in hospital *i*.

The fitted model considered for this study was as follows:

$$\text{Dimension}_{ij} = b_i + \beta_1 E1(1)_i + \beta_2 E1(2)_i + \beta_3 E1(3)_i + \beta_4 E1(4)_i + \beta_5 G1_i + \beta_6 H1_i + \beta_7 H2_i + \beta_8 H3_i + \beta_9 H4(1)_i + \beta_{10} H4(2)_i + \beta_{11} H4(3)_i + \beta_{12} H4(4)_i + \beta_{13} H4(5)_i + \beta_{14} H4(6)_i + \beta_{15} H4(7)_i + \beta_{16} H4(8)_i + \beta_{17} H4(9)_i + \beta_{18} H4(10)_i + \beta_{19} H4(11)_i + \beta_{20} H5(1)_i + \beta_{21} H6_i + \beta_{22} \text{Type(AZ)}_i + \beta_{23} \text{Type(PZ)}_i + \beta_{24} \text{Taal(B)}_i + \beta_{25} \text{Taal(F)}_i + \beta_{26} \text{measurement(1)}_i + \beta_{27} \text{measurement*Type(AZ)}_i + \beta_{28} \text{measurement*Type(PZ)}_i + \varepsilon_{ij} \quad \text{.....(3).}$$

Where measurement represent the first and second time points with measurement(1) indicating the first time point.

The restricted maximum likelihood (REML) method was used as an estimation method over the Maximum likelihood (ML) method since the later uses a likelihood function calculated from a transformed set of data so that the nuisance parameter has no effect. Complete case (CC) analysis and MI were used for the analysis via LMM. These Analysis were done using PROC MIXED in SAS 9.2

The intra-cluster correlation indicates the portion of the total variance which occurs between the hospitals. These correlations were computed as follows;

$$\rho = \frac{\tau^2}{\sigma^2 + \tau^2} \quad \text{.....(4)}$$

Where τ^2 and σ^2 are the variance components for the random intercept (random hospital effect) and error respectively. τ^2 is a measure of the unexplained random differences between respondents in different hospital. As explained in the exploratory analysis different random effect variances were assigned to the different hospital types since they differ in terms of care and hence safety towards patients. σ^2 (residual variability) on the other hand is a measure of the remaining variability when all other sources of variability has been accounted for. It includes measurement error and model specification errors.

4 RESULTS:

4.1.0 Exploratory Data Analysis.

Table 3 below shows the means and standard errors of the dimensional scores. The highest mean value was 3.6921 (D3score) for the PZ hospital. Average dimensional scores for the PZ hospitals were higher for almost all the dimensions as compared to the mean value of the AZ and SP hospitals. The AZ hospitals had the lowest mean values. On an average, the mean dimensional scores were >3 and were measured on a scale of 5. D7score had the lowest mean value in the AZ hospitals. Overall, the highest mean value was 3.6182 for the D3score and the lowest was 3.04585 for the D7score. The distribution of the mean dimensional scores were skewed; but based on the central limit theorem (CLT), sample means of moderately large samples are often well approximated by normal distribution even if the data are not normally distributed.

Table 3: Mean, Standard Error of Dimensional Scores

Dimension	Mean (μ_{AZ}) TYPE AZ (S.E)	Mean (μ_{PZ}) TYPE PZ (S.E)	Mean (μ_{SP}) TYPE SP(S.E)	Overall Mean (μ) (S.E)
D1score	3.5335(0.6771)	3.6886(0.6106)	3.5886(0.6384)	3.5561(0.0112)
D2score	3.4673(0.5989)	3.5312(0.5540)	3.5682(0.5677)	3.4779(0.0118)
D3score	3.6051(0.6827)	3.6921(0.6295)	3.6815(0.6497)	3.6182(0.0154)
D4score	3.5343(0.7402)	3.5808(0.6949)	3.4891(0.6953)	3.5393(0.0090)
D5score	3.3327(0.8517)	3.4068(0.7727)	3.3665(0.8366)	3.3432(0.0118)
D6score	3.1295(0.7434)	3.2851(0.7218)	3.2753(0.7290)	3.1534(0.0127)
D7score	3.0057(0.7156)	3.2484(0.6363)	3.2404(0.6644)	3.0485(0.0159)
D8score	3.1399(0.7022)	3.2919(0.6722)	3.4308(0.5686)	3.1666(0.0237)
D9score	3.1546(0.5735)	3.3095(0.5749)	3.3803(0.5786)	3.1801(0.0119)
D10score	3.0395(0.6413)	3.1142(0.6354)	3.2360(0.6401)	3.0539(0.0111)
O1score	3.2770(0.6564)	3.3750(0.6180)	3.4107(0.5546)	3.2931(0.0158)
O2score	3.1748(0.9807)	3.2724(0.9017)	3.1879(0.9246)	3.1880(0.0130)

Box plots of dimensional scores and the year covariate investigating trends across the years during which measurements were carried out are shown in appendix (figure A.1a, A.1b, A.1c). Submissions of dimensional scores in the acute hospitals were done from 2005 to 2011 excluding 2010. Meanwhile these submissions only started in 2007 for the long term care and psychiatric hospitals; the PZ hospitals submitted dimensional scores from 2007 to 2011 excluding 2009 and the SP hospitals submitted in 2007, 2008 and 2011. For all the hospitals, most of the median scores were highest in 2011. These scores for most of the dimensions were lower during the first year of measurement, staying slightly constant across the other years and increasing in the last year of measurement (2011) To get insight about the distribution of the different hospitals (AZ, PZ, SP) with respect to the language spoken, box plots of the distribution of the dimensions by language were created. For the acute hospitals, all the languages (Dutch, French and hospitals where both French and Dutch are spoken) were represented. In the long term care and psychiatric hospitals, only Dutch and French speaking hospitals were represented. Dutch speaking hospitals had a higher response rate than the

French speaking hospitals. The plots obtained were similar for all the dimensions. As a result only plots of dimension d1score have been presented in appendix A.2 for illustration. The results presented above only give an idea of what to expect when the dimensional scores are analyzed taking into account other factors such as covariate effect, missing observations and clustering resulting from similarities between respondents within the same hospital.

4.1.1 Exploring Missingness:

Results of frequencies and proportions of missing observations are presented in Table 4 below. A lot of missingness is observed for the AZ hospitals. The a0 variable which measures respondents work area records the highest level of missing observation for this hospital with a proportion of up to 9.79%. The level of missingness is also higher amongst the covariates containing respondent's background information as compared to the dimensional scores. Even though the dimensional scores also indicate some level of missingness, dimension 02score had the highest proportion of missing observations (6.30%) for the long term care hospitals. The highest proportion of missing observation in the psychiatric hospitals was observed for the a0 (21.15%)

Table 4: Frequency and Proportions for Missing Observations.

		TYPE=AZ
Dimensional Score	Frequency of missing observations	Proportion of missing observations
D1score	1015	1.39
D2score	190	0.26
D3score	110	0.15
D4score	620	0.85
D5score	600	0.822
D6score	994	1.36
D7score	104	0.14
D8score	743	1.02
D9score	681	0.93
D10score	900	1.23
o1score	1019	1.40
O2score	2380	3.26
Covariates		
A0	7143	9.79
E1	3402	4.66
G1	3127	4.29
H1	2307	3.16
H2	2371	3.25
H3	2662	3.26
H4	5598	7.67
H5	4068	5.58
H6	3997	5.48

		TYPE=SP
Dmensional Score	Freq of missing obs	Prop of missing obs
D1score	30	1.50
D2score	6	0.30
D3score	2	0.10
D4score	18	0.90
D5score	14	0.70
D6score	8	0.40
D7score	2	0.10
D8score	20	1.00
D9score	21	1.05
D10score	26	1.30
o1score	2	0.10
O2score	126	6.30
Covariates		
A0	100	5.01
E1	102	5.11
G1	92	4.61
H1	18	0.90
H2	24	1.20
H3	26	1.30
H4	79	3.96
H5	31	1.55
H6	39	1.95

		TYPE=PZ
Dmensional Score	Freq of missing obs	Prop of missing obs
D1score	105	0.93
D2score	11	0.10
D3score	6	0.05
D4score	65	0.58
D5score	70	0.62
D6score	13	0.12
D7score	4	0.04
D8score	62	0.55
D9score	86	0.76
D10score	69	0.61
o1score	7	0.06
O2score	507	4.51
Covariates		
A0	2380	21.15
E1	625	5.56
G1	411	3.65
H1	142	1.26
H2	160	1.42
H3	185	1.64
H4	1657	14.73
H5	205	1.82
H6	234	2.08

4.1.2 Missing Data Pattern

When there are missing data, another important step is to understand not just how much data are missing, but also the pattern of missing value. The pattern of missing values can sometimes suggest why the values are missing (Patrick E., et al 2007). Investigation of the missing data patterns (not shown) reveals the following results for the different hospital types. Lots of missing observations was observed for a0 and h1, h2, h3, h4, h5 and h6 variables. Specifically, the acute hospitals had 919 missing patterns. Of these 919 groups, 72.58% (52951) complete profiles were observed for the first group (group=1). Ie 72.58% of the respondents had no missing information. The last 5 groups of the pattern (915-919) had no covariate information. The long term care hospitals had 103 missing data patterns with 79.06% (1578) complete profiles. In this hospital, lots of missingness was seen in the a0 variable. Meanwhile the psychiatric hospitals recorded 275 missing data patterns with 67.63% (7609) complete profiles. The last 5 groups of this pattern had no information for all the covariates including the dimensional scores o1score and o2score.

