2014•2015
FACULTY OF SCIENCES
Master of Statistics

Master's thesis

The trajectories of resilience, anxiety and depression in breast cancer patients and control groups

Supervisor :
dr. Jurgen CLAESEN

Supervisor :
Dr. WARD SCHROOTEN
Mrs. SABINE MARKOVITZ

Negasi Asres Mesfin

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics

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

2014•2015
FACULTY OF SCIENCES
Master of Statistics

Master's thesis

The trajectories of resilience, anxiety and depression in breast cancer patients and control groups

Supervisor :
dr. Jurgen CLAESEN

Supervisor :
Dr. WARD SCHROOTEN
Mrs. SABINE MARKOVITZ

Negasi Asres Mesfin

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics

universiteit hasselt | Maastricht University

Certification

I declare that this thesis was written by me under the guidance and counsel of my supervisors.

.....Date.....
Negasi Asres Mesfin Student

We certify that this is the true thesis report written by **Negasi Asres Mesfin** under our supervision and we thus permit its presentation for assessment.

.....Date.....
Dr. Jurgen CLAESEN Internal Supervisor

.....Date.....
Prof. Dr. Ward SCHROOTEN
External Supervisor

..... Date.....
Sabine MARKOVITZ External Supervisor

Acknowledgements

Firstly, I would like to thank almighty GOD and virgin MARY for the measureless blessing of my day-to-day activity. Next, my special thanks goes to my internal supervisor Dr. Jurgen CLAESEN for his continuous professional support and constructive suggestions starting from the initial contacts until the end of the project. My next special thank goes to the external supervisors Prof. Dr. Ward SCHROOTEN and Sabine MARKOVITZ for their generous support, useful comments, remarks and encouragements for all the time and providing reference journals. Next, I cannot find words to express my gratitude to the Vlaamse Interuniversitare Raad (VLIR) university cooperation development program for sponsoring me financially, so that I can focus not only on my thesis research but also on my whole study without any extra difficulties in life. I am also grateful to thank all my master course professors for transferring their practical knowledge in a clear, understandable and easer manner and to all my best friends for sharing experience during the time we spent together.

Lastly, I would also like to thank my lovely family, Moma, Baba, two elder sisters, and younger brothers. They were always supporting me, encouraging me and praying for my success.

Negasi Asres

September, 2015

Genk, Belgium

Abstract

This study was conducted to investigate the trajectories of resilience, anxiety and depression scores for breast cancer patients and control group and the predictors of those trajectories. We used longitudinal data that consist of 211 people in the control group and 274 people in the breast cancer group, in total 485 individuals were followed at three time points for resilience and at two time points for anxiety and depression. A linear mixed model was fitted to see how resilience, anxiety and depression evolves over time. Next, a Latent Class Growth model (LCGM) was used to determine the optimal trajectories in each response and for each group separately. Furthermore, the significant predictors for those trajectories were studied using Multinomial logistic regression model. Results of the linear mixed model indicated that the evolution of resilience, anxiety and depression scores over time depend on cancer and control group (i.e. interaction with time was significant). Using LCGM analysis, we obtained three classes of resilience score (i.e. trajectories of dramatically decreasing, slight decreasing and stable high score for the cancer group and trajectories of low stable, medium stable and high stable score for the control group) and two classes of anxiety and depression score (i.e. trajectories of highly increasing and stable high score for the cancer group and trajectories of highly increasing and slight increasing score for the control group). The analysis of predictors showed that for the resilience score the baseline resilience, optimism, neuroticism, positive emotion and negative emotion in the breast cancer patients and the baseline resilience score, neuroticism and positive emotion in the control group were risk factors (which were identified as significant predictors and related to the grouping or class memberships). For the anxiety score, education, optimism, neuroticism and negative emotion for cancer group and education, age, neuroticism, positive emotion and negative emotion for control women were risk factors related to the class memberships. Finally, for the depression score, positive emotion, negative emotion and baseline resilience score in the cancer group and positive emotion, negative emotion, age, marital status, optimism, and neuroticism in the control group were significant predictors related to the class memberships.

***Keywords**: Resilience, Anxiety, Depression, Latent Class Growth Model*

Contents

List of Tables

List of Figures

1 Introduction

1.1 Background

Life events are defined as discrete experiences that interrupt humans day-to-day activities, causing a substantial change and readjustment. An example of life events are divorce, illness or injury, failing from doing something, changing conditions, losing a job or lover (Jang, Yuri, Haley and William, 2002). Exposure to negative life events has been associated with a variety of adverse physical and psychological health outcomes, including symptoms of depression, anxiety, minimizing life enjoyments or satisfactions (King and Ogle, 2014). Furthermore, research has shown that during the normal lifetime, a majority of individuals are exposed to at least one and often several stressful and life-treating events (Norris, 1992),(King and Ogle, 2014) and (Bonanno, 2004). Although strong negative life events are highly distressing and often potentially debilitating, it is well established that not every individual reacts in the same way (Bonanno, et al., 2012).

Some individuals experience highly distress from which they can not able to recover. And some others suffer less intensely and for a much shorter period of time. Other people also recover very fast or quickly but then begin to experience unexpected health problems. However, many more people manage the temporary suffering or potentially hard situations or events in a very well manner with no apparent disturbance in their ability to function at work or with their family (Bonanno, 2004). Breast cancer diagnosis and treatment are considered as one of the possible traumatic or stressor negative life events and researchers have stressed that dealing with breast cancer could result in poor psychological outcomes (Koutrouli, Anagnostopoulos, and Potamianos, 2012).

Large proportions of people who have experienced cancer appear to find benefit in the experience like improved quality of life, better interpersonal relationships, and changes in values and priorities as a result of their experience (Aspinwall and MacNamara, 2005). With this large overall survival ratio in cancer patients over the last decades, research shifted towards the emotional wellbeing of long-term cancer survivors (Aspinwall and MacNamara, 2005) and (Woodgate, 1999). These studies showed that many patients remain psychologically healthy and higher general happiness (Markovitz, Schrooten, Arntz and Peters, 2015) notwithstanding of the multiple challenges of the disease. Nowadays, interest has grown in finding protective factors of patients which maintain emotional stability in times of adversity (Seligman, and Csikszentmihalyi, 2000) and (Aspinwall, and

Staudinger, 2003). Recent study (Markovitz, Schrooten, Arntz and Peters, 2015) confirmed that resilience is a partially protective factor against emotional distress in cancer patients. Resilience can be define as an individuals ability or capacity for adapting well in the face of negative life events or adversity, health or relationship problems. It helps to maintain or regain mental health following exposure to stress (Rutter, 1993). It means returning back from difficult experiences and situations. There were discrepant views regarding the question whether resilience can be best conceptualized as a latent personality trait that becomes manifest during adversity or as a dynamic process that develops as a result of experiencing adversity (Luthar, Cicchetti, and Becker, 2000). A cross sectional study was conducted by including a control group (Markovitz, Schrooten, Arntz and Peters, 2015) and confirmed that cancer patients didn't differ from the control group in terms of resilience. The level of resilience were the same in both groups. And adversity did not seem to have an impact on the level of resilience. Again the study suggested that resilience might be a relatively stable trait that is not affected by adversity. However, given the cross-sectional nature of this study conclusions need to be confirmed with a longitudinal design.

Currently, investigators are also interested in research aiming on the trajectories of psychological outcome for people exposed to negative life events. For traumatic injury event, the trajectories of depression and other psychological measures were studied in (deRoon-Cassini, Mancini, Rusch and Bonanno, 2010). They identified four latent classes or trajectories of depression symptoms with respect to traumatic injury. For Spinal Cord Injury, the trajectories of depression and anxiety measures were studied in (Bonanno, et al., 2012) and identified four trajectories with respect to depression and three trajectories with respect to anxiety level. There are also other related literatures which were done to study the trajectory pattern in psychological outcome variables.

In this study, we extended this focusing on the trajectories of resilience, depression and anxiety level as well as predictors of those trajectories for people with breast cancer (cancer group) and without breast cancer (control group). Here, including healthy group in this report helps to draw inference in both groups and compare results such as patterns and number of trajectories.

1.2 Objectives

The main goal of this thesis was to study the longitudinal trajectory of resilience,anxiety and depression in breast cancer patients as well as in the control group and to determine predictors of those trajectories. Moreover, to investigate how resilience, anxiety and depression evolves over time.

1.3 Data Description

The data used in this study was obtained from a study conducted in the year between 2009 and 2013, 284 women diagnosed with primary breast cancer were recruited consecutively at time of hospitalization for breast surgery in a hospital in the province of Limburg, Belgium (Markovitz, Schrooten, Arntz and Peters, 2015). Informed consent was obtained from 284 patients, 10 of these did not return the questionnaires thus, a total of 274 patients remained for analyses. Further, a control group of 211 healthy womens were included in order to study the trajectory of resilience in both groups and compare. Except for the diagnosis of cancer, inclusion criteria were identical as for the clinical group. Thus, the data set consists of two groups the breast cancer and control group in total 485 individuals involved. In addition, different psychological variables were measured during the follow up period but, in this study the focus was only on three responses resilience, anxiety and depression. Psychological resilience was measured with the 25-items Connor-Davidson Resilience Scale (CDR)(Connor and Davidson, 2003). Higher scores indicate higher degrees of resilience (Markovitz, Schrooten, Arntz and Peters, 2015). To assess anxiety and depression symptoms, it was used the hospital anxiety (HADSA) and depression Scale (HADSD)(Zigmond and Snaith, 1983). Higher scores reflect higher score of depression and anxiety (Markovitz, Schrooten, Arntz and Peters, 2015). Resilience score of individuals were measured at points such as before surgery (T1), 6 months after surgery (T2) and at the end of the study (T3) (i.e. at one and half years). However, for anxiety and depressive symptom measurements were taken only at two time points at first (T1) and one and half years later (T3). Next to the repeated measurements subjects baseline information were available as well. Table 1 shows all the variables with their corresponding descriptions.

Table 1: Description of predictor and outcome variables in the study

Variable	**Type**	**Description**
ID		Identification number of individuals in the study
Predictor variables		
age	Continuous	baseline age of the patients
Group	Categorical	1=breast cancer patients (S), 2= control group (C)
Martital status	Categorical	1=Married, 2=divorced, 3=widow, 4= single
Education	Categorical	1=primary school, 2= secondary school 6-12y 3= higher education, 4=university degree
posaff	Continuous	Positive emotions of the patient at time T1
negaff	Continuous	Negative emotions of the patient at time T1
LES	Continuous	Life Events Scale at time T1
LOS	Continuous	Life orientation scale (Optimism) at time T1
EPQ	Continuous	Neuroticism at time T1
T1CDR	Continuous	Baseline resilience score
Outcome variables		
CDR, HADSA, HADSD	Continuous	Connor-davidson resilience, anxiety and depression score

2 Methods

2.1 Exploratory Data Analysis

In order to have a general idea in model building process and choosing the preliminary fixed-effects structure, the random-effects structure and the residual covariance structure, the exploratory data analysis will be performed. We use descriptive statistics and graphical techniques to explore the entire variables, individual profiles and average evolutions.

2.2 Logistic Regression Model

The simplest technique to analyze data which is collected over time is by summarizing the measurements at different time points into one measurement per subject and apply the ordinary analyses methods. A number of summary measures are available (Verbeke and Mulenberghs, 2000) one simple case is a paired t-test. However, the choice of this summary depends on the question of interest and the nature of the data. In this report, since almost half of the observation are missed in the second time point (T2) for resilience score in the control group, we ignore difference with the second time point (T2). Furthermore, for anxiety and depression, we have only two measurements at T1 and T3. Thus, the analysis was performed by taking the difference between the two time periods (T3) and (T1) and creating class based on the sign of the difference (T3-T1) for each response. The difference was calculated as follows:

$$\Delta_i = Y_{i3} - Y_{i1}, \qquad for \qquad\qquad i = 1, 2, ...k \tag{1}$$

where, k is number of patients in each group and Y_{i3} and Y_{i1} are responses measured at T3 and T1 respectively. Next, negative values of Δ_i including zero are categorized as decreasing class and positive values of Δ_i as improving class. Although, this categorization was very strict and rough, we applied before appropriate method. Moreover, we need to see if there were relationship between these classes and possible covariates. Then, a logistic regression was fitted considering the decreasing class as a reference. This can be written as:

$$\mathbf{logit}[\pi(\mathbf{x})] = \alpha_{\mathbf{0}} + \sum_{\mathbf{j=1}}^{\mathbf{p}} \beta_{\mathbf{j}} * \mathbf{X_j} \tag{2}$$

where, $\pi(x) = P(Y = 1)$ is probability of improving or positive difference and α is a constant i.e. an intercept and the vector $\mathbf{X_j}$ consists of demographic and intrapersonal predictors.

2.3 Linear Mixed Model (LMM)

Summary measure approaches have a drawback of not optimize the use of all available information (Verbeke and Molenberghs, 2000). Thus, the linear mixed effects model was fitted by assuming the vector of repeated measurements on each subject follows a linear regression model where some of the regression parameters are population specific whereas others are subject-specific. Hence, the model allows fixed effects and subject specific effects. The linear mixed model can be written as:

$$\mathbf{Y_i} = \mathbf{X_i}\boldsymbol{\beta} + \mathbf{Z_i}\mathbf{b_i} + \boldsymbol{\varepsilon_i} \tag{3}$$

where, $\mathbf{b_i} \sim \mathbf{N}(\mathbf{0}, \mathbf{D})$, $\varepsilon_\mathbf{i} \sim N(\mathbf{0}, \Sigma_\mathbf{i})$, $\mathbf{b_1} \ldots \mathbf{b_N}$ and $\varepsilon_\mathbf{1} \ldots \varepsilon_\mathbf{N}$ are independent

$\mathbf{Y_i}$ is the n_i-dimensional response vector for subject i, $1 < i < N$, N is the number of subjects in the study. $\mathbf{X_i}$ and $\mathbf{Z_i}$ are $(n_i \times p)$ and $(n_i \times q)$ dimensional matrix's of known covariates, $\boldsymbol{\beta}$ is a p-dimensional vectors containing the fixed effects, $\mathbf{b_i}$ is a q-dimensional vector of random effects and $\varepsilon_\mathbf{i}$ is n_i-dimensional vector of residual components. Finally, $\mathbf{D}$ is a general $q \times q$ covariance matrix with (i, j) elements of $d_{ij} = d_{ji}$ and Σ_i is $n_i \times n_i$ covariance matrix which depend on i only through its dimension n_i. It follows that from equation 3 conditional on the random effect b_i, Y_i is normally distributed with mean vector $\mathbf{X_i}\boldsymbol{\beta} + \mathbf{Z_i}\mathbf{b_i}$ and covariance matrix Σ_i. Further, b_i assumed to be normally distributed with mean vector $\mathbf{0}$ and covariance $\mathbf{D}$. Let $f(y_i|b_i)$ and $f(b_i)$ be the corresponding density function. The marginal density function of Y_i is then given by $f(y_i) = \int f(y_i|b_i)f(b_i)db_i$ which can easily be shown to be density function of an n_i-dimensional normal distribution with mean $\mathbf{X_i}\boldsymbol{\beta}$ and with covariance matrix $\boldsymbol{V_i} = \boldsymbol{Z_i}\boldsymbol{D}\boldsymbol{Z_i^t} + \boldsymbol{\Sigma}$. (Verbeke and Molenberghs, 2000).