A lot of missing values for the a0 and H variables may have been due to sensitivity of the survey questions. These variables either sort to know the work area or back ground information of the respondents. Answers to the questions related to these variables could easily be linked to the respondents who were rather reluctant to complete specific units. Hence the researchers suggest that the reason for missingness was because respondents wanted to stay anonymous.

4.1.3 Relationship between Dimensional Scores.

Correlations and covariances (equation 1) between the dimensional scores were computed to check if there were associations between them. Table 5 below shows the pair wise correlation of the dimensional scores. Based on the hypothesis test in section **3.1.1**, a significant p value of 0.0001 for all the dimensions indicates that there is correlation between the dimensional scores. This may be expected since all the dimensional scores are geared towards measuring patient safety in the hospitals. Most of these correlations were< 0.5 but for fairly higher values indicating possible association between d1score and d4score in the acute hospital (0.51). In other words supervisor/manager expectations and actions promoting safety may be explaining communication openness. D4score is again associated with the D5score dimension with a correlation of 0.58. Like the previous correlation, this suggests that communication openness may have the same meaning as feedback and communication about error. A rather moderate correlation of 0.54 and 0.57 was also observed for the d4score and d5score dimension in long term care and psychiatric hospitals respectively. D9score and D10score also have a correlation of 0.57 for all hospital types. This again indicates a possible association between dimensions measuring teamwork across units and handoffs and transitions in the hospitals. The rest of the pair wise correlations for the dimensional scores were always low indicating the dimensions are related but the relationships are not fixed (uncertain).

A scatter plot matrix for the first 5 dimensions for the different hospitals types is shown in the appendix (figure A.3a, A.3b and A.3c) to illustrate the relationship between the dimensional scores. The plots show that associations between dimensional scores were strongest in the acute hospitals and weakest in the long term care hospitals. This means that dimensional scores in the long term care hospitals were more independent.

Table 5: Pair wise Correlation of Dimensional Scores (lower triangle) and Covariances (diagonal +upper triangle)

						TYPE AZ						
Dimension	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	O1	O2
D1score	**0.46**	0.17	0.16	0.26	0.27	0.18	0.12	0.17	0.11	0.11	0.18	0.16
D2score	0.41	**0.36**	0.15	0.16	0.22	0.11	0.06	0.14	0.09	0.07	0.13	0.15
D3score	0.34	0.36	**0.47**	0.21	0.19	0.16	0.11	0.11	0.12	0.09	0.12	0.11
D4score	0.51	0.36	0.41	**0.56**	0.36	0.23	0.13	0.16	0.13	0.12	0.17	0.20
D5score	0.46	0.44	0.34	0.58	**0.73**	0.18	0.09	0.19	0.14	0.13	0.19	0.34
D6score	0.35	0.24	0.32	0.43	0.28	**0.55**	0.17	0.14	0.12	0.12	0.17	0.11
D7score	0.24	0.15	0.22	0.24	0.16	0.33	**0.51**	0.13	0.09	0.11	0.19	0.04
D8score	0.35	0.32	0.22	0.31	0.33	0.28	0.27	**0.49**	0.17	0.14	0.19	0.13
D9score	0.29	0.25	0.30	0.31	0.29	0.27	0.21	0.42	**0.32**	0.21	0.133	0.09
D10score	0.26	0.19	0.23	0.26	0.24	0.24	0.21	0.31	0.57	**0.41**	0.14	0.11
O1score	0.41	0.32	0.28	0.35	0.33	0.35	0.39	0.04	0.34	0.32	**0.34**	0.14
O2score	0.24	0.25	0.17	0.28	0.40	0.16	0.05	0.19	0.17	0.17	0.22	**0.96**

						TYPE SP						
Dimension	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	O1	O2
D1score	**0.41**	0.14	0.14	0.21	0.25	0.17	0.13	0.19	0.13	0.14	0.13	0.13
D2score	0.38	**0.32**	0.13	0.14	0.23	0.08	0.07	0.16	0.11	0.10	0.09	0.13
D3score	0.35	0.35	**0.42**	0.16	0.17	0.11	0.14	0.13	0.15	0.12	0.09	0.10
D4score	0.49	0.36	0.36	**0.48**	0.31	0.21	0.12	0.18	0.15	0.14	0.14	0.17
D5score	0.47	0.49	0.32	0.54	**0.70**	0.14	0.13	0.26	0.18	0.18	0.14	0.30
D6score	0.36	0.19	0.25	0.42	0.23	**0.53**	0.14	0.15	0.09	0.10	0.15	0.11
D7score	0.30	0.19	0.31	0.25	0.23	0.28	**0.44**	0.12	0.11	0.14	0.13	0.04
D8score	0.46	0.42	0.29	0.39	0.46	0.31	0.27	**0.43**	0.18	0.17	0.15	0.17
D9score	0.35	0.35	0.39	0.37	0.36	0.22	0.29	0.48	**0.33**	0.21	0.11	0.12
D10score	0.34	0.28	0.26	0.32	0.34	0.21	0.32	0.39	0.57	**0.41**	0.11	0.16
O1score	0.38	0.28	0.23	0.35	0.29	0.37	0.35	0.40	0.33	0.30	**0.31**	0.10
O2score	0.22	0.24	0.17	0.26	0.39	0.17	0.07	0.28	0.23	0.27	0.19	**0.85**

						TYPE PZ						
Dimension	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	O1	O2
D1score	**0.37**	0.14	0.14	0.20	0.22	0.16	0.09	0.15	0.11	0.11	0.16	0.14
D2score	0.42	**0.31**	0.13	0.14	0.19	0.11	0.05	0.14	0.09	0.08	0.12	0.13
D3score	0.36	0.37	**0.39**	0.18	0.18	0.17	0.08	0.11	0.12	0.10	0.12	0.11
D4score	0.48	0.35	0.42	**0.48**	0.31	0.24	0.10	0.15	0.14	0.14	0.16	0.18
D5score	0.46	0.44	0.36	0.57	**0.59**	0.18	0.07	0.19	0.15	0.14	0.17	0.28
D6score	0.36	0.26	0.34	0.47	0.32	**0.52**	0.14	0.15	0.13	0.14	0.17	0.12
D7score	0.23	0.13	0.19	0.24	0.15	0.30	**0.40**	0.10	0.08	0.09	0.13	0.03
D8score	0.36	0.36	0.25	0.33	0.36	0.30	0.24	**0.45**	0.17	0.15	0.17	0.11
D9score	0.33	0.29	0.33	0.34	0.33	0.31	0.21	0.44	**0.33**	0.21	0.13	0.10
D10score	0.29	0.22	0.26	0.31	0.28	0.29	0.25	0.36	0.57	**0.40**	0.14	0.10
O1score	0.42	0.34	0.31	0.38	0.35	0.39	0.32	0.41	0.36	0.36	**0.38**	0.12
O2score	0.25	0.25	0.19	0.29	0.41	0.19	0.04	0.18	0.19	0.17	0.22	**0.81**

4.2.1 Complete Case (CC) Analysis.

Results of the CC are shown on table 6 below. Significant covariates are indicated with an (*). The model considered is as shown below.

$$\text{Dimension}_{ij} = b_i + \beta_1 E1(1)_i + \beta_2 E1(2)_i + \beta_3 E1(3)_i + \beta_4 E1(4)_i + \beta_5 G1_i + \beta_6 H1_i + \beta_7 H2_i + \beta_8 H3_i + \beta_9 H4(1)_i + \beta_{10} H4(2)_i + \beta_{11} H4(3)_i + \beta_{12} H4(4)_i + \beta_{13} H4(5)_i + \beta_{14} H4(6)_i + \beta_{15} H4(7)_i + \beta_{16} H4(8)_i + \beta_{17} H4(9)_i + \beta_{18} H4(10)_i + \beta_{19} H4(11)_i + \beta_{20} H5(1)_i \beta_{21} H6_i + \beta_{22} \text{Type(AZ)}_i + \beta_{23} \text{Type(PZ)}_i + \beta_{24} \text{Taal(B)}_i + \beta_{25} \text{Taal(F)}_i + \beta_{26} \text{measurement(1)}_i + \beta_{27} \text{measurement*Type(AZ)}_i + \beta_{28} \text{measurement*Type(PZ)}_i + \varepsilon_{ij.}$$

Where Dimension_{ij} is the response referring observation *j* in hospital *i* for the different 12 dimensions. While E1 G1, H1, H2, H3, H4 H5, H6, Type, measurement, measurement* Type are the fixed effects (covariates). $\beta_1 \ldots\ldots\ldots \beta_{28}$ are the fixed effect coefficients. b_i is the random effect coefficient for the hospitals (random intercept). ε_{ij} is the error of observation *j* in hospital *i*. Description of the covariates is found on table xxx and details of the model are described in the methodology section. For this analysis (CC), the outcome dimensions (o1score and o2score) where considered to have an over view of the modeling.

O1score measures patient safety via overall perception of patient safety. For this dimension, patient safety grade (E1), number of events reported (G1), number of hours typically worked per week (H3), staff position in the hospital (H4), staff position with respect to direct interaction or contact with patient (H5), duration of service in current specialty or profession (H6), Type of hospital (Type), language spoken in the hospital (Taal) measurement time (measurement) had a significant effect on this response. A significant effect of type implies a significant difference in the safety dimension for different hospital type. Similarly, a significant effect of Taal means a significant difference in the dimension for the different languages spoken. Meanwhile a significant effect of measurement implies a significant difference in the dimension over time (first and second measurement periods).

Considering the o2score dimension which measures patient safety via frequency of events reported, E1, G1, duration of service in work area /unit (H2), H4, H5, H6, Type and Taal had a significant effect on the dimension.

Based on the variance components, there is a variation in the patient safety within and between hospitals. The intra class correlation for these outcome dimension ranges from 0.0009 for the acute (AZ) hospitals to 0.0206 in the psychiatric (PZ) hospitals. This indicates that there is low clustering with the respondents within the hospitals

Tabel 6: Parameter Estimate and Standard Error- Complete Cases (CC) Analysis.

Effect	Parameter	O1score Parameter estimates (s.e)	O2score Parameter estimates (s.e)
Intercept	β_0	2.4486(0.0311)*	2.5618(0.0692)*
E1(1)	β_1	1.8213(0.0249)*	1.1696(0.0444)*
E1(2)	β_2	1.4195(0.0206)*	0.8284(0.0364)*
E1(3)	β_3	0.9224(0.0205)*	0.4392(0.0363)*
E1(4)	β_4	0.2975(0.0219)*	0.0376(0.0387)
G1	β_5	-0.0083(0.0005)*	0.0273(0.0009)*
H1	β_6	-0.0003(0.0042)	0.0007(0.0006)
H2	β_7	-0.0005(0.0004)	-0.0028(0.0007)*
H3	β_8	-0.0006(0.0002)*	0.0004(0.0003)
H4(1)	β_9	-0.0390(0.0093)*	0.0785(0.0166)*
H4(2)	β_{10}	-0.0449(0.0130)*	0.0534(0.0228)*
H4(3)	β_{11}	-0.0278(0.0115)*	0.2619(0.0207)*
H4(4)	β_{12}	0.0654(0.0118)*	-0.0800(0.0208)*
H4(5)	β_{13}	0.0699(0.0140)*	-0.0467(0.0248)*
H4(6)	β_{14}	-0.0271(0.0263)	-0.1157(0.0470)*
H4(7)	β_{15}	0.1126(0.0269)*	0.0506(0.0426)
H4(8)	β_{16}	-0.0084(0.0243)	0.1039(0.0426)*
H4(9)	β_{17}	0.0249(0.0159)	-0.1497(0.0289)*
H4(10)	β_{18}	0.0829(0.0135)*	0.1478(0.0238)*
H4(11)	β_{19}	0.0396(0.0116)*	-0.1640(0.0206)*
H5(1)	β_{2o}	-0.0274(0.0086)*	-0.1186(0.0154)*
H6	β_{21}	-0.0012(0.0004)*	0.0029(0.0007)*
Type(AZ)	β_{22}	-0.0709(0.0214)*	-0.0356(0.0557)
Type(PZ)	β_{23}	-0.0401(0.0242)	0.0441(0.0608)
Taal(B)	β_{24}	-0.1762(0.0379)*	-0.0028(0.0526)
Taal(F)	β_{25}	-0.2454(0.0114)*	-0.0644(0.0173)*
Measurement(1)	β_{26}	-0.0371(0.0286)	-0.0270(0.0794)
Measurement*type(AZ)	β_{27}	0.0065(0.0311)	0.0733(0.0811)
Measurement*type(PZ)	β_{28}	0.0191(0.0360)	0.0586(0.0896)
Variance components			
Residual	σ^2	0.2813	0.8513
Random effect	τ_{AZ}^2	0.0047	0.0075
Random effect	τ_{PZ}^2	0.0046	0.0179
Random effect	τ_{SP}^2	0.0005	0.0149
Intracluster correlation	ρ_{AZ}	0.0164	0.0009
Intracluster correlation	ρ_{PZ}	0.0161	0.0206
Intracluster correlation	ρ_{SP}	0.0002	0.0172

Emperical standard errors; significant at 5% level;

4.2.2 Multiple Imputation (MI) Analysis

Results of the analysis using the multiple imputed dataset described in the methodology section are shown below in Table 7. A similar model to that of the CC-analysis was considered as follow with similar description for the model parameters.

$$\text{Dimension}_{ij} = b_i + \beta_1 E1(1)_i + \beta_2 E1(2)_i + \beta_3 E1(3)_i + \beta_4 E1(4)_i + \beta_5 G1_i + \beta_6 H1_i + \beta_7 H2_i + \beta_8 H3_i + \beta_9 H4(1)_i + \beta_{10} H4(2)_i + \beta_{11} H4(3)_i + \beta_{12} H4(4)_i + \beta_{13} H4(5)_i + \beta_{14} H4(6)_i + \beta_{15} H4(7)_i + \beta_{16} H4(8)_i + \beta_{17} H4(9)_i + \beta_{18} H4(10)_i + \beta_{19} H4(11)_i + \beta_{20} H5(1)_i \ \beta_{21} H6_i + \beta_{22} \text{Type(AZ)}_i + \beta_{23} \text{Type(PZ)}_i + \beta_{24} \text{Taal(B)}_i + \beta_{25} \text{Taal(F)}_i + \beta_{26} \text{measurement(1)}_i + \beta_{27} \text{measurement*Type(AZ)}_i + \beta_{28} \text{measurement*Type(PZ)}_i + \varepsilon_{ij.}$$

When patient safety was measured via supervisor/manager expectations and actions promoting safety (D1score); E1, G1, H1, 2, H3, H4, H5, H6 Type, Taal and measurement had a significant effect on the dimension.

Considering organizational learning-continuous improvement (D2score); E1, G1, H1, H2, H3, H4, H5, Type and Taal were significant. There was also a significant difference in patient safety over time for the type of hospital (measurement*type interaction).

When team work within units was the cards (D3score) E1, G1, H2, H3, H4, H5, H6, Type and Taal had a significant effect on the dimension.

Measuring patient safety via communication openness (D4score); E1, G1, H1, H2, H3, H4, H6, Taal, Type, measurement and measurement*type had a significant effect.

D5score measures feedback and communication about error pertaining to patient safety. E1, G1, H1, H2, H3, H4, H5, H6, Type, Taal, measurement and measurement*Type all had a significant effect on the dimension.

When non-punitive response to error was the case (D6score); E1, G1, H1, H2, H3, H4, H5, H6, Type, Taal, measurement and measurement*Type had a significant effect.

Staffing (D7score) was used to measure patient safety. Similar to the D5score and D6score; . E1, G1, H1, H2, H3, H4, H5, H6, Type, taal, measurement and measurement*Type had a significant effect on the dimension.

D8score records management support for patient safety. For this dimension, E1, G1, H1, H2, H3, H4, H5, H6, Type, Taal and measurement were significant.

Promoting patient safety via team work across units (D9score), E1, G1, H1, H2, H3, H4, H5, H6, Type and Taal had a significant effect the dimension.

Considering handoffs and transitions (D10score) in measuring patient safety; E1, G1, H3, H4, H, H6 Type, Taal and measurement had a significant effect on the dimension.

Considering the outcome dimensions, o1score records overall perception of patient safety; E1, G1, H1, H2, H3, H4, H5, H6, Type, Taal and measurement had a significant effect.

Meanwhile for frequency of events reported (o2score), similar results like in the o1score were observed. E1, G1, H1, H2, H3, H4, H5, H6, Type, Taal and measurement were significant.

A minimum and a maximum intra class correlation of 0.0001 (AZ hospital) and 0.0926 (SP hospital) for D10score and D8score indicates some clustering with the respondents within the hospitals.

Tabel 7: Parameter Estimates and Standard Errors of Multiple Imputation (MI) Analysis