2.4 Latent Class Growth Model (LCGM)

Standard growth analyses estimate a single trajectory that averages the individual trajectories of all participants in a given sample. This average trajectory contains an averaged intercept and slope for the entire sample (Andruff, Carraro, Thompson and Gaudreau, 2009). This approach captures

individual differences by estimating a random coefficient that represents the variability surrounding this intercept and slope. Growth models are useful in case research questions for which all individuals in a given sample are expected to change in the same direction across time with only the degree of change varying between people (Andruff, Carraro, Thompson and Gaudreau, 2009). However, some psychological phenomena may follow a multinomial pattern in which both the strength and the direction of change are varying between people (Nagin, 2002). Therefore, an alternative modeling strategies are available that consider multinomial heterogeneity in change and such approach is a group-based statistical technique known as Latent Class Growth Modeling (Andruff, Carraro, Thompson and Gaudreau, 2009).

LCGM is a semi-parametric technique used to identify distinct subgroups of individuals following a similar pattern of change over time on a given variable. Although each individual has a unique developmental course, the heterogeneity or the distribution of individual differences in change within the data is summarized by a finite set of unique polynomial functions each corresponding to a discrete trajectory (Nagin, 2005). Unlike standard latent growth modeling techniques in which individual differences in both the slope and intercept are estimated using random coefficients, LCGM fixes the slope and the intercept to equality across individuals within a trajectory. Such an approach is acceptable given that individual differences are captured by the multiple trajectories included in the model (Andruff, Carraro, Thompson and Gaudreau, 2009).

Although the model is widely applicable, the rating scale of the instrument used to measure the variable of interest dictates the specific probability distribution used to estimate the parameters. Psychometric scale data necessitate the use of the censored normal model distribution (Andruff, Carraro, Thompson and Gaudreau, 2009). In the censored normal model, each trajectory is described as a latent variable y_{it}^{*j} that represents the predicted score on a given dependent variable of interest (Y) for a given trajectory membership group (j) at a specific time (t) for subject i, and is defined by the following function:

$$y_{it}^{*j} = \beta_0^j + \beta_1^j time_{it} + \beta_2^j time_{it}^2 + \varepsilon_{ij} \tag{4}$$

where, j=1, 2, 3,..possible classes, $time_{it}$ and $time_{it}^2$ are the subject i's time variable entered in linear or quadratic term. ε_{ij} is a disturbance term assumed to be normally distributed with a mean

of zero and a constant variance. Lastly, β_0^j, β_1^j and β_2^j are the parameters defining the intercept and slopes (i.e., linear, quadratic) of the trajectory for a specific subgroup (j). As demonstrated in the above polynomial function, the trajectories are most often modeled using either a linear (time), quadratic ($time^2$) or cubic trend ($time^3$), depending on the number of time points measured. Notwithstanding the numerous advantages of LCGM, one limitation concerns with the number of measurements. As with all growth models, minimum of three time points is preferred for proper estimation of trajectories and to estimate more complex models involving trajectories following cubic or quadratic trend (Curran and Muthen, 1999).

Next, in order to correctly specify the model and obtain proper number of classes or correctly estimate class proportions, we extended the unconditional model (4) by including important baseline covariates (i.e. conditional model) as recommended in (Muthen, 2003).

2.4.1 Selecting Number of Trajectories

In deciding the number of trajectories, investigators evaluate which model provides the best fit to the data by interpreting and comparing both the fit statistics and the posterior probabilities for each model tested. One possible choice for testing the hypothesis of the number of components in a mixture is the likelihood ratio test. However, the null hypothesis is on the boundary of the parameter space, and hence the classical asymptotic results do not hold (Jones, Nagin and Roeder, 2001). To overcome this problem, one can use the change in the BIC between models as an approximation to the log of the Bayes factor (Jones, Nagin and Roeder, 2001). Thus, the Bayesian Information Criterion (BIC) value was used to compare models that include different numbers of trajectories and the approximation can be written as follows: $2log(B_{10}) = 2(\Delta BIC)$ as illustrated in (Jones, Nagin and Roeder, 2001). where, (ΔBIC) value is calculated by subtracting the BIC of simpler model from the complex model. Furthermore, for interpreting the estimate of the log Bayes factor, a guideline of values ranging from 0 to 2 as weak evidence, values ranging from 2 to 6 as moderate evidence, values ranging from 6 to 10 as strong evidence, and values greater than 10 are interpreted as very strong evidence for the more complex model (Jones, Nagin and Roeder, 2001) was applied in this report. At first, depending on the number of time points, the linear, quadratic or cubic functions of each trajectory can be tested. To ensure parsimony as illustrated in (Andruff, carraro, Thompson

and Gaudreau, 2009) the non-significant cubic and quadratic terms are removed from trajectories in a given model, but linear parameters are retained irrespective of significance. Once non-significant terms have been removed, each model is retested and resulted to a new BIC value. This process of comparing the fit of each subsequent, more complex model, to the fit of the previously tested, simpler model, continues until there is no substantial evidence for improvement in model fit. All these were done using PROC TRAJ in SAS software.

2.5 Multinomial Logistic Regression Model

In this study, we are interested in studying how risk factors were associated with the different trajectories or classes identified by using LCGA model above for each response and in both groups separately. Or to determine which covariates are predictive to those extracted classes. Thus, a multinomial logistic regression model was fitted considering the classified groups as a response variable with different categories. Therefore, in case where response variable has J nominal classes or categories, a baseline category logit model are used. The baseline-category model sets the J^{th} category as a reference and it simultaneously describes log odds for all pairs of categories compared to the baseline category, often the last one. The model can be written as follows (Agresti, 2002).

$$logit[P(G = j|\mathbf{x})] = \mathbf{log}(\frac{\pi_{\mathbf{j}}(\mathbf{x})}{\pi_{\mathbf{J}}(\mathbf{x})}) = \alpha_{\mathbf{j}} + \beta_{\mathbf{j}}^{'} * \mathbf{x_j} \qquad \mathbf{j = 1, 2, ...J - 1} \tag{5}$$

where, $\pi_j(x)$ is the probability of the j class or trajectory of the response at a fixed setting $\mathbf{x}$ for explanatory variable, G is the possible group extracted using LCGA and $\mathbf{x}$ is a vector of covariates. But, for responses which had only two trajectories a simple logistic regression was replaced instead. Deciding which covariates to be kept in the statistical model has always been a difficult task for data analysts. Here, in this case to select the best candidate risk factors, stepwise selection procedure was used after including all the main effects and possible interactions. But, statistical significant should not be the only reason for inclusion or exclusion of a variable from the model. Variables known to be important, but not significant could be also included in the model (Agresti, 2002).

2.6 Concept of Missingness

It is often difficult to obtain complete measurement for all subjects in the study specially in data which is measured over time like longitudinal data. This issue can exist for known or unknown reasons. However, depending on the missingness mechanism, a different approach should be used. Thus, the conditional distribution for the missing mechanism could be one of the following possibilities: Missing Completely At Random (MCAR), Missing At Random (MAR) and Missing Not At Random (MNAR).

2.6.1 Missing Completely At Random (MCAR)

Data is said to be missing completely at random (MCAR) when the probability that an observation being missing is unrelated to either the specific values that, in principle, should have been obtained or the set of observed responses. In practice this means that, under MCAR, the analysis of only those units with complete data gives valid inferences (Molenberghs and Verbeke, 2005).

2.6.2 Missing at Random (MAR)

When the probability that an observation being missing depends on the set of observed data but is conditionally unrelated to the specific unobserved data then, it is said to be missing at random. Under this mechanism, analyses based on the direct likelihood are valid under both the linear and generalized linear mixed model (Molenberghs and Verbeke, 2005).

2.6.3 Missing Not at Random (MNAR)

The missingness is assumed to be missing not at random (MNAR) if it is neither MCAR nor MAR. The probability of a measurement being missing depends on unobserved data and simplification of the joint distribution of the full data is not possible (Molenberghs and Verbeke, 2005).

2.6.4 Methods for Handling Missing Data

One simple methods to analyze missing data are using complete case (CC) analysis or only those subjects who have a complete profile are included in the analysis. However, this results in a loss of information and has impact on precision and power or low efficiency. Moreover, this method relies

on the strong (and often unrealistic) assumption of MCAR and if the missingness mechanism is not MCAR, it introduces biased results (Molenberghs and Verbeke, 2005). Despite the limitation of this method, we used complete case analysis only for the purpose of comparison with other analysis in this report. The direct likelihood method for likelihood-based models such as linear mixed model to longitudinal data leads to valid inference under Missing at random (MAR) assumption without modeling the missingness process (Verbeke and Molenberghs 2000, Molenberghs and Verbeke 2005). Furthermore, PROC TRAJ also uses the maximum likelihood method to estimate parameters, including group sizes and shapes of trajectories (Niyonkuru, et al., 2013). Therefore, in this study subjects with missing data are included in the analysis. And multiple imputation method was used for comparison. Multiple imputation (MI) procedure replace each missing value with a set of M plausible values. The imputed data sets are then analyzed by using standard procedures for complete data and combining the results from these analyses. Finally, results were compared under the three approaches complete case (CC), direct likelihood (DL) and multiple imputation (MI). However, parameter estimates can be biased if the missing data are (MNAR). Therefore, it is important to asses the sensitivity of the conclusions to unverifiable assumptions. There are different ways of doing sensitivity analysis, in this report we used a multiple imputation based sensitivity analysis. This method was used to check if MAR assumptions are plausible. The Missing values were imputed under a plausible scenario for which the missing data are missing not at random (MNAR). If this scenario leads to a conclusion different from that of inference under MAR, then the MAR assumption is questionable (Yuan, 2014).

2.7 Software Used

The SAS statistical package version 9.4 and R version 3.0.2 were used for statistical analysis and data management respectively. All the tests were done at the 5% level of significance.

3 Results

3.1 Exploratory Data Analysis

3.1.1 Descriptive Statistics

Descriptive statistics was provided for each explanatory and response variables as shown in Table 2 and Table 3. With respect to education in the control group majority of the individuals 125 (59.12 %) completed higher education (no university), and majority of the breast cancer patients 114 (41.61 %) completed secondary school (12-18 year). In terms of Marital status from Table 2, most of the individuals 175 (81.52 %) in the control group and 216 (78.83%) in the cancer group were married. Again, 3 cancer patients did not provide their educational and marital status information in the study.

Table 2: The number (proportion) of people included in the cancer and control group across different levels of covariates.

	Education count (%)				
Group	**Missing**	**1**	**2**	**3**	**4**
C	0	16 (7.58%)	53 (25.12%)	125 (59.12%)	17 (8.06%)
S	3(1.09%)	49 (17.88%)	114(41.61%)	98 (35.77%)	10 (3.65%)
	MaritalStatus count (%)				
	Missing	**1**	**2**	**3**	**4**
C	0	172 (81.52%)	18 (8.53%)	6 (2.84%)	15 (7.11%)
S	3 (1.09%)	216 (78.83%)	23 (8.39%)	17 (6.2%)	15 (5.47%)

A summary statistics for each response and age of respondents based on their group was given in Table 3. As shown in the table, the average age in the cancer and control group was 54.07 and 45.76 years respectively. With respect to resilience, minimum score in the cancer and control group was 12 and 32 respectively. Again the maximum anxiety score observed in cancer and control group was 20 and 19 respectively. Furthermore, the maximum depression score in cancer and control group was 20 and 15 respectively. Overall, it seems that older age, higher depression, higher anxiety score in the cancer group as compared to control group. In addition, summary statistics for each intrapersonal covariates showed that the average baseline resilience score for cancer and control group was 68.85 and 68.03, respectively which seems slightly different and the average life event scale, Neuroticism

for cancer and control group was 1.3, 4.96 and 1.15, 4.22 respectively. Again on average, the life orientation scale (optimism) was 19.87 for cancer and 20.71 for control group.

Table 3: Summary statistics for each response and age in both groups

Group	Variable	Minimum	Median	Maximum	Mean	STD
C	Age	22	47	71	45.76	12.34
	resilience	32	68	100	68.38	11.45
	anxiety	0	7	19	6.93	3.45
	depression	0	6	15	5.85	3.65
S	Age	24	54	78	54.07	9.93
	resilience	12	67	100	66.93	15.72
	anxiety	0	8	20	8.07	3.93
	depression	0	7	20	6.13	3.95

3.1.2 Individual Profile

The individual profile plot presented in Figure 1, Figure 2 and Figure 3 is helpful to see how an individuals resilience, anxiety and depression score evolves over time in breast cancer patients and in the control group. From the plot, it seems that high between subject and within subject variability in both cancer and control group for each response. As it shown in Figure 1, for resilience score the within subject variability in the cancer group (left panel) seems a bit higher as compare to the control (right panel) group. However, for the anxiety Figure 2, the between subject variability seems almost the same in both groups. Lastly for the depression, the between subject variability seems a bit higher in the breast cancer patients compare to control group. Furthermore, there seems a multinomial pattern in which both the strength and the direction of change are varying between people in each plot. The direction of change showed highly decreasing profile for some people, increasing profile for some other people and stable profile for others across time. This might be at least an indication that the population is not perfectly homogeneous whereby all individuals in a given sample are expected to change in the same direction across time with only the degree of change varying between people.