Effect	Parameter	D1scoreParameter estimates (s.e)	D2score Parameter estimates (s.e)	D3score Parameter estimates (s.e)	D4score Parameter estimates (s.e)
Intercept	β_0	3.1859(0.0223)*	3.2594(0.0137)*	3.5326(0.0225)*	3.3407(0.0246)*
E1(1)	β_1	0.6476(0.0121)*	0.5269(0.0071)*	0.5673(0.0082)*	0.7358(0.0088)*
E1(2)	β_2	0.3334(0.0064)*	0.3471(0.0038)*	0.3005(0.0044)*	0.4029(0.0047)*
E1(3)	β_3	-0.0448(0.0063)*	0.0808(0.0037)*	0.0028(0.0043)	-0.0072(0.0046)
E1(4)	β_4	-0.4326(0.0085)*	-0.3257(0.0051)*	-0.3428(0.0057)*	-0.4153(0.0062)*
G1	β_5	0.0253(0.0003)*	0.0086(0.0002)*	-0.0022(0.0002)*	0.0041(0.0003)*
H1	β_6	0.0022(0.0003)*	0.0037(0.0002)*	-0.0003(0.0002)	0.0008(0.0002)*
H2	β_7	-0.0032(0.0002)*	-0.0016(0.0002)*	-0.0027(0.0002)*	-0.0014(0.0002)*
H3	β_8	-0.0051(0.0001)*	0.0012(0.0001)*	-0.0003(0.0001)*	0.0005(0.0001)*
H4(1)	β_9	-0.0712(0.0056)*	0.0706(0.0034)*	0.0463(0.0039)*	0.0123(0.0041)*
H4(2)	β_{10}	0.0627(0.0088)*	0.2498(0.0052)*	0.2517(0.0060)*	0.3848(0.0064)*
H4(3)	β_{11}	0.2335(0.0074)*	0.0736(0.0044)*	0.0052(0.0051)	0.0425(0.0054)*
H4(4)	β_{12}	-0.0846(0.0077)*	0.0142(0.0046)*	0.0820(0.0053)*	0.1468(0.0057)*
H4(5)	β_{13}	-0.0431(0.0096)*	0.0525(0.0057)*	0.0939(0.0066)*	0.1471(0.0070)*
H4(6)	β_{14}	-0.1192(0.0191)*	-0.0889(0.0114)*	-0.0141(0.0130)	-0.0091(0.0139)
H4(7)	β_{15}	0.1147(0.0192)*	0.2914(0.0113)*	0.1502(0.0129)*	0.4041(0.0128)*
H4(8)	β_{16}	0.1179(0.0164)*	-0.0413(0.0097)*	-0.1548(0.0112)*	-0.1715(0.0119)*
H4(9)	β_{17}	-0.0936(0.0106)*	0.0187(0.0063)*	-0.0095(0.0072)	-0.0342(0.0077)*
H4(10)	β_{18}	0.1639(0.0089)*	-0.0510(0.0053)*	-0.0202(0.0061)*	0.0755(0.0065)*
H4(11)	β_{19}	-0.1598(0.0075)*	-0.0564(0.0045)*	0.0768(0.0052)*	-0.0414(0.0056)*
H5(1)	β_{2o}	-0.0783(0.0053)*	-0.0332(0.0031)*	0.0291(0.0036)*	-0.0065(0.0039)
H6	β_{21}	0.0027(0.0003)*	-0.0001(0.0002)	-0.0018(0.0002)*	-0.0007(0.0002)*
Type(AZ)	β_{22}	-0.0302(0.0208)	-0.1027(0.0132)*	-0.0734(0.0222)*	0.0161(0.0240)
Type(PZ)	β_{23}	0.0478(0.0232)*	-0.0716(0.0143)*	0.0195(0.0242)	0.0174(0.0256)
Taal(B)	β_{24}	-0.0345(0.0221)	0.1616(0.0185)*	0.1676(0.0239)*	0.0828(0.0180)*
Taal(F)	β_{25}	-0.0755(0.0069)*	0.1897(0.0054)*	0.2008(0.0074)*	0.0247(0.0058)*
Measurement(1)	β_{26}	0.0033(0.0289)	0.0047(0.0179)	-0.0228(0.0307)	-0.0690(0.0337)*
Measurement*type(AZ)	β_{27}	0.0369(0.0297)	-0.0459(0.0188)*	0.0059(0.0315)	0.0418(0.0342)
Measurement*type(PZ)	β_{28}	0.0484(0.0329)	0.0454(0.0203)*	0.0169(0.0343)	0.0819(0.0363)*
Variance components					
Residual	σ^2	0.8644	0.3020	0.3970	0.4597
Random effect	τ_{AZ}^2	0.0067	0.0059	0.0104	0.0049
Random effect	τ_{PZ}^2	0.0156	0.0057	0.0187	0.0132
Random effect	τ_{SP}^2	0.0089	0.0037	0.0149	0.0183
Intracluster correlation	ρ_{AZ}	0.0008	0.0192	0.0255	0.0105
Intracluster correlation	ρ_{PZ}	0.0177	0.0185	0.0449	0.0279
Intracluster correlation	ρ_{SP}	0.0102	0.0121	0.0362	0.0383

Emperical standard errors; significant at 5% level.

Effect	Parameter	D5scoreParameter estimates (s.e)	D6score Parameter estimates (s.e)	D7score Parameter estimates (s.e)	D8score Parameter estimates (s.e)
Intercept	β_0	3.3611(0.0375)*	3.1938(0.0224)*	3.3080(0.0303)*	3.3824(0.0334)*
E1(1)	β_1	0.7891(0.0101)*	0.5455(0.0091)*	0.5403(0.0085)*	0.7193(0.0081)*
E1(2)	β_2	0.4464(0.0054)*	0.2960(0.0048)*	0.3180(0.0045)*	0.4467(0.0043)*
E1(3)	β_3	-0.0189(0.0053)*	-0.0021(0.0049)	-0.0146(0.0044)*	0.0792(0.0042)*
E1(4)	β_4	-0.5088(0.0071)*	-0.3260(0.0064)*	0.3106(0.0059)*	-0.4531(0.0057)*
G1	β_5	0.0115(0.0003)*	0.0039(0.0003)*	-0.0056(0.0003)*	-0.0058(0.0002)*
H1	β_6	0.0009(0.0002)*	-0.0006(0.0002)*	-0.0005(0.0002)*	-0.0005(0.0002)*
H2	β_7	-0.0021(0.0002)*	0.0005(0.0002)*	-0.0024(0.0002)*	-0.0014(0.0002)*
H3	β_8	0.0007(0.0001)*	-0.0003(0.0001)*	-0.0029(0.0001)*	-0.0016(0.0001)*
H4(1)	β_9	-0.0146(0.0048)*	0.0320(0.0044)*	-0.0649(0.0041)*	-0.1210(0.0039)*
H4(2)	β_{10}	0.1657(0.0074)*	0.4547(0.0067)*	0.1262(0.0063)*	0.1188(0.0059)*
H4(3)	β_{11}	0.1323(0.0062)*	-0.0106(0.0056)	-0.1214(0.0053)*	0.0433(0.0051)*
H4(4)	β_{12}	-0.1551(0.0065)*	0.1934(0.0059)*	0.1303(0.0056)*	-0.0763(0.0053)*
H4(5)	β_{13}	-0.0342(0.0081)*	0.2173(0.0073)*	0.0789(0.0068)*	-0.0400(0.0065)*
H4(6)	β_{14}	-0.1798(0.0161)*	0.0521(0.0145)*	-0.0332(0.0136)*	-0.1506(0.0129)*
H4(7)	β_{15}	0.2353(0.0161)*	0.4395(0.0144)*	0.0848(0.0135)*	0.1252(0.0128)*
H4(8)	β_{16}	-0.1295(0.0137)*	-0.0126(0.0124)	-0.0112(0.0116)	0.0119(0.0110)
H4(9)	β_{17}	-0.0736(0.0089)*	0.0382(0.0080)*	0.0455(0.0075)*	0.1063(0.0072)*
H4(10)	β_{18}	0.0222(0.0075)*	-0.0327(0.0068)*	0.1916(0.0064)*	-0.0471(0.0061)*
H4(11)	β_{19}	-0.1985(0.0064)*	0.1028(0.0058)*	0.1075(0.0054)*	-0.0619(0.0052)*
H5(1)	β_{2o}	-0.1128(0.0045)*	-0.0426(0.0040)*	-0.0089(0.0038)*	-0.0851(0.0036)*
H6	β_{21}	0.0006(0.0002)8*	-0.0005(0.0002)*	0.0006(0.0002)*	0.0039(0.0002)*
Type(AZ)	β_{22}	-0.0876(0.0372)*	-0.1377(0.0219)*	-0.3091(0.0301)*	-0.1710(0.0336)*
Type(PZ)	β_{23}	-0.0216(0.0376)	-0.3448(0.0246)	-0.0371(0.0311)	-0.0579(0.0355)
Taal(B)	β_{24}	-0.0188(0.0254)	-0.0282(0.0245)	0.0767(0.0266)*	-0.1501(0.0333)*
Taal(F)	β_{25}	-0.0656(0.0074)*	-0.0621(0.0077)*	0.1778(0.0079)*	-0.0779(0.0101)*
Measurement(1)	β_{26}	-0.0854(0.0521)	-0.0952(0.0303)*	-0.1202(0.0419)*	-0.0301(0.0465)
Measurement*type(AZ)	β_{27}	0.0696(0.0527)	0.0430(0.0312)	0.1127(0.0427)*	-0.0765(0.0476)
Measurement*type(PZ)	β_{28}	0.0653(0.0533)	0.0865(0.0348)*	0.1098(0.0441)*	-0.0593(0.0501)
Variance components					
Residual	σ^2	0.6088	0.4898	0.4292	0.3850
Random effect	τ_{AZ}^2	0.0109	0.0106	0.0131	0.0224
Random effect	τ_{PZ}^2	0.0064	0.0237	0.0138	0.0311
Random effect	τ_{SP}^2	0.0482	0.0137	0.0308	0.0393
Intracluster correlation	ρ_{AZ}	0.0176	0.0212	0.0296	0.0549
Intracluster correlation	ρ_{PZ}	0.0104	0.0462	0.0312	0.0747
Intracluster correlation	ρ_{SP}	0.0734	0.0272	0.0669	0.0926

Emperical standard errors; significant at 5% level.

Effect	Parameter	D9scoreParameter estimates (s.e)	D10score Parameter estimates (s.e)	O1score Parameter estimates (s.e)	O2score Parameter estimates (s.e)
Intercept	β_0	3.3333(0.0272)*	3.1704(0.0249)*	3.3277(0.0117)*	3.0162(0.0222)*
E1(1)	β_1	0.4863(0.0069)*	0.5278(0.0078)*	0.9378(0.0071)*	0.6603(0.0121)*
E1(2)	β_2	0.2767(0.0037)*	0.2968(0.0041)*	0.5438(0.0038)*	0.3393(0.0064)*
E1(3)	β_3	0.0041(0.0036)	-0.0069(0.0042)	0.0535(0.0037)*	-0.0411(0.0063)*
E1(4)	β_4	-0.2756(0.0049)*	-0.2912(0.0055)*	-0.5639(0.0049)*	-0.4262(0.0085)*
G1	β_5	-0.0051(0.0002)*	-0.0084(0.0002)*	-0.0094(0.0002)*	0.0251(0.0003)*
H1	β_6	0.0005(0.0002)*	-0.0014(0.0002)*	-0.0006(0.0002)*	0.0014(0.0003)*
H2	β_7	-0.0012(0.0002)*	-0.0016(0.0002)*	-0.0002(0.0001)	-0.0029(0.0002)*
H3	β_8	-0.0008(0.0001)*	-0.0019(0.0001)*	-0.0006(0.0001)*	0.0004(0.0001)*
H4(1)	β_9	-0.0046(0.0034)	0.0751(0.0037)*	-0.0385(0.0033)*	0.0712(0.0056)*
H4(2)	β_{10}	0.1565(0.0051)*	0.1409(0.0057)*	0.0847(0.0052)*	0.0627(0.0088)*
H4(3)	β_{11}	0.0160(0.0044)*	0.0570(0.0049)*	-0.0288(0.0044)*	0.2335(0.0074)*
H4(4)	β_{12}	0.0953(0.0046)*	-0.0319(0.0050)*	0.0709(0.0046)*	-0.0846(0.0077)*
H4(5)	β_{13}	0.0559(0.0056)*	-0.0218(0.0063)*	0.0765(0.0057)*	-0.0431(0.0096)*
H4(6)	β_{14}	0.0149(0.0111)	-0.0857(0.0124)*	-0.0298(0.0113)*	-0.1192(0.0192)*
H4(7)	β_{15}	0.0972(0.0112)*	-0.2694(0.0123)*	0.1231(0.0112)	0.1147(0.0191)*
H4(8)	β_{16}	-0.1150(0.0095)*	-0.2268(0.0106)*	0.0172(0.0096)*	0.1179(0.0164)*
H4(9)	β_{17}	0.0520(0.0062)*	0.0144(0.0069)*	0.0439(0.0062)*	-0.0936(0.0106)*
H4(10)	β_{18}	-0.0486(0.0052)*	-0.0845(0.0058)*	0.0998(0.0053)*	0.1639(0.0089)*
H4(11)	β_{19}	-0.0037(0.0045)	-0.1460(0.0049)*	0.0426(0.0045)*	-0.1598(0.0075)*
H5(1)	β_{20}	-0.0163(0.0031)*	0.0168(0.0034)*	-0.0234(0.0031)*	-0.0783(0.0053)*
H6	β_{21}	0.0012(0.0002)*	0.0026(0.0018)*	-0.0013(0.0002)*	0.0027(0.0003)*
Type(AZ)	β_{22}	-0.2058(0.0271)*	-0.1854(0.0246)*	-0.0777(0.0110)*	-0.0302(0.0208)*
Type(PZ)	β_{23}	-0.0485(0.0286)	-0.1060(0.0266)*	-0.0409(0.0122)*	0.0478(0.0232)*
Taal(B)	β_{24}	-0.1535(0.0219)*	-0.1972(0.0209)*	-0.1899(0.0173)*	-0.0345(0.0221)
Taal(F)	β_{25}	-0.1140(0.0068)*	-0.0950(0.0066)*	-0.2610(0.0051)*	-0.0755(0.0069)*
Measurement(1)	β_{26}	-0.0501(0.0377)	0.0113(0.0343)	-0.0311(0.0148)*	0.0033(0.0289)
Measurement*type(AZ)	β_{27}	0.0476(0.0383)	0.0227(0.0349)	0.0022(0.0157)	0.0369(0.0297)
Measurement*type(PZ)	β_{28}	0.0240(0.0404)	0.0320(0.0377)	0.0231(0.0173)	0.0484(0.0329)
Variance components					
Residual	σ^2	0.2884	0.3597	0.2970	0.8644
Random effect	τ_{AZ}^2	0.0089	0.0077	0.0050	0.0067
Random effect	τ_{PZ}^2	0.0176	0.0198	0.0049	0.0157
Random effect	τ_{SP}^2	0.0256	0.0199	0.0017	0.0089
Intracluster correlation	ρ_{AZ}	0.0299	0.0001	0.0166	0.0008
Intracluster correlation	ρ_{PZ}	0.0575	0.0522	0.0162	0.0178
Intracluster correlation	ρ_{SP}	0.0815	0.0524	0.0006	0.0102

Emperical standard errors; significant at 5% level.

4.2.3 Testing for the Need of Random Effect

In order to check whether there was a need for the random effect (random intercept); a mixture of chi-squares was used. For the null hypothesis of H_0: $d_{11}= 0$, the obtained log-likelihood ratio statistics from REML method was used (Table 8). The likelihood ratio statistics has a distribution that is approximately 50:50 mixture of two chi-square distributions. The p value is calculated as a comparism of model 2 versus model 1. Considering the D1score for example; p= $P(\chi^2_{0:1}>3194.7) = \frac{1}{2} P(\chi^2_0>3194.7)+ \frac{1}{2}P(\chi^2_1>3194.7)$. But a χ^2_0 random variable gives a probability mass of 1 to the value 0 and a mass of 0 to every

other value. Hence the contribution of this term to the p value of the test statistics is zero. Table 9 shows the observed $-2\text{In}(\lambda_N)$. Most of the values are >90 and yields p values below 0.0001 (Verbeke and Molenberghs, 2000). It can be concluded the covariance structure should not be reduced by removing the random intercept from the model. In a similar manner, the need for separate random effects for the hospital types was investigated comparing model 1 versus model 3 $-2\text{In}(\lambda_N)$ values suggested a separate random effect may not be needed for the D2score dimension. The P value for this dimension is calculated as follow; p= $P(\chi^2_{1:1}>3194.7) = \frac{1}{2} P(\chi^2_1>2.8)+ \frac{1}{2}P(\chi^2_1>2.8)$. This yields a P value of approximately 0.1000

Table 8: Likelihood Ratio Test for the Need of Random Effect

Dimension	REML Estimates Model 1: Intercept	REML Estimates Model 2:_____	REML estimates Model3: Intercept +group =type
D1score	814932.5	818127.2	814912.6
D2score	709630.9	714206.9	709628.1
D3score	828044.4	836116.8	828015.2
D4score	890464.8	893333.4	890380.8
D5score	1011761.0	1016013.0	1011666.0
D6score	918483.5	924479.4	918428.5
D7score	861668.8	870281.7	861646.1
D8score	815661.0	837218.6	815643.0
D9score	690686.1	698456.3	690624.1
D10score	785493.8	791792.3	785405.4
O1score	702281.1	705487.6	702270.1
O2score	1162200.0	1164124.0	1162159.0

Table 9: $-\mathbf{2In}(\boldsymbol{\lambda_N})$ for the Dimensions

Dimension	Hypothesis	$-2\text{In}(\lambda_N)$	Hypothesis	$-2\text{In}(\lambda_N)$
D1score	Model 2versus model1	3,194.7	Model 1 versus Model 3	19.9
D2score	Model 2versus model1	4,576.0	Model 1 versus Model 3	2.8
D3score	Model 2versus model1	8,072.4	Model 1 versus Model 3	29.2
D4score	Model 2versus model1	2,868.6	Model 1 versus Model 3	84.0
D5score	Model 2versus model1	4,252.0	Model 1 versus Model 3	95.0
D6score	Model 2versus model1	5,995.9	Model 1 versus Model 3	55.0
D7score	Model 2versus model1	8,612.9	Model 1 versus Model 3	22.7
D8score	Model 2versus model1	21,557.6	Model 1 versus Model 3	18.0
D9score	Model 2versus model1	17,770.2	Model 1 versus Model 3	62.0
D10score	Model 2versus model1	6,298.5	Model 1 versus Model 3	88.4
O1score	Model 2versus model1	3,208.5	Model 1 versus Model 3	11
O2score	Model 2versus model1	1,924.0	Model 1 versus Model 3	41

5 CONCLUSSIONS AND RECOMMENDATIONS

The mean dimensional score for all the dimensions were >3 measured on scale of 5. The median scores for these dimensions were highest in 2011, the second measurement period. There was a low level of pairwise association between the dimensions. This association was more evident between the D4score and D5score which measures communication openness and feedback and communication about errors. D9score and D10score also presented with a pair wise correlation slightly higher than the other dimensions.

Results from the CC analysis were not far from those of the MI analysis. The number of significant covariates depended on the dimension and the approach of analysis considered. E1 which measures patient safety grade, number of events reported (G1), staff position in the hospital (H4) and language spoken in the hospital (Taal) were significant for all the dimensions for both the CC and MI analysis. The MI analysis in general had more significant covariates as compared to the CC analysis. This might mean that the MI analysis is more sensitive in accounting for the variability induced by missingness. For example considering the o1score and o2score; measurement (time effect) turn out to be significant for the MI analysis but not for the CC analysis. This was an important key factor that could be missed out if missingness was not taken in to account.