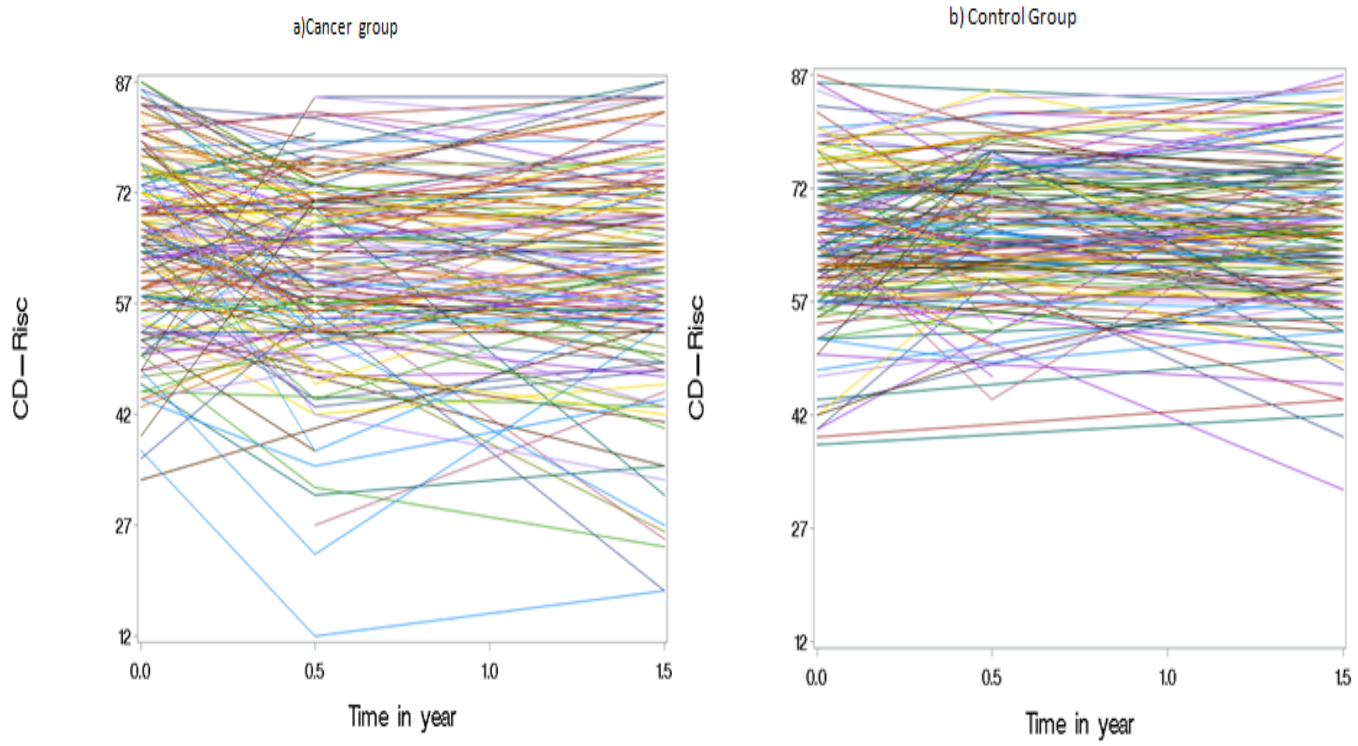

Figure 1: resilience scale individual profiles for cancer and control group

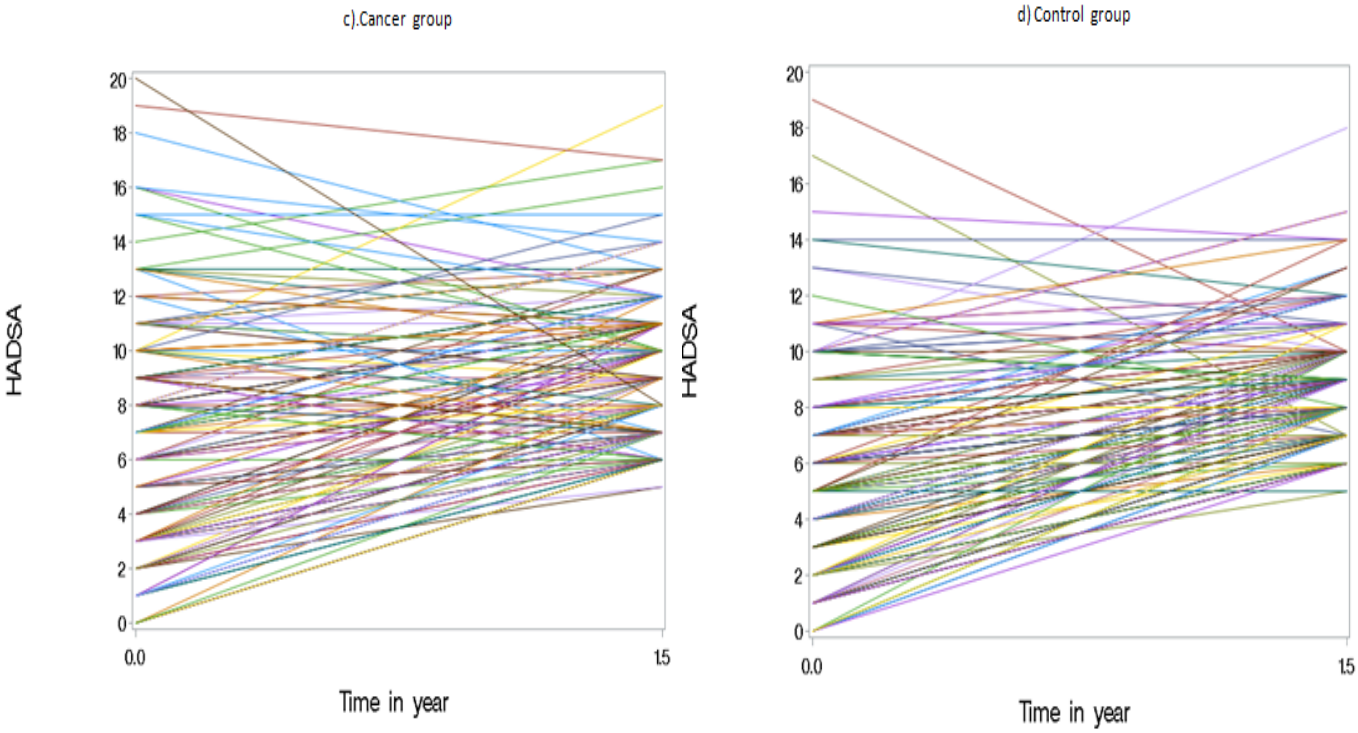

Figure 2: anxiety individual profiles for cancer and control group

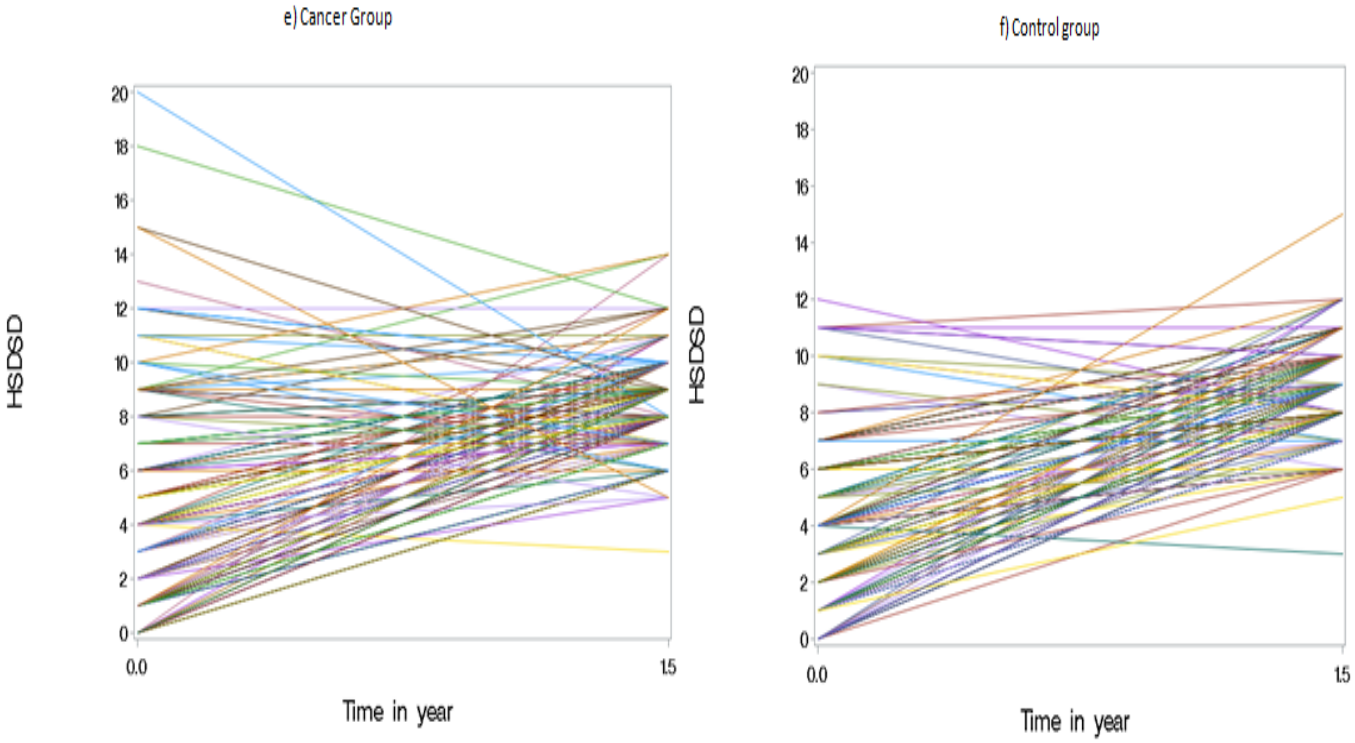

Figure 3: depression individual profiles for cancer and control group

3.1.3 Average Evolution

To have an idea on how the two groups evolved over time and specify a plausible fixed effect structure for linear mixed model, graphical exploration for average evolution of resilience, anxiety and depression score for breast cancer and control groups were presented. From Figure 4, the average resilience score seems higher at baseline and decrease over time for the breast cancer group. But, this was not the case for the control group as it seems increasing first and start declining after the second time point. Again from Figure 5, the average anxiety score was increasing over time in both groups, but this was higher for cancer group in all the time. And lastly, the depression score of cancer patients was higher at the beginning as compare to control group, but in the end it seems to coincide.

3.1.4 Exploring Missingness

Missing data reduce the representativeness of the sample and can, therefore, distort inferences about the population. In this data set, the missingness are explored for both groups in each

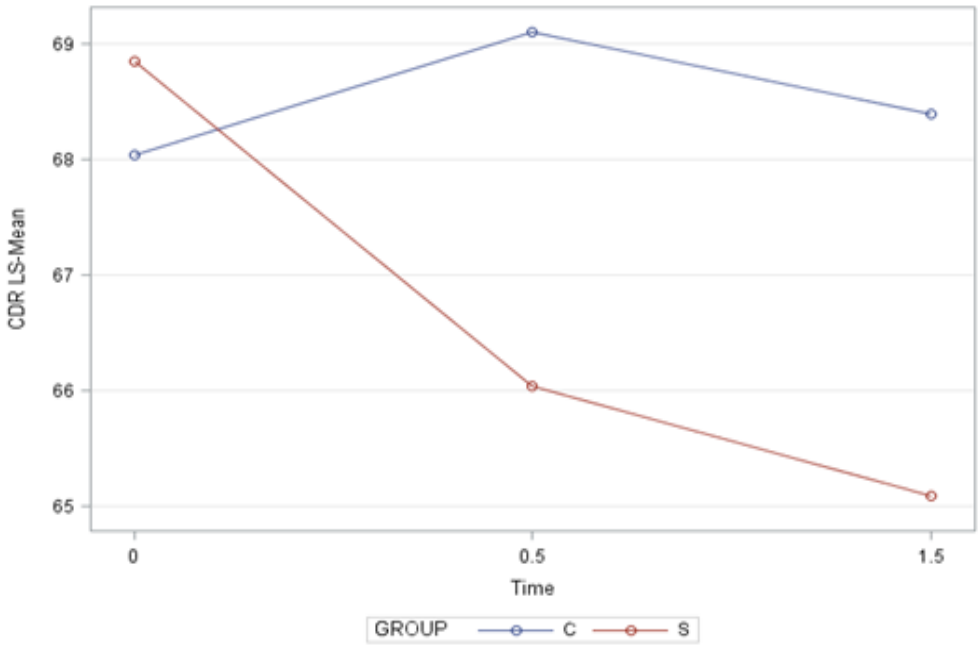

Figure 4: Average evolution over time in years for resilience score in breast cancer and control groups

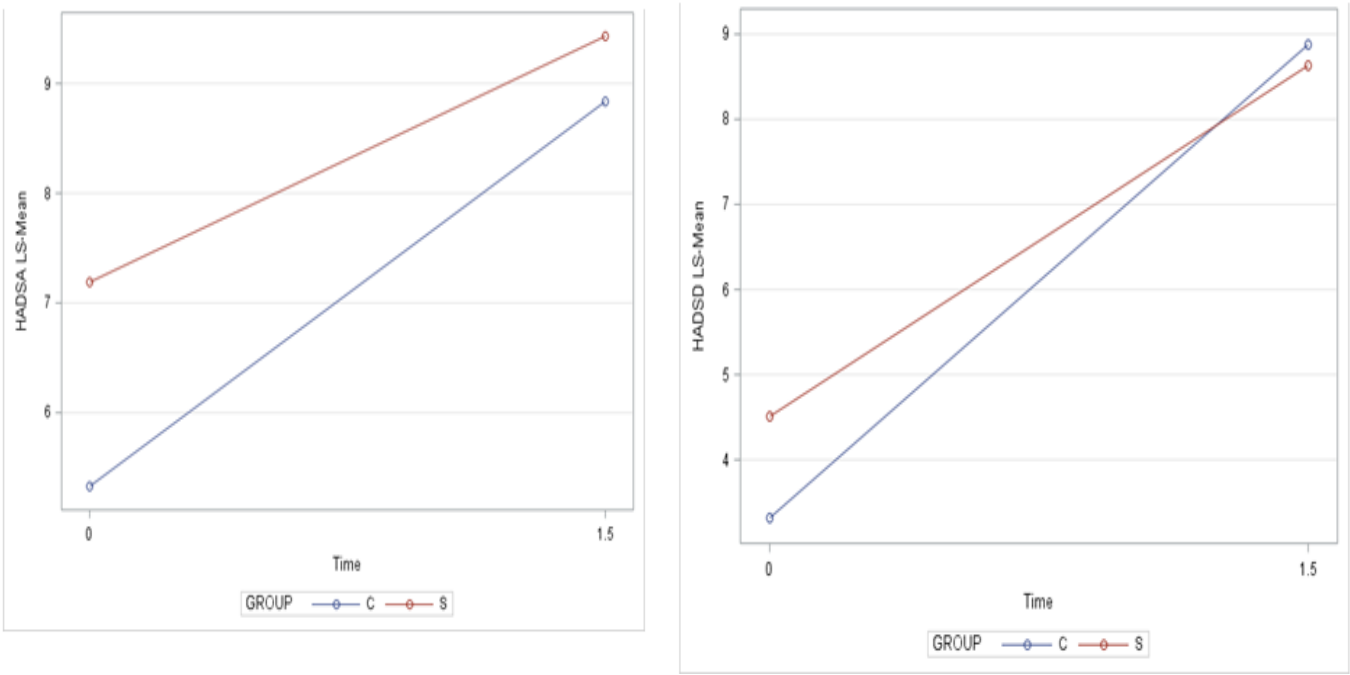

Figure 5: Average evolution over time in years for anxiety (left panel) and depression (right panel) in breast cancer and control groups

response. From Table 4, it was clear that 80 individuals out of 211 (i.e. 37.91%) had complete measurements in the control group and 87 (41.23%) individuals are missed in the second time point and the remaining 44 (20.86%) individuals are missing at first or last time period. Furthermore, in the cancer group, 143 (52.19%) patents out of 274 completed the study successfully. Again from

Table 4: Missingness characteristics for resilience score in control and cancer group

measurment at			resilience count(%)	
T1	T2	T3	Control	Cancer
O	O	O	80 (37.91%)	143 (52.19%)
O	O	M	15 (7.11%)	40 (14.6%)
O	M	M	18(8.53%)	55(20.07%)
M	M	M	2 (0.93%)	8(2.92%)
O	M	O	87(41.23%)	12(4.38%)
M	O	O	1(0.47%)	8 (2.92%)
M	O	M	1 (0.47%)	7(2.55%)
M	M	O	7 (3.32%)	1 (0.36%)
Total			211	274

O= Observed Response, M= Missed Response

Table 5 for anxiety, it can be seen that for control group, 172 out of 211 (i.e. 81.52%) individuals completed the study successfully and for cancer group, 152 (55.47%) patients completed the study. Lastly, in the depression score from Table 5, 172 (81.52%) individuals had complete measurements and for cancer group, 156 (56.93%) patents had full measurements. The percentage of missingness for resilience in the control group was high because approximately half of the control group have not had a measurement at the second time. However, the missed data percentage for the anxiety and depression score in cancer and control groups was less than 50%.

Table 5: Missingness characteristics for anxiety and depression level in control and cancer group

measurment at		anxiety count(%)		depression count(%)	
T1	T3	Control	Cancer	Control	Cancer
O	O	172(81.52%)	152(55.47%)	172(81.52%)	156(56.93%)
O	M	34(16.11%)	102(37.23%)	35 (16.59%)	100(36.5%)
M	O	1(0.47%)	12(4.38%)	1(0.47)	9 (3.28%)
M	M	4 (1.9%)	8(2.93%)	3(1.42)	9 (3.28%)
Total		211	274	211	274

O= Observed Response, M= Missed Response

3.2 Logistic Regression Model

At first, a kind of summary statistics was performed by taking difference of measurements recorded at two time periods (T3) and (T1) and creating two classes based on the sign of the difference (T3-T1) i.e. Δ_i. Thus, for resilience score a negative Δ_i including zero as decreasing class, positive Δ_i as improving class. Then a logistic regression was applied to these two classes considering the decreasing class as reference. Parameter estimates and standard error for all responses are presented in Table 6. From Table 6, baseline resilience score was the only significant covariate

Table 6: Parameter estimates with their standard errors based on logistic regression for each response

	resilience		**anxiety**		**depression**	
	cancer	**control**	**cancer**	**control**	**cancer**	**control**
Parameter	**Est.(SE)**	**Est.(SE)**	**Est.(SE)**	**Est.(SE)**	**Est.(SE)**	**Est.(SE)**
Intercept	2.68(2.22)	2.39(1.91)	0.46(2.76)	3.95(2.91)	-3.88(3.08)	2.17(3.94)
Age	-0.01(0.02)	0.01(0.02)	-0.03(0.02)	-0.02(0.03)	0.02(0.03)	-0.01(0.03)
ms	0.19(0.48)	0.89(0.48)	0.84(0.59)	0.44(0.75)	-0.16(0.71)	0.39(1.26)
edu	0.36(0.40)	-0.73(0.44)	0.12(0.49)	0.32(0.73)	-0.13(0.56)	-1.88(0.97)
T1CDR	**-0.07(0.02)***	**-0.07(0.02)***	0.04(0.02)	-0.03(0.03)	0.04(0.02)	-0.02(0.04)
LES	-0.10(0.12)	0.22(0.13)	0.14(0.15)	-0.15(0.18)	-0.06(0.18)	0.33(0.40)
LOT	0.06(0.05)	0.08(0.05)	0.07(0.06)	0.15(0.08)	0.11(0.07)	0.12(0.11)
EPQ	-0.05(0.07)	-0.13(0.08)	-0.04(0.08)	-0.21(0.12)	-0.16(0.09	-0.17(0.16)
posaff	0.09(0.05)	0.04(0.06)	-0.03(0.06	0.04(0.10)	0.12(0.07)	0.19(0.13)
negaff	-0.04(0.04)	0.04(0.06)	**-0.15(0.05)***	**-0.18(0.08)***	-0.04(0.05)	-0.16(0.10)

* are significant at 5% level

with the improving resilience class in both cancer and control group. Again, for the anxiety and depression score, the same procedure was done. And, only negative emotion was significant covariate in both groups for the increasing anxiety score and there were no any significant covariates with the probability of increasing depression score in both groups. Nevertheless, as mentioned, the main problem of doing summary statistics, i.e. very crude, rough cut of points and information losses, we did this just to get insight before applying appropriate method.

3.3 Linear Mixed Model (LMM)

Considering resilience score, anxiety and depression level as continuous variable, a linear mixed model which takes into account the correlation between measurements at different time points within each subject were fitted. In order to identify a model that fits the data best, we started our model building process with initial model that includes all the independent variables and their pairwise interactions as a fixed effect and random intercept and random slope for time as random effects.

In order to reduce the model in seek of parsimonious model, the backward selection approach was used basing on likelihood ratio tests (LRT). The restricted maximum likelihood (REML) technique was used to check for possible reduction of covariance structure. The REML produces unbiased estimating equation for the variance parameters. On the other hand, MLE technique was used for reducing mean structure. We have tested if any possible reduction of random effects is possible and finally continue for the reduction of mean structure. The model assuming the unstructured covariances structure was fitted. Based on mixture of chi-squire $\chi^2_{1:2} = 13.1$ and p-value=0.001, the random slope is important in the model for resilience score. Next, age, Marital status, education were removed respectively from the initial model for resilience score. Again the same procedure was done for anxiety and depression score and based on mixture of chi-squire $\chi^2_{1:2} = 12$ and p-value=0.002, the random slope should be kept in the model for anxiety score. Again for the depression score the need for random slope $\chi^2_{1:2} = 65.3$ and p-value=0.0001, the random slope is important for the model in depression score. Moreover, marital status and education, were not important and removed respectively from the initial model for anxiety and depression score as well. Thus, the final model for each response can be written as follows:

$$\begin{aligned} R_{ij} = \beta_0 + b_{1i} + \beta_1 Group_i + \beta_2 time_{ij} + \beta_3 LOT_i + \beta_4 EPQ_i + \beta_5 posaff_i + \beta_6 negaff_i \\ + \beta_7 LES_i + (\beta_8 Group_i + b_{2i}) \times time_{ij} + \varepsilon_{ij} \end{aligned} \tag{6}$$

$$\begin{aligned} A_{ij} = \beta_0 + b_{1i} + \beta_1 age_i + \beta_2 Group_i + \beta_3 time_{ij} + \beta_4 LOT_i + \beta_5 EPQ_i + \beta_6 posaff_i \\ + \beta_7 negaff_i + \beta_8 LES_i + (\beta_9 Group_i + b_{2i}) \times time_{ij} + \varepsilon_{ij} \end{aligned} \tag{7}$$

$$D_{ij} = \beta_0 + b_{1i} + \beta_1 age_i + \beta_2 Group_i + \beta_3 time_{ij} + \beta_4 LOT_i + \beta_5 EPQ_i + \beta_6 posaff_i + \beta_7 negaff_i + \beta_8 LES_i + (\beta_9 Group_i + b_{2i}) \times time_{ij} + \varepsilon_{ij} \quad (8)$$

where, $Group_i$ is the respondents group whether breast cancer patients or control group at the study and $time_{ij}$ is the time at which j^{th} measurements was taken for patient i, LOT is the baseline optimism of subjects, EPQ is neuroticism of the individual, posaff is positive emotion, negaff is negative emotion and ε_{ij} is the measurement error. The results based on the final model are summarized in Table 7. Thus, Table 7 shows that life orientation scale or optimism and positive emotion were positively associated with resilience score and statistically significant baseline effect on resilience score (p-value < 0.0001). This implies that an increase in optimism or higher positive emotion resulted to improve resilience score keeping other things constant. Again there was significant difference between breast cancer and control group at baseline i.e. main effect of group was significant (p-value < 0.0001). Further, neuroticism and negative emotion was negatively related i.e. a unit increase in the neuroticism of the individual resulted to a decrease in resilience score by 0.68 fixing other variables constant. It was also clear that, through time resilience score of patients decrease and there was difference in evolution between breast cancer patients and control group i.e. the interaction between group by time was significant (p-value=0.0005). Results from Table 7 for anxiety showed that, there was no significant difference between breast cancer and control group at baseline. Again, optimism was negatively associated with the anxiety score (p-value < 0.0001). This implies that an increase in optimism resulted to decrease anxiety score keeping other things constant. Again there was significant negative effect of age, neuroticism and negative emotion with anxiety score. Moreover, the anxiety score of patients increase through time and the evolution between breast cancer and control groups was significantly different i.e. interaction with time was significant. Similarly for the depression score age, optimism, neuroticism, positive emotion and negative emotion had significant baseline effect on depression score (p-value=0.001). This implies that for example an increase in optimism resulted to decrease in depression score keeping other things constant. Furthermore, there was no difference between the two groups at baseline for depression score. However, the evolution depends on group status.

Table 7: Parameter estimates with their standard errors based on linear mixed models for resilience, anxiety, depression

	resilience	**anxiety**	**depression**
Effect	**Est.(SE)**	**Est.(SE)**	**Est.(SE)**
Intercept	42.96(3.81)*	4.31(0.98)*	3.37(0.75)*
Age		0.02(0.01)*	0.02(0.01)*
Group	8.42(1.24)*	0.65(0.35)	0.73(0.37)
Time	0.67(0.57)	2.35(0.17)*	3.68(0.19)*
LOT	0.43(0.13)*	-0.08(0.03)*	-0.04(0.02)*
EPQ	-0.68(0.20)*	0.26(0.04)*	0.05(0.03)
posaff	1.43(0.14)*	-0.06(0.03)*	-0.03(0.02)
negaff	-0.43(0.13)*	0.15(0.03)*	0.05(0.02)*
T3LES	0.17(0.33)	0.20(0.07)*	0.03(0.05)
Group*Time	-2.98(0.84)*	-0.81(0.25)*	-0.95(0.28)*
var(b_{1i})	36.33(8.47)	2.28(0.71)	3.16(0.88)
var(b_{2i})	0.46(9.96)	0.50(0.40)	0.34(0.63)
cov(b_{1i}, b_{2i})	13.45(5.63)	-0.95(0.48)	-2.59(0.52)
σ^2	63.66(6.91)	5.67($1e^{-7}$)	6.90($1e^{-7}$)

* are significant covariates at 5% level of significant

3.4 Latent Class Growth Model (LCGM)

3.4.1 Unconditional Model

First, the unconditional model was fitted for each response of interest without the inclusion of the possible covariates. Thus, for the resilience score of individuals, we compared a model with one to five trajectories in both groups as recommended for model testing procedure. From Table 8, the change in Bayesian information criteria 2*(ΔBIC) showed improvement until three-class solutions and adding a fourth trajectory did not improve the fit of the model in both groups. Accordingly, we selected three-trajectories for breast cancer and control group as optimal choose. As can be seen in Figure 6, the three-class solution in breast cancer group across the 1.5 year follow up identified three distinct trajectories of resilience score and the average posterior membership probabilities of belonging to a trajectory were 0.82, 0.83, and 0.89 for the three trajectories, respectively. As shown also in Table 9, the intercept and slope were significant only for the first and second classes. Thus, the first class, capturing 18.5% of the sample, with low initial resilience score described a trajectory of decreasing dramatically. The second and more common class capturing 52.2% of the

Table 8: Fit statistics for one to five LCGM for resilience score in both groups without including covariates

	Breast cancer		Control group	
#Trajectories	BIC	2*(ΔBIC)	BIC	2*(ΔBIC)
1	-2532.79		-1828.05	
2	-2461.36	**142.86**	-1783.1	**89.9**
3	-2439.17	**44.38**	-1765.9	**34.4**
4	-2439.40	-.46	-1768.36	-4.92
5	-2450.68	-22.56	-1774.89	-13.06

sample, with medium initial resilience score described a trajectory of slight decreasing over time. The third class which captured 29.3% of the sample, with high resilience initial score described a trajectory of stable or constant over time. None of the quadratic functions were significant in any of the models. Furthermore, as shown in Figure 6, the three-class solution in the control group across the 1.5 year follow up identified three distinct trajectories of resilience score and the average posterior membership probabilities of belonging to a trajectory were 0.85, 0.90 and 0.92 for the three trajectories, respectively. As can be seen from Table 9, the slope for all the three trajectories were not significant. Thus, the first-class, capturing 13.7% of the sample, with low initial resilience score described a trajectory of stable low resilience score. The second and more common class capturing 64.4% of the sample, with medium initial resilience score described a trajectory of medium stable or constant over time. The third class which captured 21.9% of the sample, with high initial resilience score described a trajectory of high stable resilience. In general, we found the same number of trajectories in both the cancer and control group but, the low to medium resilient breast cancer patients decrease their resilience over time while all the trajectories in the control group not changed or constant over time and patient's who have had high resilience score in cancer group showed the same evolution as healthy group. Similarly for the anxiety score of an individual, based on Table 10, we selected three-classes for breast cancer group as well as control group as optimal choose. As shown in Figure 7, the three-class solution in breast cancer group across the 1.5 year follow up identified three distinct trajectories of anxiety score and the average posterior membership probabilities of belonging to a trajectory were 0.90, 0.81, and 0.87 for the three trajectories, respectively. And from Table 11, the slope was significant for the first and second classes but not for the third class in

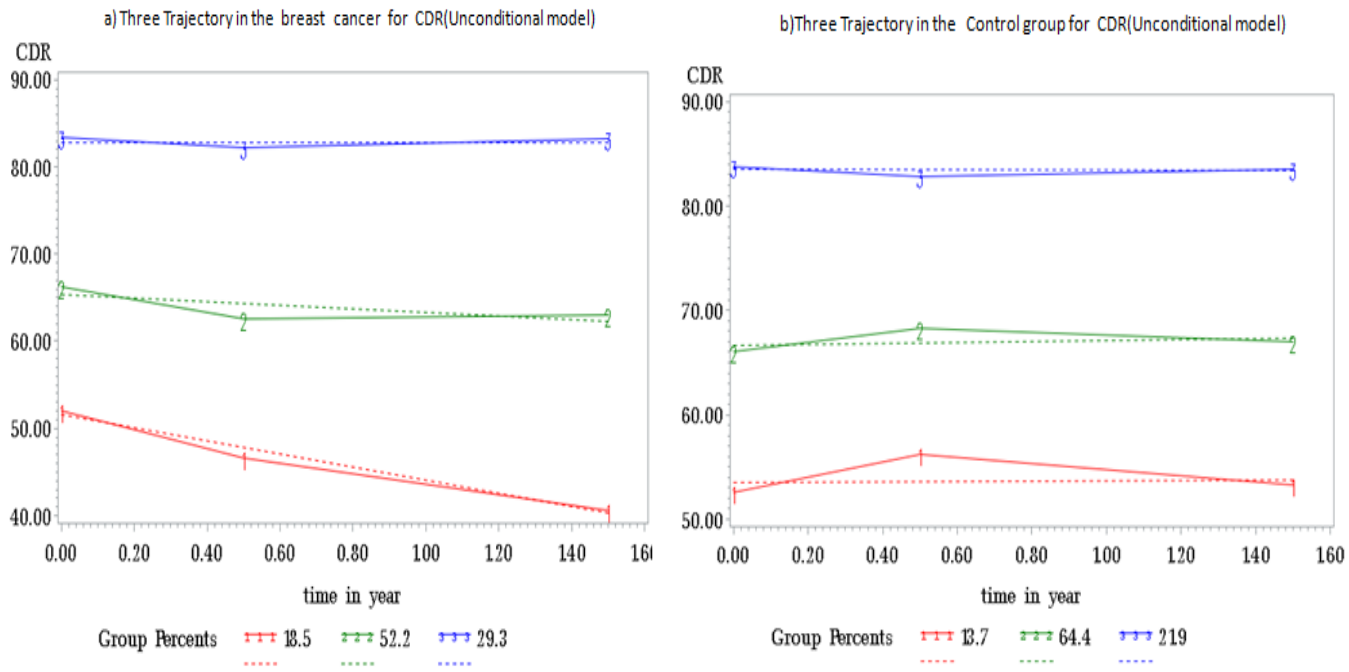

Figure 6: Three trajectory LCGM in a Breast cancer group and in the Control group for resilience scale Unconditional model