For most of the dimensions there was a significant effect of time (measurement). There was also a significant effect of the type of hospital over time. There was also an indication of a small intra cluster correlation within the hospitals. A higher between variability was seen for the long term care hospitals. The highest intra cluster correlation was 0.0926 for this hospital (SP). The likelihood ratio test also confirmed the need for the random effect to account for the possible correlations between respondents of the same hospital. This analysis also indicated the need for separate random effects for the different hospital types.

Even though the MI analysis via PROC MI seemed more optimal as compared to the complete case analysis, a sensitivity analysis under a different imputation model was done to ensure there were no model defects. This analysis was done using the AMELIA software which operates from the R software under a similar assumption of multivariate normality for the variables. Similar results obtained from this analysis as compared to those described in the methodology section 3.2.2; suggests the model was considerable ok under those assumptions.

This analysis was carried out using data that was voluntarily submitted to the database. Hence, the number of samples in the different hospital types was not balanced. In order to have comparable sample sizes for the different hospitals, a defined sampling approach could be included in the design. Application of weights could contribute to remedy this problem. This study could also be continued on a more regular basis in a well defined time frame so that in feature possible trends in the timing could be investigated.

6 REFERENCES

1. Vlayen, A., Hellings, J., Claes N, Peleman, H., Schrooten, W. (2011). A nationwide Hospital Survey on Patient Safety Culture in Belgian hospitals: setting priorities at the launch of a 5-year patient safety plan. *BMJ Qual Saf*.

2. Instruments for the Belgian Hospital Survey on Patient Safety Culture. Switch Language for French or German Versions. Available at: http://www.health.belgium.be/eportal/Healthcare/Healthcarefacilities/Patientsafety/Coordinationpatientqualityands/Pillar1(SGS)/index.htm?fodnlang=nl. [Acessed on 4 September 2012]

3. Health and Safety Commission (of Great Britain) (1993). Organising for Safety: Third Report of the ACSNI (Advisory Committee on the Safety of Nuclear Installations) *Study Group on Human Factors*.Sudbury, England: HSE Books.

4. Clarke, S. and Ward, K. (2006).The role of leader influence, tactics and safety climate in engaging Employees' safety participation Risk Analysis, 26, 1175 - 1186.

5. Glendon, A. I., Clarke, S. G. & Mckenna, E. F. (2006). *Human Safety and Risk Management.* Florida, CRC Press.

6. Reason, J. (1998). Achieving a safe culture: theory and practice Work and Stress, 12, 293 – 306.

7. Georgie, P. (2009). My life: in safe hands?. Agency for Healthcare Research and Quality (AHRQ) Available at: http://www.ahrq.gov/qual/patientsafetyculture/pscintusers.htm. [Accessed on 6 September 2012].

8. Sammer, C. (2009*). Culture of Safety in Hospitals: A Three-Part Analysis of Safety Culture, Evidence-Based Practice Guidelines, and PatientOutcomes* Fort Worth, Tx: University of North Texas Health Science Center.

9. Morello RT, Lowthian JA, Barker AL, McGinnes R, Dunt D, Brand C. BMJ Qual Saf. (2012). Strategies for improving patient safety culture in hospitals: a systematic review. Epub ahead of print.

10. Paul, T. and von Hippel (2007). *Regression with missing y's: an improved strategy for analyzing multiple imputed data, Sociological Methodology*, volume 37.

11. Graham, J., Olchowski, A. and Gilreath, T. (2007) How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory. *8:206.*

12. Patrick, E. McKnight, et al, (2007). *Missing Data: a gentle introduction*, The Guilford Press.

13. Molenberghs, G. and Kenward, M. (2007). *Missing Data in Clinical Studies*, John Wiley & Sons, Ltd.

14. Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*, Chapman & Hall/CRC.

15. Little, R. and Rubin, D. (2002). *Statistical Analysis with Missing Data*, 2nd Edition.

16. Allison, P. (2001). *Missing Data Thousand Oaks, CA:* Sage Publications.

17. Cohen, J. and Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.).Hillsdale, NJ: Erlbaum.

18. Cohen, J., Cohen, P., West, S. and Aiken, L. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd edition. Mahwah, N.J.: Lawrence Erlbaum.

19. Jones, M.(1996). Indicator and stratification methods for missing explanatory variables in multiple linear regression Journal of the American Statistical Association, 91,222-230.

20. Scheuren, F. (2005). Multiple imputation: How it began and continues. The American Statistician, 59, : 315-319.

21 Schafer, J. and Olsden, M. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. Multivariate Behavioral Research, 33, 545-571.

22. Little, R. and Rubin, D (1983). On jointly estimating parameters and missing data by maximizing the complete data likelihood. *The American Statistician, 37,* 218-220.

23. Rubin, D. (1987) *Multiple Imputation for Nonresponse in Surveys*. J. Wiley & Sons, New York

24. Dempster, A., Laird, N., Rubin, D. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society. Series B (Methodological) 39 (1): 1–38. JSTOR 2984875. MR 0501537.

25. Sundberg, R. (1974). Maximum likelihood theory for incomplete data from an exponential family. *Scandinavian Journal of Statistics* **1** (2): 49–58. 4615553. MR 381110.

26. Verbeke, G. and Molenberghs, G. (2000) *Linear Mixed Models for Longitudinal Data*. New York :Springer.

27. Fitzmaurice, G., Laird, N., and Ware, J. (2004) . *Applied Longitudinal Analysis*. New Jersey: John Wiley & Sons,

28. Mallinckrodt, C., Clark, S., Carroll, R., and Molenberghs, G. (2003). Assessing response profiles from incomplete longitudinal clinical trial data under regulatory considerations. *J Biopharm Stat*. 13(2): 179-90.

29. Molenberghs, G. and Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. New York: Springer.

7 APPENDIX

Figure A.1a: Box plots showing Trends in Dimensional Scores Over years- AZ Hospitals

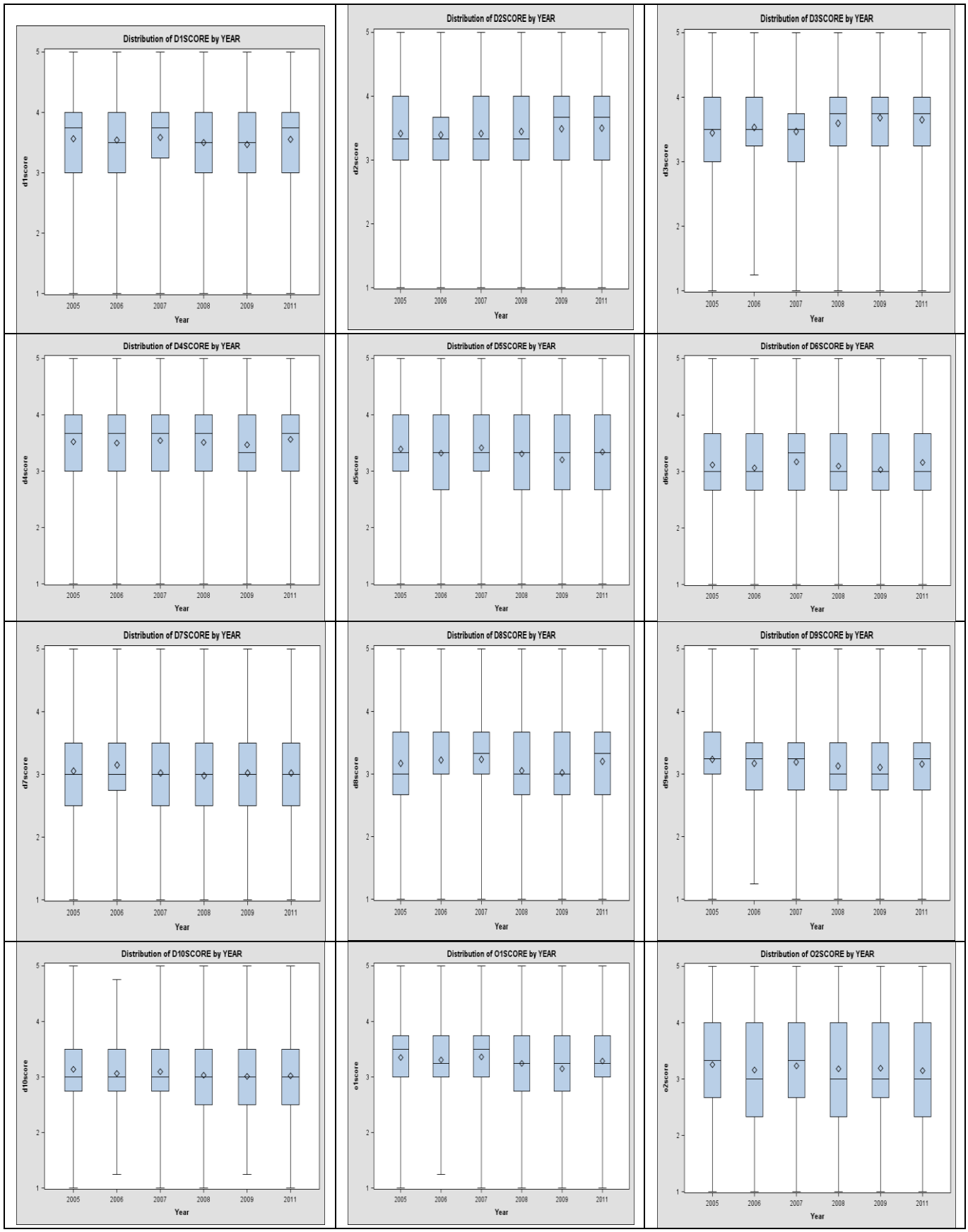


Figure A.1b: Box plots showing Trends in Dimensional Scores Over years- PZ Hospitals