Table 9: Parameter estimates for resilience score in each group

resilience Scale	**Intercept**		**Slope**	
Trajectory of	*Est(SE)*	*p-value*	*Est(SE)*	*p-value*
Cancer group				
decrease dramatically	53.53(1.69)	< 0.0001	-7.49(1.99)	0.0002
slight decrease	65.29(1.40)	< 0.0001	-2.01(1.02)	0.04
high stable	83.01(1.32)	< 0.0001	-0.08(1.32)	0.95
Control group				
low stable	53.49(1.96)	< 0.0001	0.18(1.55)	0.9
medium stable	66.60(0.85)	< 0.0001	0.48(0.66)	0.4
high stable	83.64(1.40)	< 0.0001	-0.13(1.20)	0.9

Table 10: Fit statistics for one to four groups using unconditional LCGM for anxiety and depression in both groups

anxiety score				
	Breast cancer		Control group	
# of Trajectories	BIC	2*(ΔBIC)	BIC	2*(ΔBIC)
1	-1142.74		-956.00	
2	-1116.17	**53.14**	-921.46	**69.08**
3	-1107.10	**18.14**	-911.97	**18.98**
4	-1109.99	-5.78	-913.63	-3.32
depression score				
	Breast cancer		Control group	
#of Trajectories	BIC	2*(ΔBIC)	BIC	2*(ΔBIC)
1	-1081.22		-863.46	
2	-1035.12	**92.2**	-833.54	**61.2**
3	-1019.64	**30.96**	-835.82	-4.56
4	-1021.98	-4.68		

both groups. Thus, the first class, capturing 41.7% of the sample, with very low initial score of anxiety described a trajectory of dramatically increasing anxiety. The second and more common class capturing 47.5% of the sample, with medium initial anxiety score described a trajectory of slightly increasing anxiety over time. The third and small in size class which captured 10.8% of the sample, with high anxiety initial score described a trajectory of stable high anxiety. In the same

Table 11: Parameter estimates for anxiety score in each group Unconditional model

anixiety level	**Intercept**		**Slope**	
Trajectory of	*Est(SE)*	*p-value*	*Est(SE)*	*p-value*
Cancer group				
dramatically increasing	3.26 (0.36)	<0.0001	3.01(0.28)	<0.0001
slightly increasing	8.79(0.47)	<0.0001	0.89(0.28)	0.0018
stable high	14.97(0.88)	<0.0001	-1.34(0.71)	0.06
control group				
dramatically increasing	2.48(0.29)	<0.0001	3.48(0.22)	<0.0001
slightly increasing	6.77(0.42)	<0.0001	1.70(0.25)	<0.0001
stable high	11.92(0.56)	<0.0001	-0.05(0.41)	0.9

Figure 7, the three-class solution in the control group identified three distinct trajectories of anxiety score like in the cancer group and the average posterior membership probabilities of belonging to a

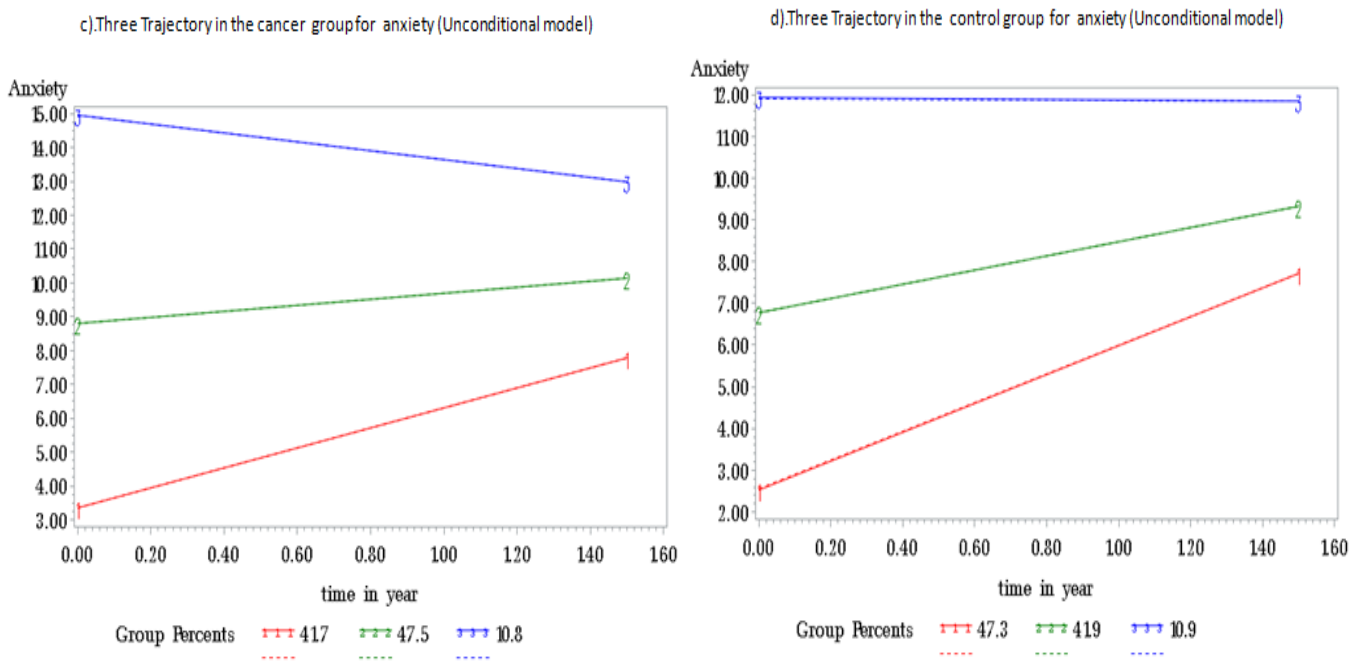

Figure 7: Three trajectory latent class growth model in a breast cancer group and in the control group for anxiety score unconditional model

trajectory were 0.89, 0.85, and 0.86, respectively. Thus, the first and more common class, capturing 47.3% of the sample, with very low initial score of anxiety described a trajectory of dramatically increasing anxiety. The second class capturing 41.9% of the sample, with medium initial anxiety score described a trajectory of slightly increasing anxiety over time. The third and small in size class which captured 10.9% of the sample, with high initial anxiety score described a trajectory of stable high anxiety.

Again for the depression score based on Table 10, three-trajectory model was retained as most parsimonious model for breast cancer group and two-trajectory model as optimal choose for the control group in the unconditional model. As shown in Figure 8, the three-class solution in breast cancer group identified three distinct trajectories of depression score and the average posterior membership probabilities of belonging to a trajectory were 0.91, 0.93, and 0.92, respectively. As shown also in Table 12, the slope was significant for all the classes in both groups except the third class in the cancer group. Thus, the first and most common class, capturing 61.7% of the sample, with very low initial score of depression described a trajectory of dramatically increasing in depression. The second class capturing 33.7% of the sample, with medium initial depression score

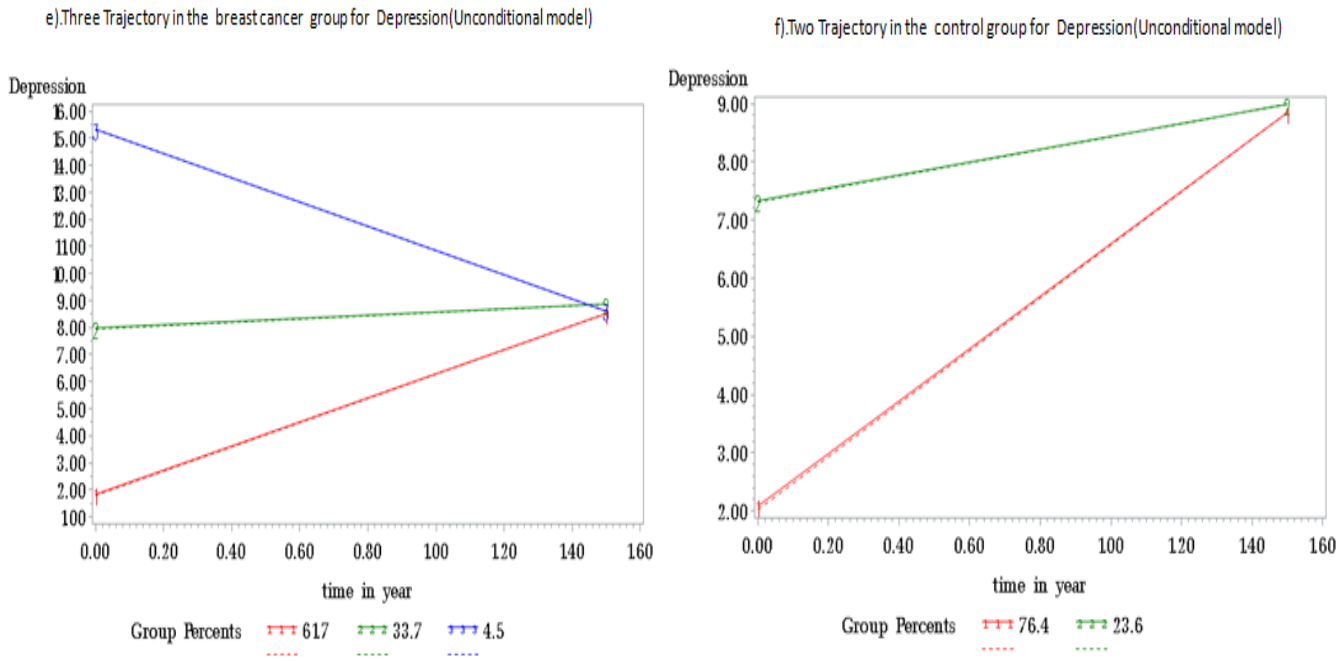

Figure 8: Three trajectory latent class growth model in a Breast cancer group and Two trajectory in control group for depression

described a trajectory of slightly increasing in depression over time. The third but very small in size class which captured 4.5% of the sample, with highly depressed initial score described a trajectory of stable high depression.

In the same Figure 8, the two-class solution in the control group across the follow up identified two distinct trajectories of depression score and the average posterior membership probabilities of belonging to a trajectory were 0.95 and 0.88 for the two trajectories, respectively. Hence, the first and highest class in size consists of 76.4% of the sample, with very low initial score of depression described a trajectory of dramatically increasing or becoming worse in depression. The second class capturing 23.6% of the sample, with medium initial depression score described a trajectory of slightly increasing over time.

3.4.2 Conditional Model

To see if there is an improvement in model fit, all the possible covariates that allow model convergence were included in the model. Before starting the procedure, it is important here to mention that, PROC TRAJ does not allow for distinctions between continuous and categorical predictors.

Table 12: Parameter estimates for depression score in each group Unconditional model

depression score	**Intercept**		**Slope**	
Trajectory of	*Est(SE)*	*p-value*	*Est(SE)*	*p-value*
Cancer group				
dramatically increasing	1.56(0.21)	<0.0001	4.63(0.19)	<0.0001
slightly increasing	7.95(0.34)	<0.0001	0.62(0.28)	0.03
stable high	5.33(0.92)	<0.0001	-4.47(0.96)	0.08
control group				
dramatically increasing	1.87(0.21)	<0.0001	4.65(0.18)	<0.0001
slightly increasing	7.32(0.44)	<0.0001	1.11(0.36)	0.0023

Therefore, it was necessary to create dummy-coded variables for education and marital status. Furthermore, some of the categories of these variables contain very few observation as shown in the exploratory section Table 2. Thus, dummy variable was created from the given categories by collapsing categories with small number of observations i.e. marital status, married vs others, for education primary+secondary vs higher education. This helps us to have enough observation and get power in the analysis. However, due to converges problem in resilience and anxiety score some covariates were excluded from the conditional model. Next, best model was identified by comparing models until the optimal number of trajectories. For the resilience score in breast cancer and control group, we compared a model with one to four trajectories. Results in Table 13, showed that there was no evidence to go beyond the three classes. Accordingly, three classes were chosen as optimal number of trajectories in both cancer and control group and this was consistent with the unconditional model where we found the same number of trajectories. As shown in Figure 9, the structure of the three-class conditional (with covariate) solution was exactly similar to the three-class unconditional model in both groups (i.e. Figure 9 has the same interpretation as in the unconditional model). The only slight change in the inclusion of covariate for breast cancer group were the proportion of the first class decrease to 15.6% from 18.5% and the proportion of the second class decrease to 51.6% from 52.2% and the proportion of the third class increase to 32.7% from 29.2%. The average posterior membership probabilities of belonging to a trajectory were 0.87, 0.90 and 0.91 for the three trajectories, respectively. This very high probability indicate that the allocation was good. Again, the slight change in the inclusion of covariate for healthy group were the proportion of the first class decrease to 8.9% from 13.7% and the proportion of the second class increase to 66.6%

Table 13: Fit indexes for one to four conditional LCGM for resilience score in breast cancer and control group.

	Breast cancer		Control group	
# of Trajectory	BIC	$2*(\Delta BIC)$	BIC	$2*(\Delta BIC)$
1	-2552.37		-1826.73	
2	-2009.4	1085.94	-1732.34	188.78
3	-2002.47	**13.86**	-1724.56	**15.56**
4	-2031.83	-58.72	-1729.06	-9

Note BIC = Bayesian Information Criterion.

from 64.4% and the proportion of the third class increase to 24.6% from 21.9%. And the average posterior membership probabilities of belonging to a trajectory were 0.95, 0.96, and 0.93 for the three trajectories, respectively. Parameter estimates were also given in Table 14 and shows the same conclusion as we found in the unconditional model.

Table 14: Parameter estimates for resilience score in each group based on conditional LCGM

resilience Scale	**Intercept**		**Slope**	
Trajectory of	*Est(SE)*	*p-value*	*Est(SE)*	*p-value*
Cancer group(n=208)				
decrease dramatically	51.19(2.04)	<0.0001	-7.42(2.48)	0.003
slight decrease	65.03(1.33)	<0.0001	-2.44(1.04)	0.03
high stable	82.71(1.27)	<0.0001	-0.56(1.34)	0.67
σ	9.73(0.36)			
Control group(n=205)				
low stable	49.83(2.21)	<0.0001	2.23(1.81)	0.22
medium stable	65.92(0.73)	<0.0001	0.19(0.69)	0.78
high stable	82.52(1.55)	<0.0001	-0.13(1.20)	0.90
σ	7.71(0.31)			

Further, in the conditional model for anxiety score, we compared a model with one to three trajectories for breast cancer and control group. Results in Table 16, showed that there was no evidence to go beyond the second class. Accordingly, two classes were chosen as optimal number of trajectories in both cancer and healthy group. As shown in Figure 10, the two-class solution in breast cancer group with the inclusion of covariates across the 1.5 year follow up identified two distinct trajectories of anxiety score and the average posterior membership probabilities of belonging to a

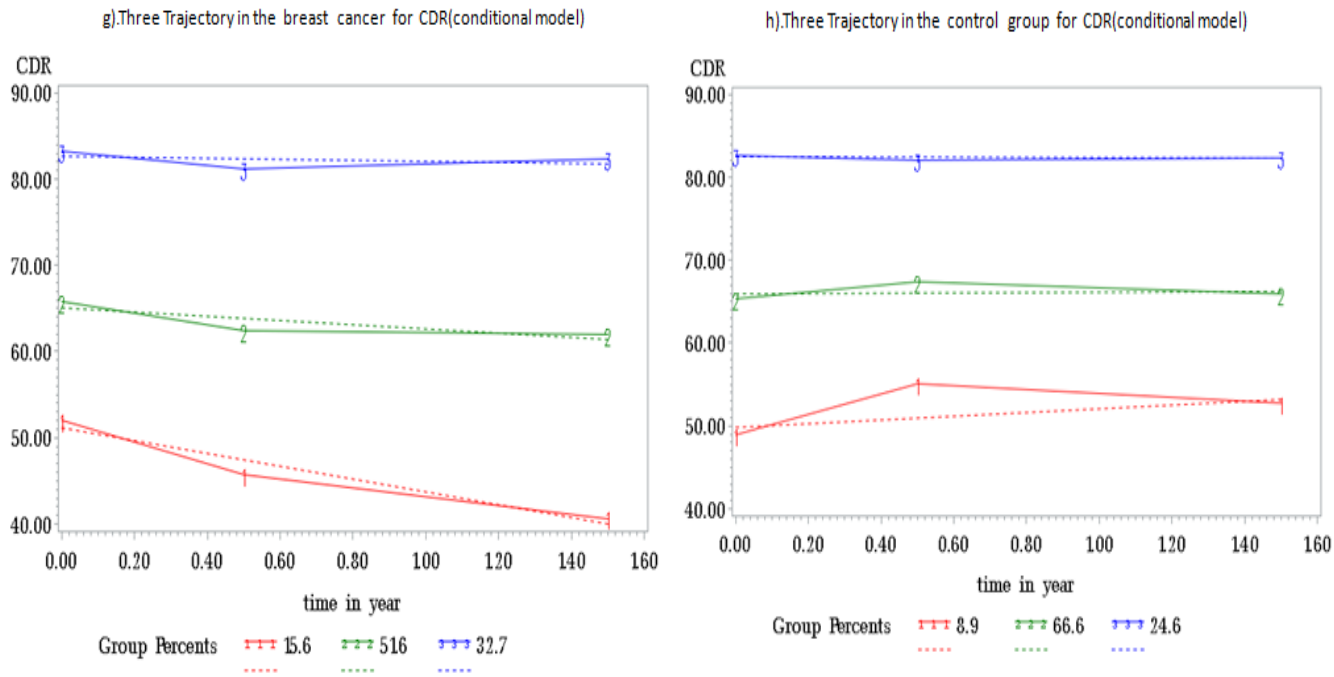

Figure 9: Three trajectory of Conditional LCGM in cancer group and in the control group for resilience score

trajectory were 0.97 and 0.99 for the two trajectories, respectively. Thus, the first class, contain 62.8% of the sample, with low initial score of anxiety described a trajectory of increasing anxiety score over time. Results from Table 15, also showed that the slope was significant for this class. The second class capturing 37.2% of the sample, with high anxiety initial score described a trajectory of high stable anxiety and non-significant slope for change across time. In the same Figure 10, the two-class solution in the control group across the 1.5 year follow up identified two distinct trajectories of anxiety score like in the cancer group for the conditional model and the average posterior membership probabilities of belonging to a trajectory were 0.97 and 0.96 for the two trajectories, respectively. Thus, the first class, capturing 56.0% of the sample, with low initial score of anxiety described a trajectory of highly increasing anxiety. And from Table 15, the slope was significant for this class. The second class capturing 44.0% of the sample, with high initial anxiety score described a trajectory of slightly increasing anxiety over time i.e. the slope was also statistically significant. In general, the third and small in size class which was observed in the unconditional model and captured around 10% of the sample was distributed to the two class when we include covariates. Hence, with respect to anxiety score in the conditional model, we found two trajectories for both

breast cancer and control groups.

Table 15: Parameter estimates for anxiety score in each group for conditional model

anxiety level	**Intercept**		**Slope**	
Trajectory of	*Est(SE)*	*p-value*	*Est(SE)*	*p-value*
Cancer group(n=211)				
increasing anxiety	4.46(0.28)	<0.0001	2.66(0.28)	<0.0001
stable high	11.41(0.38)	<0.0001	-0.30(0.36)	0.40
σ	2.88(0.12)			
control group (n=196)				
highly increasing	2.94(0.25)	<0.0001	3.23(0.22)	<0.0001
slightly increasing	8.33(0.27)	<0.0001	1.13(0.24)	<0.0001
σ	2.16(0.09)			

Table 16: Fit statistics for one to three groups using conditional LCGM for anxiety and depression in both groups

anxiety score				
	Breast cancer		Control group	
#of trajectory	BIC	$2*(\Delta BIC)$	BIC	$2*(\Delta BIC)$
1	-1151.99		-956	
2	-897.21	509.56	-840.37	231.26
3	-1027.31	-260.2	-970.97	-261.2
depression score				
	Breast cancer		Control group	
#of trajectory	BIC	2log(B10)	BIC	2log(B10)
1	-1090.62		-863.53	
2	-764.53	**652.18**	-768.31	**190.44**
3	-910.25	-291.44	-884.73	-232.84

Note: BIC = Bayesian information criterion.

The conditional model for depression score was also fitted in a similar procedure with the above steps and results from Table 16 indicate that no evidence or improvement by adding extra class to the two class shown. As a result, two classes were chosen as optimal number of trajectory in both cancer and healthy group. As shown also in Figure 11, the two-class solution in breast cancer group with the inclusion of covariates across the follow up identified two distinct trajectories of depression score and the average posterior membership probabilities of belonging to a trajectory were 0.97 and 0.97

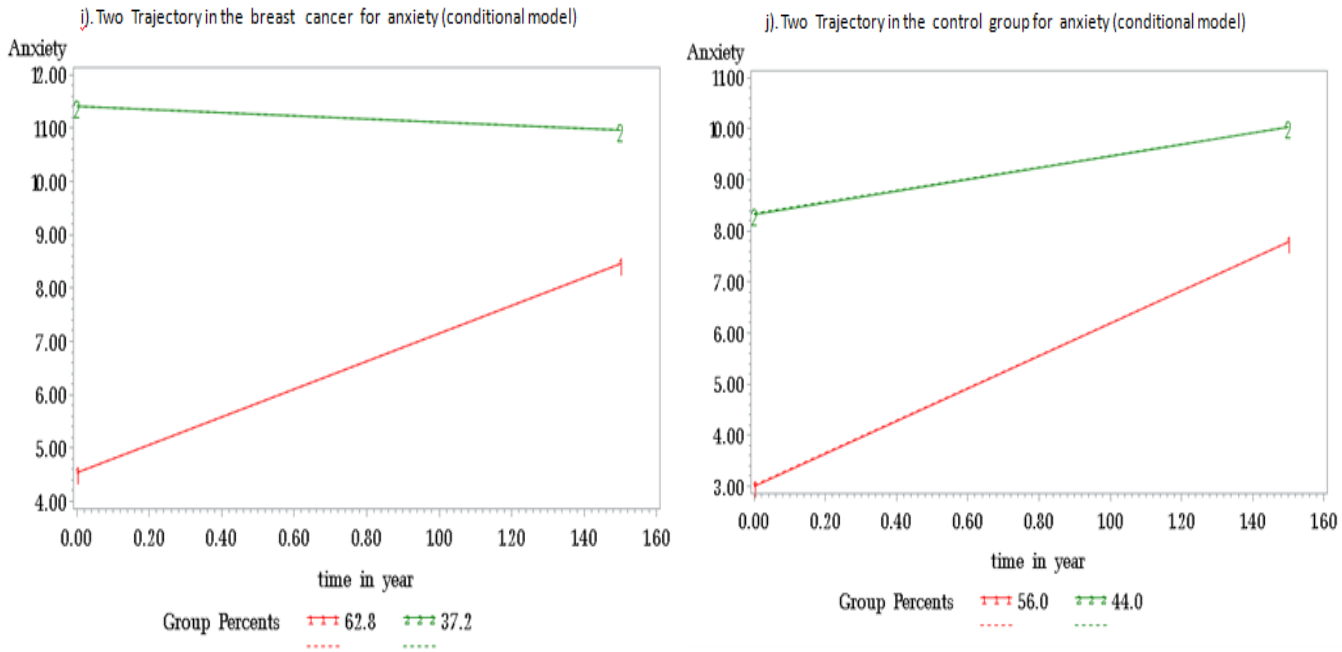

Figure 10: Conditional two trajectory latent class growth model in a breast cancer group and in the control group for anxiety score

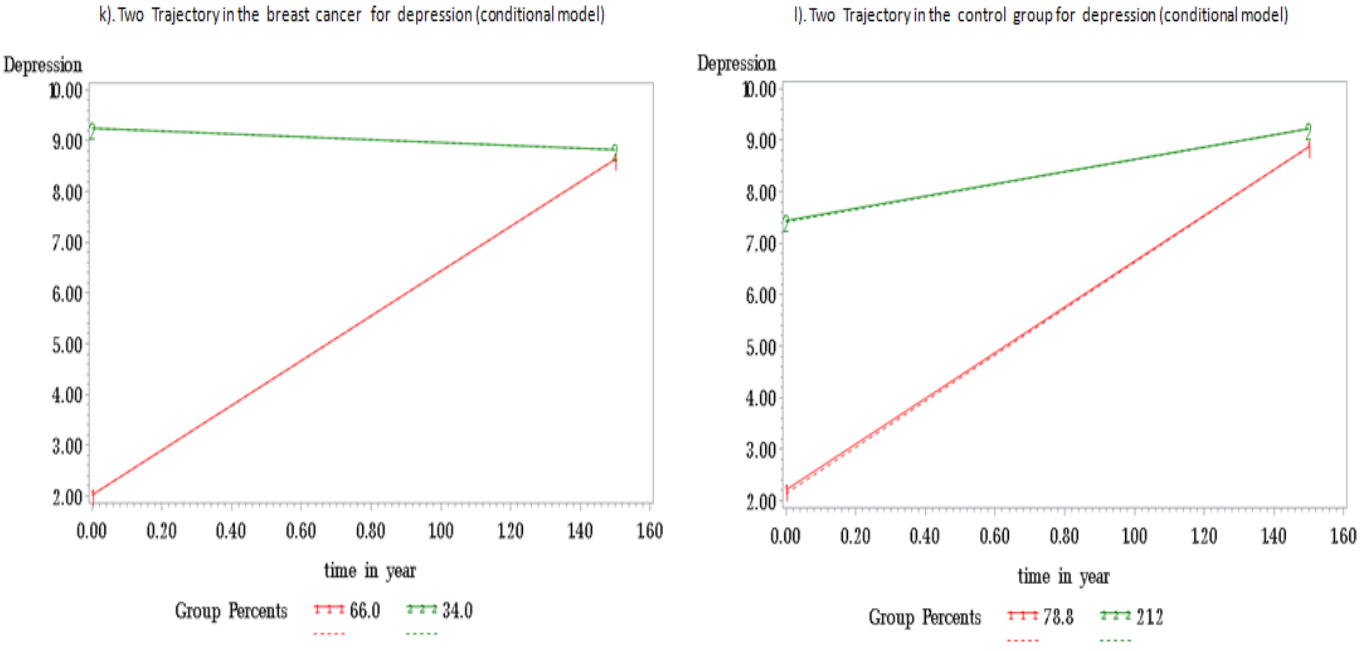

Figure 11: Conditional two trajectory latent class growth model in a breast cancer group and in the control group for depression score

for the two trajectories, respectively. Again from Table 17, we can see that the slope was significant for this class. Thus, the first class, contain 66.0% of the sample, with low initial score of depression described a trajectory of increasing depression score over time. The second class capturing 34.0% of the sample, with high initial score described a trajectory of high stable depression and statistically non-significant slope. In the same Figure 11, the two-class solution in the control group across the 1.5 year follow up identified two distinct trajectories of depression score. And the average posterior membership probabilities of belonging to a trajectory were 0.99 and 0.98 for the two trajectories, respectively. And from Table 17, the slope was significant for both classes. Thus, the first class, capturing 78.8% of the sample, with low initial score of depression described a trajectory of highly increasing depression. The second class capturing 21.2% of the sample, with high initial depression score described a trajectory of slightly increasing depression over time. In general, with respect to depression score in the conditional model, we found two trajectories for both breast cancer and control patients.

Table 17: Parameter estimates for depression score in each group conditional model

Depression level	**Intercept**		**Slope**	
Trajectory of	*Est(SE)*	*p-value*	*Est(SE)*	*p-value*
Cancer group(n=190)				
highly increasing	1.66(0.26)	<0.0001	4.65(0.25)	<0.0001
stable high	9.25(0.38)	<0.0001	-0.29(0.37)	0.43
σ	2.46(0.11)			
control group (n=197)				
highly increasing	1.98(0.16)	<0.0001	4.58(0.16)	<0.0001
slightly increasing	7.43(0.32)	<0.0001	1.19(0.31)	0.0031
σ	1.95(0.08)			

3.5 Multinomial Logistic Regression Model

Using stepwise model selection procedure for the resilience score in the breast cancer group we obtained a final model which consists of main effects of baseline resilience, life event scale, optimism, neuroticism, negative emotion and positive emotion as best candidates among the available common predictors. A multinomial logistic regression result for predictors of class membership in the resilience score for cancer group was given in Table 18. It can be seen that baseline resilience

score, optimism, neuroticism and positive emotion were risk factors that are identified as significant predictors and significantly related to the grouping or class memberships in the breast cancer patients. Hence, group one had significantly lower baseline resilience score than group three and group two had significantly lower baseline resilience score than groups three. These results indicate that trajectories containing patients with dramatically decreasing resilience score (G1), were made up of patients with significantly lower resilience scores than trajectories of high stable resilience score (G3). Furthermore, trajectories containing patients with slightly decreasing resilience score (G2), were made up of patients with significantly lower resilience scores as compare to trajectories containing patients with high stable resilience score (G3). In-short patients with high stable trajectory (G3) have had significantly higher baseline resilience score than class one and class two. From Table 18, group one had significantly lower baseline optimism score than group three and there was no difference between group two and group three with respect to optimism score. These results indicate that trajectories containing patients with dramatically decreasing resilience score (G1), were made up of patients with significantly lower optimism scores than trajectories of high stable resilience score (G3). Conversely trajectory of stable high resilience score were significantly made up of patients with high Optimism score. But, there was no significant different between the second trajectory (G2) and third trajectory (G3) with respect of optimism score. Furthermore, neuroticism of patients had also significantly related with the class memberships. Thus, group one patients were significantly more neurotic than group three and there was no difference between group two and group three with respect to neuroticism score. These results indicate that trajectories containing patients with dramatically decreasing resilience score (G1), were made up of patients with significantly more neurotic than trajectories of high stable resilience score (G3). The other important predictor we found was the positive emotion of the patients. Thus, group one patients had significantly less positive emotion than group three and group two patients had also significantly less positive emotion compare to group three. These results indicate that trajectories containing patients with dramatically decreasing resilience score (G1), were made up of patients who had low positive emotion than trajectories of high stable resilience score (G3). Again, trajectories containing patients with slightly decreasing resilience score (G2), were made up of patients who had low positive emotion compare to trajectories of high stable resilience score (G3). Finally, it can be seen also that baseline life event scale was not related with class memberships.