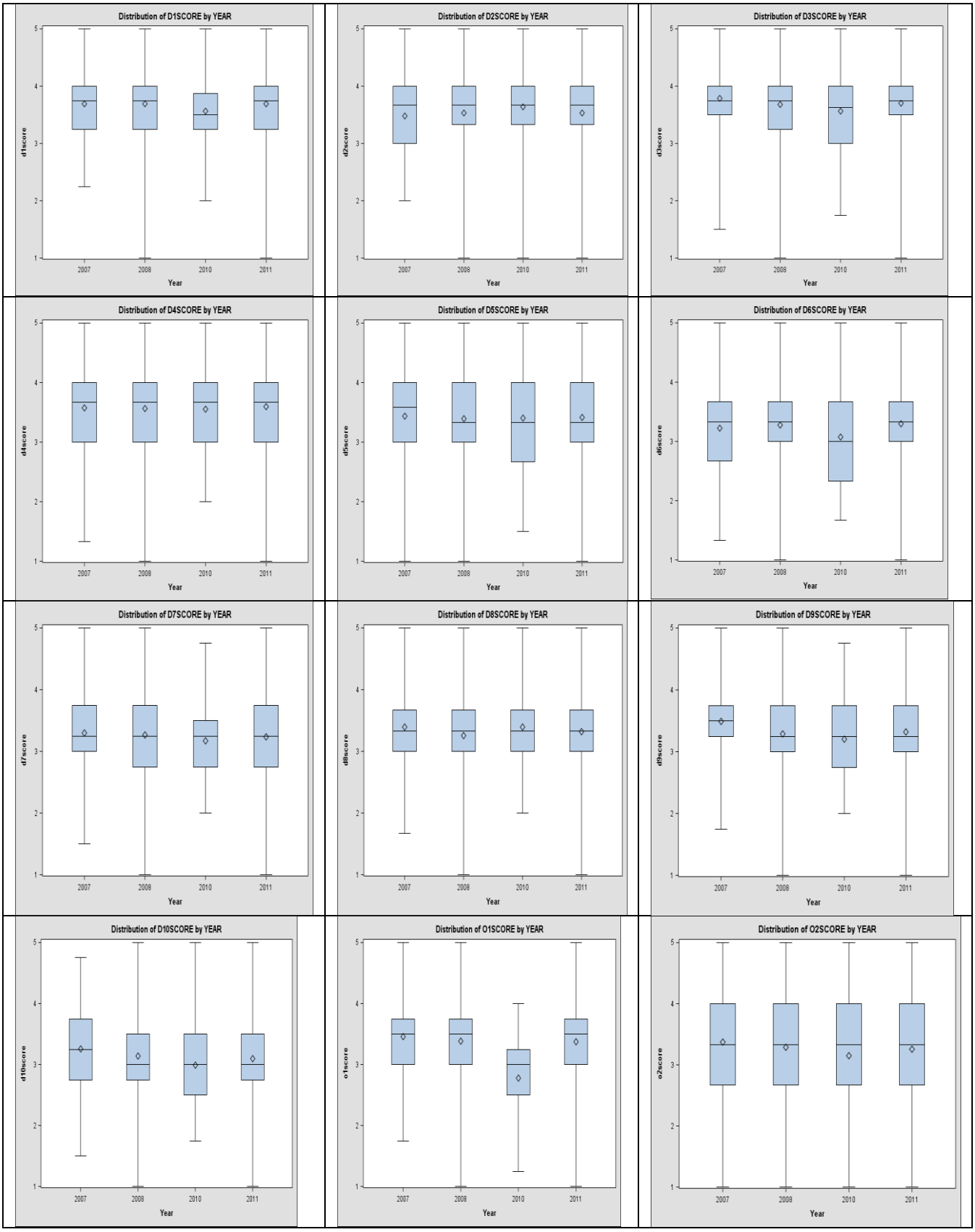


Figure A.1c: Box plots showing Trends in Dimensional Scores Over years- SP Hospitals

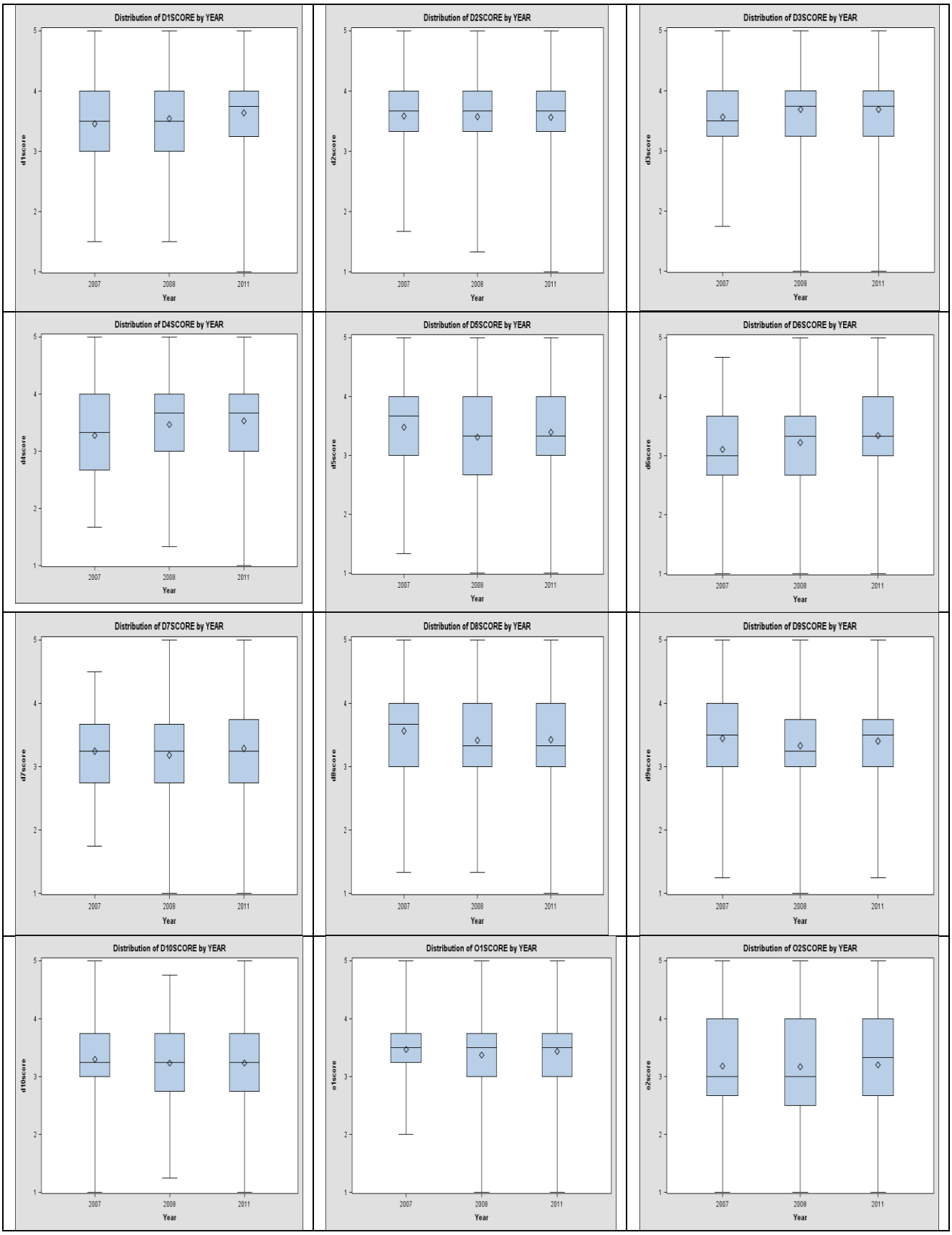
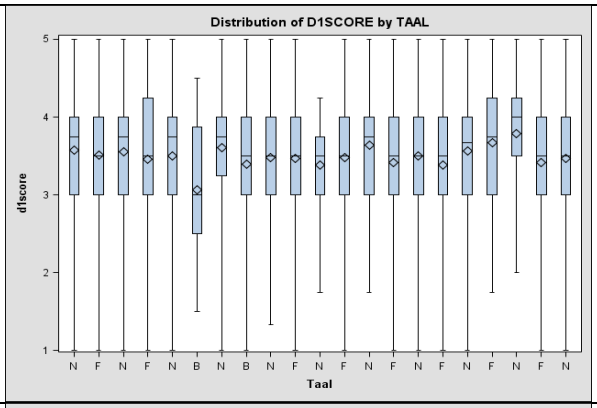
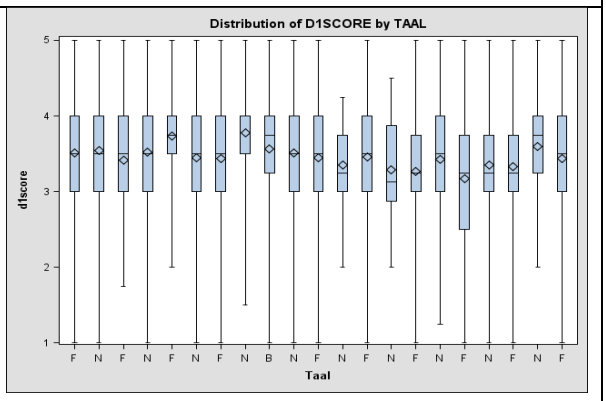
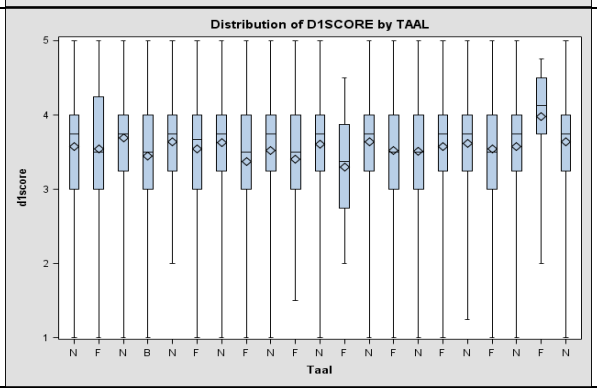