Next, for the healthy group we applied similar model selection procedure and obtained a model which consists of main effects of resilience, neuroticism and positive emotion as best candidates

among all the available predictors. However, we included life event scale, negative emotion and optimism for the sake of comparison with the cancer group. Results are presented in Table 18. It can be seen that, the baseline resilience score, neuroticism and positive emotion were risk factors that are identified as significant predictors and significantly related to the grouping or class memberships in the control group. Therefore, group one had significantly lower baseline resilience score than group three and group two had significantly lower baseline resilience score than groups three. This indicate that trajectories containing patients with low stable resilience score (G1), were made up of patients with significantly lower resilience scores than trajectories of high stable resilience score(G3). Furthermore, neuroticism of patients had also significantly related with the class memberships. Thus, group one patients were significantly more neurotic than group three and group two patients were significantly more neurotic as compare to group three. The baseline positive emotion of patients had significantly related with the class memberships. Thus, group one patients had significantly less positive emotion than group three and group two patients had also significantly less positive emotion as compare to group three. These results showed that trajectories containing patients with low stable resilience score (G1), were made up of patients who had low positive emotion than trajectories of high stable resilience score (G3). Again, the trajectories containing patients with medium stable score (G2), were made up of patients who had low positive emotion compare to trajectories of high stable resilience score (G3). Finally, it can be seen also that baseline Life event scale(LES) and optimism were not significantly related to the class memberships in the control group.

Since we had only two trajectories for anxiety and depression score a logistic regression analysis using the stable high anxiety (G2) as the reference group was fitted. Results of the fit are provided in Table 19. From the table education, optimism, neuroticism and negative emotion were risk factors identified as predictors and significantly related to the class memberships in the breast cancer patients. This results indicate that, trajectories with stable high anxiety score (G2), were significantly made up of patients with primary to secondary school education level than trajectories containing increasing anxiety level (G1). With respect to optimism we found, group one had significantly higher baseline optimism score than group two. These results indicate that trajectories containing patients with highly increasing anxiety score(G1), were made up of patients with significantly higher optimism scores than trajectories of stable high anxiety score(G2). Furthermore, group one patients were significantly less neurotic than group two and trajectories containing patients with highly increasing anxiety score(G1), were made up of patients with significantly less

Table 18: A Multinomial logistic regression model for predictors of resilience score class memberships

	Cancer group		Control group	
	High Stable (G3)vs		**high stable(G3)vs**	
	D.decreasing(G1)	S.decreasing (G2)	low stable(G2)	medium stable (G2)
Predictors	Est.(SE)	Est.(SE)	Est.(SE)	Est.(SE)
Intercept	38.36(5.89)**	25.49(5.63)**	76.89(16.92)**	54.58(13.92)**
T1CDR	-0.40(0.06)**	-0.25(0.05)**	-1.05(0.21)**	-0.53(0.13)**
LOT	-0.26(0.11)*	-0.11(0.07)	-0.17(0.25)	-0.24(0.15)
EPQ	0.35(0.16)*	0.18(0.12)	1.58(0.42)**	0.83(0.27)*
posaff	-0.65(0.14)**	-0.44(0.11)**	-1.25(0.37)**	-0.56(0.23)*
negaff	0.02(0.09)	0.10(0.06)	0.08(0.23)	-0.10(0.17)
LES	-0.02(0.33)	0.28(0.21)	-0.29(0.57)	0.02(0.19)

G1 = Dramatically decreasing, G2=slightly decreasing

$^{*}p < 0.05,\ ^{**}\ p < 0.001$

neurotic than trajectories of stable high anxiety score(G2). Finally, group one patients had significantly less negative emotion than group two. Next, in the anxiety score for control group, it can be seen that age, neuroticism, positive emotion and negative emotion, were significantly related to the grouping or class memberships. Thus, group one patients were significantly less neurotic than group two and group one patients had significantly less negative emotion and high positive emotion as compare to patients in group two. From Table 19, for depression score in cancer patients, we obtained also that positive emotion, negative emotion and optimism were significantly related to the class memberships. It can be seen that, group one patients had significantly less negative emotion than group two. This means that trajectories containing patients with highly increasing depression score (G1), were made up of patients who had low negative emotion compare to trajectories of stable high depression score (G2). Again group one patients had significantly high positive emotion compare to group two i.e. trajectories containing patients with highly increasing depression score (G1), were made up of patients who had high positive emotion compare to trajectories of stable high depression score (G2). Next for the depression score in the control group, positive emotion, negative emotion, age, marital status, optimism, and neuroticism were significant predictors related to the class memberships. Thus, group one patients had significantly less negative emotion and high positive emotion than group two. This implies that trajectories containing patients with highly increasing depression score (G1), were made up of patients who had low negative emotion or high positive emotion compare to trajectories of slightly increasing depression score (G2). Moreover, group one patients had significantly less neuroticism, high optimism and more younger than group

two. Therefore, trajectories containing patients with highly increasing depression score (G1), were made up of patients who had less neuroticism, high optimism and more younger as compare to trajectories of higher but slightly increasing depression score (G2).

Table 19: Logistic regression for predictors of anxiety and depression class membership for cancer and control group

	anxiety		**depression**	
	cancer	**control**	**cancer**	**control**
Predictors	Est.(SE)	Est.(SE)	Est.(SE)	Est.(SE)
Intercept	-4.51(4.31)	22.24(6.26)*	-3.23(2.11)	7.49(8.01)
Age	0.06(0.04)	-0.21(0.06)*	-0.02(0.02)	-0.63(0.23)*
MS	0.12(1.20)	1.78(1.03)	0.48(0.60)	3.68(1.61)*
EDU	-2.37(0.95)*	-1.82(1.05)	-0.17(0.47)	-0.01(1.67)
LOT	1.06(0.26)*	0.19(0.12)	0.18(0.06)*	0.96(0.34)*
EPQ	-0.39(0.14)*	-1.12(0.27)*	0.002(0.08)	-2.28(0.81)*
posaff	-0.22(0.12)	0.43(0.16)*	0.16(0.06)*	1.75(0.65)*
negaff	-0.67(0.16)*	-1.81(0.40)*	-0.18(0.05)*	-0.97(0.41)*
T1CDR	-0.02(0.04)	-0.05(0.04)	-0.03(0.02)	-0.003(0.08)
LES	-0.05(0.32)	0.12(0.36)	-0.11(0.15)	-0.48(0.46)

* are significant predictors of the class membership at 5% level

3.6 Handling Missing Data

Here, we reconsidered the linear mixed model and latent class growth model analysis, to check the results if they are valid under the MAR assumption. Considering the three responses, three analysis were compared; the complete case (CC) analysis, direct likelihood (DL) and linear mixed model based on multiple imputation (MI) with number of imputations, M=15. Thus, from Table 20 appendix A for linear mixed model, the parameter estimates and standard errors for both DL and the one with MI had almost closer results than the one obtained from (CC). However, in terms of conclusion we found the same statistical significance in all the three approaches. Again, from Table 21 and Table 22 in appendix A for anxiety and depression score, the parameter estimates and standard errors in DL and MI were almost closer. Except the slight difference in magnitude of the parameter estimates, we found the same inference in all approaches here as well. Lastly, for latent class growth model, the estimates and standard errors for each trajectory in each response

was shown in Table 23, Table 24 and Table 25 appendix A, respectively. Results were similar in magnitude for each trajectory in each response for both approaches. This finding could be due to the fact that DL is valid under MAR assumption.

3.6.1 Sensitivity Analysis

In some situations, the MAR assumption might be violated. Models for the MNAR case are more general, but their inferences are typically highly dependent on implicit assumptions which cannot be assessed and tested from the data. To explore the impact of deviations from the MAR assumption on the inference, sensitivity analysis should be performed. Hence, we performed small sensitivity analysis under the missing not at random (MNAR) assumption. That is, missing values are imputed under a plausible MNAR scenario, and the results are examined. If this scenario leads to a conclusion different from that of inference under MAR, then the MAR assumption is questionable (Yuan, 2014). Then, under MNAR assumption, the missing values are imputed by using pattern-mixture model approach in which missing resilience, anxiety and depression score values in cancer group are imputed from a posterior distribution generated from observations in the healthy group, and the imputed values are adjusted to reflect the systematic difference between the distributions for missing and observed response values. Under the MNAR assumption, we adjusted the imputed values using a shift parameter. The multiple imputation under MNAR was conducted at four distinct shift parameters. At each setting, multiple imputation of size (M=15) was generated. Next, the linear mixed model was fitted to the imputed data set at each setting and compared to the likelihood based inference. The parameter estimates and standard errors obtained from LMM at different shift parameters are summarized in Table 26, Table 27 and Table 28 appendix A for each response. From the table, it can be seen that both parameter estimates and standard errors at different shift parameters are close to each other. These results were also compared to the likelihood based LMM fitted under the assumption of MAR. Therefore, based on our observation, there was no substantial evidence against the assumption of MAR for the missingness.

4 Discussion and Conclusions

This study was conducted to investigate the trajectories of resilience, anxiety and depression as well as the predictors of those trajectories for both breast cancer patients and control group. Furthermore, it was also aimed at studying how resilience, anxiety and depression evolves over time. Linear mixed model was fitted and results showed that, optimism and positive emotion were positively associated with resilience score and neuroticism and negative emotion were negatively related with resilience score. Again there was significant baseline difference between breast cancer and control group and the evolution over time for resilience score depends on health status of the individuals. Results for the anxiety score also showed that, there was significant effect of optimism, neuroticism and negative emotion with anxiety score. Moreover, the evolution between breast cancer and control group was significantly different. Similarly for the depression score, age, optimism scale, neuroticism, positive emotion and negative emotion had significant effect on depression score and the evolution of depression score depends on group. Next, a Latent Class Growth analysis was used to identify distinct subgroups of individuals. For resilience score, we obtained three classes (decreasing dramatically, slight decrease and high stable) as optimal number of trajectories in the breast cancer and three trajectories (low stable, medium stable and high stable) as optimal number of trajectories in the control group. However, the low to medium resilient class in breast cancer patients decrease their resilience score over time while all the trajectories in the control group do not change over time and those patients who have had high baseline resilience score showed the same pattern with the healthy group. For anxiety and depression score, we found two trajectories for breast cancer (i.e. trajectory of highly increasing and high stable) and two trajectories for control group (i.e. highly increasing and slightly increasing score). Lastly, a multinomial logistic regression model was fitted to determine which covariates are predictive to those extracted trajectories. Results for the resilience score showed that, baseline resilience, optimism, neuroticism, positive emotion and negative emotion were risk factors that are identified as significant predictor related to the class memberships in the breast cancer patients and baseline resilience score, neuroticism and positive emotion were risk factors that are identified as significant predictors in the control group. Furthermore, in the anxiety score education, optimism, neuroticism and negative emotion were risk factors identified as significant predictors and significantly related to the classes in the breast cancer

patients. And for anxiety score in the control group, education, age, neuroticism, positive emotion and negative emotion, were risk factors related to the class memberships. For the depression score in cancer patients, positive emotion, negative emotion and baseline resilience were significantly related to the class memberships and for the depression score in the control group, positive emotion, negative emotion, age, marital status, optimism, and neuroticism were significant predictors related to the class memberships.

5 References

1. Agresti, A. (2002). Categorical Data Analysis. Second edition. John Wiley and Sons, Inc., New Jersey.

2. Andruff, H., Carraro, N., Thompson, A. and Gaudreau, P. (2009). Latent Class Growth Modelling: A Tutorial. Tutorials in Quantitative Methods for Psychology, 5, 11-24.

3. Aspinwall, L.G.and MacNamara, A. (2005). Taking positive changes seriously: Toward a positive psychology of cancer survivorship and resilience. Cancer Supplement, 104, 2549-2556.

4. Bonanno, G. A. (2004). Loss, trauma, and human resilience: Have we underestimated the human capacity to thrive after extremely aversive events? American Psychologist, 59, 20-28. doi:10.1037/0003-066X.59.1.20

5. Bonanno, G.A., Galatzer-Levy, I.R., Kennedy, P., Elfstrom, M.L. and Lude, P. (2012) Trajectories of Resilience, Depression, and Anxiety Following Spinal Cord Injury. Rehabilitation Psychology, 57, 236-247. doi: 10.1037/a0029256

6. Connor, KM. and Davidson, JRT. (2003). Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-Risc). Depression and Anxiety, 18, 76-82.

7. Curran, P. J. and Muthen, B. O. (1999). The application of latent curve analysis to testing developmental theories in intervention research. American Journal of Community Psychology, 27, 567-595.

8. deRoon-Cassini,T. A., Mancini, A. D., Rusch, M. D. and Bonanno, G. A. (2010). Psychopathology and resilience following traumatic injury: A latent growth mixture model analysis. Rehabilitation Psychology, 55, 1-11. doi:10.1037/a0018601

9. Jones,B.L., Nagin, D.S. and Roeder, K. (2001). A SAS procedure based on mixture models for estimating developmental trajectories. Sociological Methods and Research, 29, 374-393.

10. Jang, Yuri. and Haley, William E. (2002). "Life Events and Stress." Encyclopedia of Aging. Retrieved July 20, 2015 from Encyclopedia.com: http://www.encyclopedia.com/doc/1G2-3402200230.html

11. Jacelon C.S.(1997). The trait and process of resilience. Journal of Advanced Nursing, 25, 123-129.

12. King, K. and Ogle, C. (2014). Negative Life Events Vary by Neighborhood and Mediate the Relation between Neighborhood Context and Psychological Well-Being. PLoS ONE, 9, e93539. doi:10.1371/journal.pone.0093539.

13. Koutrouli, N., Anagnostopoulos, F. and Potamianos, G.(2012). Posttraumatic stress disorder and posttraumatic growth in breast cancer patients: A systematic review. Women and Health, 52, 503-516.

14. Luthar, SS., Cicchetti, D. and Becker, B. (2000). The construct of resilience: A critical evaluation and guidelines for future work. Child Development, 71, 543-562.

15. Markovitz, S., Schrooten, W., Arntz, A., and Peters, M.(2015). Resilience as a predictor for emotional response to the diagnosis and surgery in breast cancer patients.

16. Molenberghs, G. and Verbeke, G. (2005). Models for Discrete Longitudinal Data. New York: Springer-Verlag.

17. Morano, C. (2010). Resilience and Coping With Trauma: Does Gender Make a Difference? Journal of human behavior in the social Environment, 20(4), 553-568.

18. Muthen, B. (2003). Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran. Psychological Methods, 8, 369-377. doi:10.1037/1082-989X.8.3.369

19. Nagin, D.S. (2002). Overview of a semi-parametric, group-based approach for analyzing trajectories of development. Proceedings of Statistics Canada Symposium 2002: Modelling Survey Data for Social and Economic Research.