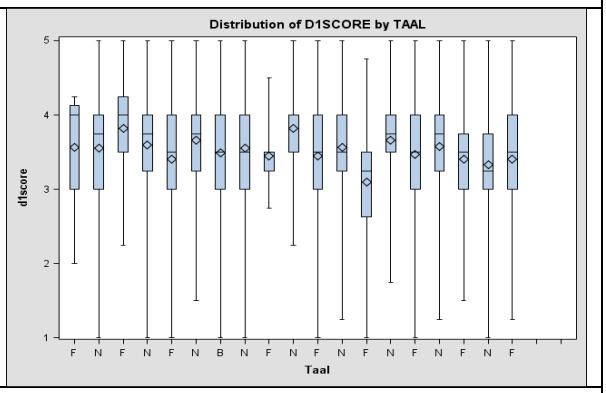
Figure A.2: Distribution of D1score by langauge for AZ, PZ and SP Hospitals

TYPE AZ

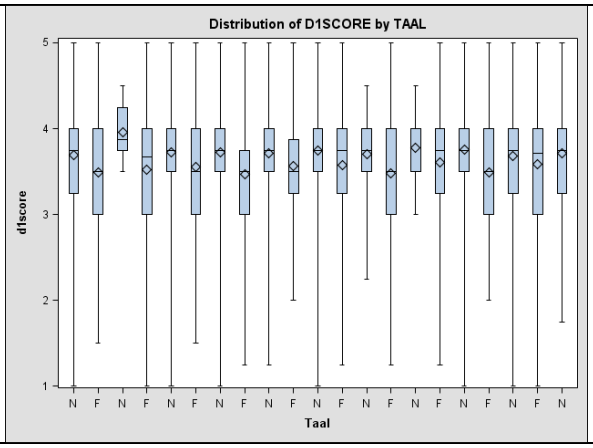








TYPE=PZ



TYPE=SP

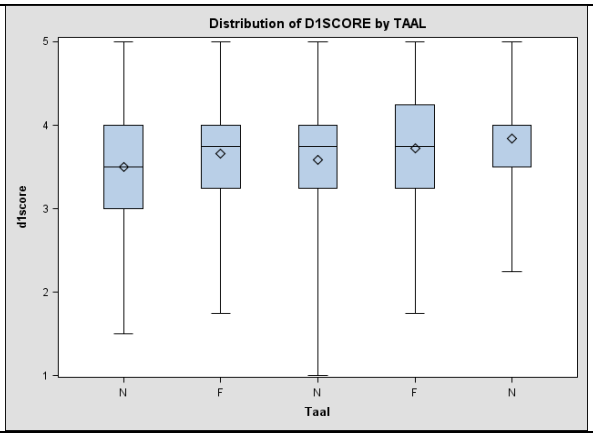


Figure A.3a, A.3b, A.3c Scatter plot Matrix for the Hospitals

Figure A.3a : type AZ

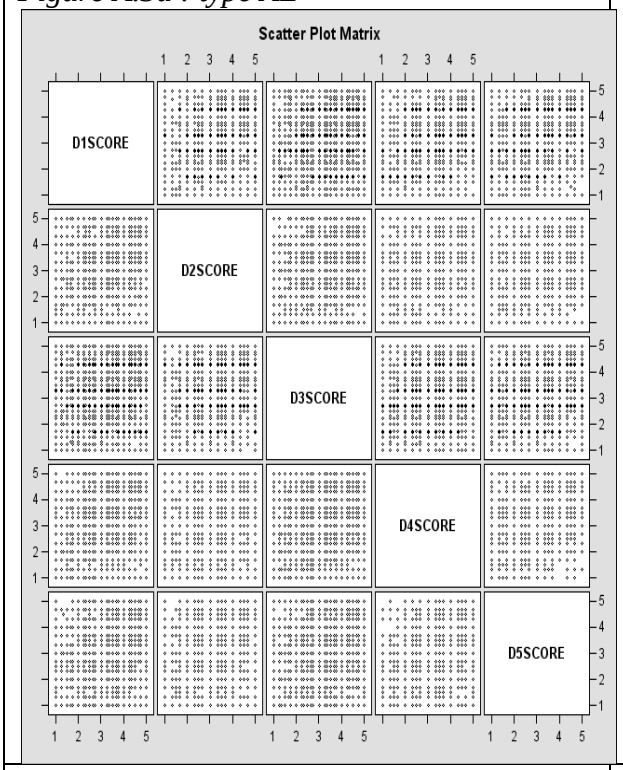


Figure A.3b: type PZ

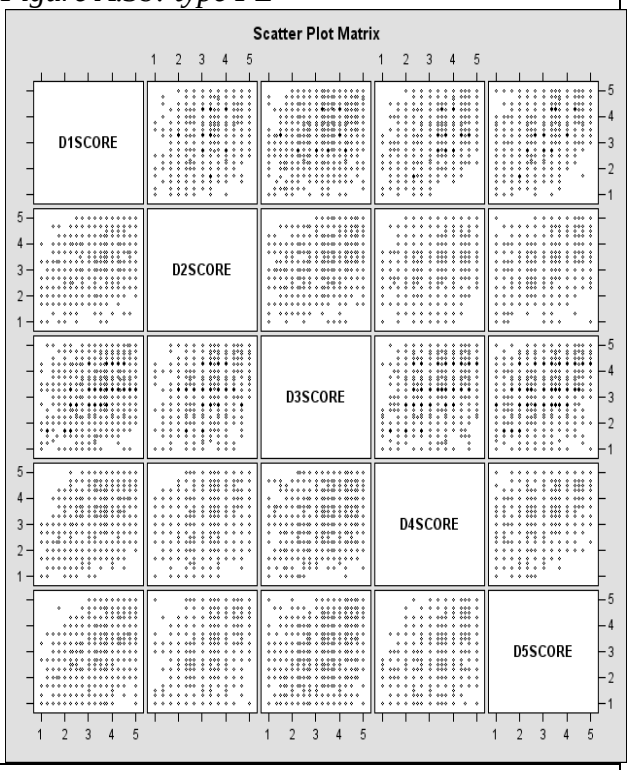
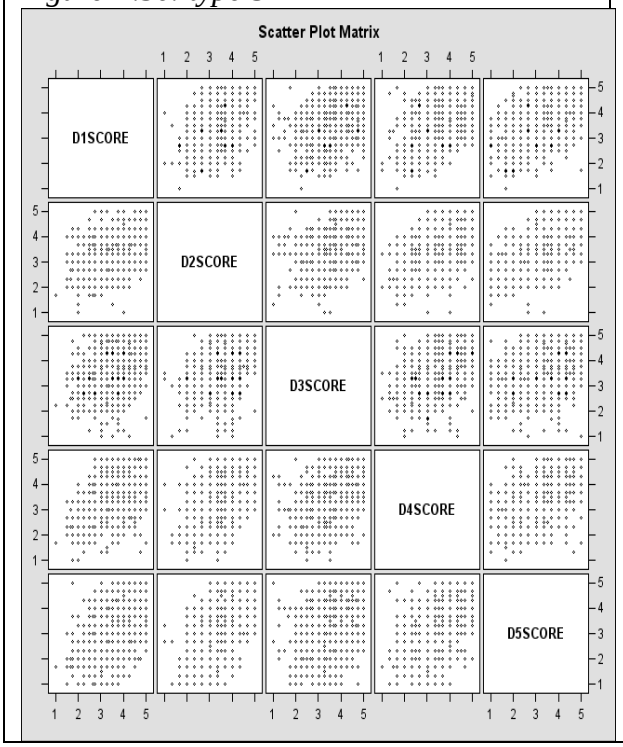


Figure A.3c: type SP



Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
After 5 years of action: has patient safety culture improved in the Belgian hospitals? Results of a benchmark database

Richting: **Master of Statistics-Biostatistics**
Jaar: **2013**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Momuluh, Quinta Nanga

Datum: **4/02/2013**