20. Nagin, D.S. (2005). Group-based modelling of development. Cambridge, MA: Harvard University Press.

21. Niyonkuru,C., Wagner, A.K., Ozawa, H., Amin, K., Goyal, A. and Fabio, A. (2013). Group-Based Trajectory Analysis Applications for Prognostic Biomarker Model Development in Severe TBI: A Practical Example. Journal of Neurotrauma, 30:938-945. doi: 10.1089/neu.2012.2578

22. Norris, F. H. (1992). Epidemiology of trauma: Frequency and impact of different potentially traumatic events on different demographic groups. Journal of Consulting and Clinical Psychology, 60, 409-418. doi:10.1037/0022-006X.60.3.409

23. Rutter, M. (1993). Resilience: Some conceptual considerations. Journal of Adolescent Health, 14, 626-631

24. Verbeke, G. and Molenberghs, G. (2000). Linear Mixed Models for Longitudinal Data. Springer-Verlag New Your, Inc., 175 Fifth Avenue, New York, NY 10010, USA.

25. Woodgate, RL (1999). Conceptual understanding of resilience in the adolescent with cancer: Part 1. Journal of Pediatric Oncology Nursing, 16, 35-43.

26. Yuan, Y. (2014) Sensitivity Analysis in Multiple Imputation for Missing Data. SAS Institute Inc., Paper SAS270.

27. Zigmond, AS. and Snaith, RP. (1983). The hospital anxiety and depression scale. Acta Psychiatrica Scandinavica, 67, 361-370.

6 Appendix

6.1 Appendix A: Tables

Table 20: Resilience, parameter estimates with their standard errors for linear mixed models based on CC,DL and MI

		CC	**Direct Liklihood**	**MI**
Effect	**Parameter**	**Est.(SE)**	**Est.(SE)**	**Est.(SE)**
Intercept	β_0	46.12(4.74)*	42.96(3.41)*	43.00(3.34)*
GROUP1	β_1	8.47(1.48)*	8.42(1.24)*	8.16(1.38)*
time	β_2	1.08(0.84)	0.68(0.57)	0.36(0.57)
LOT	β_3	0.40(0.16)*	0.43(0.13)*	0.58(0.12)*
EPQ	β_4	-0.77(0.23)*	-0.68(0.20)*	-0.60(0.18)*
posaff	β_5	1.15(0.16)*	1.43(0.14)*	1.06(0.13)*
negaff	β_6	-0.39(0.14)*	-0.43(0.12)*	-0.42(0.10)*
LES	β_7	0.17(0.38)	0.17(0.33)	0.12(0.34)
GROUP1*time	β_8	-3.57(1.05)*	-2.98(0.84)*	-2.67(0.77)*
$var(b_{1i})$	d_{11}	34.37(8.90)	36.33(8.47)	36.45(8.12)
$var(b_{2i})$	d_{22}	1.76(7.63)	0.46(9.96)	1.00(9.50)
$var(b_{1i}, b_{2i})$	d_{12}	17.43(5.98)	13.45(5.63)	13.45(5.63)
$var(\varepsilon_{ij})$	σ^2	63.66(6.19)	63.66(6.91)	63.38(5.50)

* are significant at 5%

Table 21: Anxiety, parameter estimates with their standard errors for linear mixed models based on CC, DL and MI

		CC	Direct Liklihood	MI
Effect	**Parameter**	**Est.(SE)**	**Est.(SE)**	**Est.(SE)**
Intercept	β_0	4.03(0.97)*	4.31(0.98)*	4.90(0.87)*
Age	β_1	0.02(0.01)*	0.02(0.01)*	0.02(0.01)*
GROUP	β_3	0.52(0.35)	0.65(0.35)	0.76(0.32)
time	β_4	2.32(0.17)*	2.35(0.17)*	2.33(0.16)*
T1LOT	β_5	-0.07(0.03)*	-0.08(0.03)*	-0.08(0.03)*
T1EPQ	β_6	0.27(0.04)*	0.26(0.04)*	0.27(0.04)*
T1posaff	β_7	-0.06(0.03)*	-0.06(0.03)*	-0.05(0.03)*
T1negaff	β_8	0.15(0.03)*	0.15(0.03)*	0.13(0.02)*
T3LES	β_9	0.20(0.07)*	0.20(0.33)*	0.19(0.06)*
GROUP*time	β_{10}	-0.72(0.25)*	-0.81(0.25)*	-0.84(0.23)*
$var(b_{1i})$	d_{11}	2.23(0.71)	2.28(0.71)	1.15(0.31)
$var(b_{2i})$	d_{22}	0.50(0.39)	0.50(0.40)	0.51(0.31)
$cov(b_{1i}, b_{1i})$	d_{12}	-0.93(0.47)	-0.95(0.48)	-0.98(0.31)
$var(\varepsilon_{ij})$	σ^2	5.58($1e^{-7}$)	5.67($1e^{-7}$)	5.60($1e^{-7}$)

* are significant at 5%

Table 22: Depression, parameter estimates with their standard errors for linear mixed models based on CC, DL and MI

		CC	Direct Likelihood	MI
Effect	**Parameter**	**Est.(SE)**	**Est.(SE)**	**Est.(SE)**
Intercept	β_0	3.29(0.77)*	3.37(0.75)*	3.17(0.74)*
Age	β_1	0.02(0.01)*	0.02(0.01)*	0.01(0.01)
GROUP	β_2	0.62(0.37)	0.73(0.37)	0.78(0.36)
time	β_3	3.62(0.19)*	3.68(0.19)*	3.70(0.18)*
LOT	β_4	-0.04(0.02)	-0.04(0.02)	-0.03(0.02)
EPQ	β_5	0.05(0.03)	0.05(0.03)	0.05(0.03)
posaff	β_6	-0.03(0.02)	-0.03(0.02)	-0.02(0.02)
negaff	β_7	0.05(0.02)*	0.05(0.02)*	0.05(0.02)*
LES	β_8	0.03(0.05)	0.03(0.02)	0.04(0.05)
GROUP*time	β_9	-0.85(0.28)*	-0.95(0.28)*	-0.97(0.25)*
$var(b_{1i})$	d_{11}	3.09(0.87)	3.16(0.88)	3.15(0.61)
$var(b_{2i})$	d_{22}	0.30(0.51)	0.34(0.63)	0.33(0.50)
$cov(b_{1i}, b_{2i})$	d_{12}	-2.51(0.62)	-2.59(0.52)	-2.60(0.56)
$var(\varepsilon_{ij})$	σ^2	6.79(1E-7)	6.90(1E-6)	7.85(1E-6)

* are significant at 5%

Table 23: Parameter estimates under DL and MI for resilience score in each group conditional LCGM

	Intercept		**Slope**	
	DL	**MI**	**DL**	**MI**
CDR	**est(SE)**	**est(SE)**	**est(SE)**	**est(SE)**
Cancer				
D.decrease	51.19(2.04)*	51.40(0.32)*	-7.42(2.48)*	-4.51(0.32)*
slightly decrease	65.03(1.33)*	66.71(0.28)*	-2.44(1.04)*	-2.16(0.21)*
high stable	82.71(1.27)*	82.32(0.30)*	-0.56(1.36)	-0.84(0.28)
σ	9.73(0.36)	9.64(0.06)		
control				
low stable	49.83(2.21)*	50.16(0.46)*	2.23(1.81)	0.11(0.35)
medium stable	65.92(0.73)*	66.13(0.02)*	0.19(0.69)	0.01(0.16)
high stable	82.52(0.31)*	82.51(0.28)*	-0.13(1.20)	0.53(0.27)
σ	7.71(0.31)	7.74(0.11)		

Table 24: Parameter estimates under DL and MI for anxiety in each group conditional LCGM

	Intercept		**Slope**	
	DL	**MI**	**DL**	**MI**
Anxiety	**est(SE)**	**est(SE)**	**est(SE)**	**est(SE)**
Cancer				
Increasing Anxiety	4.46(0.28)*	4.50(0.07)*	2.66(0.28)*	2.64(0.05)*
stable high	11.41(0.38)*	11.23(0.09)*	-0.30(0.36)	-0.19(0.07)
σ	2.88(0.12)	2.65(0.02)		
control				
highly increasing	2.94(0.95)*	3.19(0.06)*	3.23(0.22)*	3.14(0.05)*
slightly increasing	8.33(0.27)*	8.59(0.09)*	1.13(0.24)*	1.10(0.06)*
σ	2.16(0.09)	2.14(0.02)		

Table 25: Parameter estimates under DL and MI for depression level in each group conditional LCGM

	Intercept		**Slope**	
	DL	**MI**	**DL**	**MI**
Depression	**est(SE)**	**est(SE)**	**est(SE)**	**est(SE)**
Cancer				
Increasing Depression	1.66(0.26)*	1.72(0.05)*	4.65(0.25)*	4.61(0.04)*
stable high	9.25(0.38)*	8.86(0.07)*	-0.29(0.37)	-0.10(0.06)*
σ	2.46(0.11)	2.25(0.06)		
control				
highly increasing	1.98(0.16)*	2.09(0.04)*	4.58(0.18)*	4.51(0.04)*
slightly increasing	7.43(0.32)*	7.34(0.03)*	1.19(0.31)*	1.13(0.07)*
σ	1.95(0.08)	1.93(0.02)		

Table 26: Comparing models at different shift parameters under MNAR assumption using multiple imputation(MI) for Resilience score

		N-Shift Parameters			
	DL	**-4**	**-1**	**2**	**5**
	Est(SE)	Est(SE)	Est(SE)	Est(SE)	Est(SE)
Intercept	42.96(3.81)*	41.53(3.26)*	42.58(3.24)*	43.49(3.24)	44.28(3.27)*
group	8.42(1.24)*	7.29(1.15)*	6.42(1.15)*	5.59(1.16)	4.78(1.18)*
time	0.67(0.57)	0.19(0.59)	0.34(0.59)	0.48(0.59)	0.62(0.60)
LOT	0.43(0.13)*	0.68(0.12)	0.67(0.12)*	0.67(0.12)	0.66(0.12)*
EPQ	-0.68(0.19)*	-0.52(0.19)	-0.54(0.19)*	-0.56(0.20)	-0.58(0.20)*
posaff	1.33(0.14)*	1.04(0.12)*	1.04(0.12)*	1.04(0.11)	1.04(0.12)*
negaff	-0.43(0.13)*	-0.39(0.10)*	-0.38(0.10)*	-0.38(0.11)	-0.37(0.11)*
LES	0.17(0.33)	0.11(0.33)	0.11(0.33)	0.10(0.33)	0.09(0.34)
Group*time	-2.98(0.84)*	-2.60(0.85)*	-2.75(0.84)*	-2.92(0.85)*	-3.07(0.86)*

* are sgnificant at 5% level

Table 27: Comparing models at different shift parameters under MNAR assumption using multiple imputation (MI) for anxiety score

		N-Shift Parameters			
	DL	**-4**	**-1**	**2**	**5**
	Est(SE)	Est(SE)	Est(SE)	Est(SE)	Est(SE)
Intercept	4.31(0.98)*	4.62(0.91)*	4.75(0.85)*	4.86(0.87)*	4.91(0.97)*
age	0.02(0.01)*	0.02(0.008)*	0.01(0.008)*	**0.01(0.01)**	**0.01(0.009)**
group	0.65(0.35)	0.54(0.30)	0.48(0.28)	0.34(0.29)	0.17(0.31)
time	2.35(0.17)*	1.99(0.17)*	2.24(0.16)*	2.49(0.16)*	2.74(0.17)*
LOT	-0.08(0.03)*	-0.12(0.03)*	-0.11(0.025)*	-0.11(0.03)*	-0.11(0.03)*
EPQ	0.26(0.04)*	0.29(0.04)*	0.28(0.04)*	0.28(0.04)*	0.29(0.04)*
posaff	-0.06(0.03)*	-0.06(0.03)*	-0.05(0.02)*	-0.05(0.02)*	**-0.05(0.03)**
negaff	0.15(0.03)*	0.16(0.02)*	0.16(0.02)*	0.16(0.02)*	0.18(0.02)*
LES	0.17(0.33)	0.21(0.07)*	0.19(0.07)*	0.19(0.07)*	0.21(0.08)*
group*time	-2.98(0.84)*	-0.46(0.22)*	-0.72(0.21)*	-0.96(0.21)*	-1.22(0.23)*

*are significant at 5 level

Table 28: Comparing models at different shift parameters under MNAR assumption using multiple imputation(MI) for Depression score

		N-Shift Parameters			
	DL	**4**	**1**	**2**	**5**
	Est(SE)	Est(SE)	Est(SE)	Est(SE)	Est(SE)
Intercept	3.37(0.77)*	3.84(0.85)*	3.95(0.80)*	4.08(0.80)*	4.23(0.85)*
age	0.02(0.01)*	0.02(0.008)*	0.02(0.007)*	**0.01(0.007)**	**0.01(0.008)**
group	0.73(0.37)	0.38(0.29)	0.33(0.27)	0.24(0.28)	0.13(0.30)
time	3.68(0.19)*	3.30(0.17)*	3.59(0.17)*	3.88(0.17)*	4.18(0.18)*
LOT	-0.04(0.02)	-0.08(0.03)*	-0.08(0.03)*	-0.07(0.025)*	-0.07(0.02)*
EPQ	0.05(0.03)	0.13(0.04)*	0.12(0.04)*	0.11(0.04)*	0.11(0.04)*
posaff	-0.03(0.02)	-0.06(0.02)*	-0.06(0.02)*	-0.06(0.02)*	-0.07(0.02)*
negaff	0.05(0.02)*	0.08(0.02)*	0.08(0.02)*	0.08(0.02)*	0.09(0.02)*
LES	0.03(0.02)	0.11(0.07)	0.11(0.06)	0.09(0.07)	0.08(0.08)
group*time	-0.95(0.28)*	-0.55(0.23)*	-0.85(0.23)*	-1.14(0.23)*	-1.44(0.24)*

6.2 Appendix B: Selected SAS codes

```
*******Linear Mixed Model for Resilience score*************;
proc mixed data=mydataCDR method=ml  covtest;
title ' LMM for CDR'; class idnr  timeclass ;
model CDR=  groups  time T1LOT T1EPQ T1PANASposaff T1PANASnegaff T3LES groups*time
/solution covb ddfm=satterth;
random  intercept time /type=un subject=idnr;
repeated timeclass/ type=simple  subject=idnr;
run; quit;
*******Linear Mixed Model for Anxiety score*****************;
proc mixed data=mydataAD method=ml covtest ;
title 'lmm model for Anxiety';  class idnr  timeclass ;
model HADSA =  age groups  time T1LOT T1EPQ  T1PANASposaff   T1PANASnegaff  T3LES
groups*time / solution covb
ddfm=satterth;
random intercept time / type=un subject=idnr;
repeated timeclass/ type=simple  subject=idnr;
run; quit;
*******Linear Mixed Model for Depression score*****************;
proc mixed data=mydataAD method=ml covtest ;
title 'lmm  model for depression level';
class idnr  timeclass ;
model HADSD =Age groups  time   T1LOT T1EPQ  T1PANASposaff T1PANASnegaff  T3LES
groups*time/ solution
ddfm=satterth;
random intercept time / type=un subject=idnr;
repeated timeclass/ type=simple subject=idnr;
*ods output SolutionF=mixparms Covb=mixcovb;
run; quit;
```

```
******************** Latent Class Growth  Model************;
*For CDR cancer group we found three groups without and with covariates ;
proc traj data=trajS out=cdrcancer outstat=os outplot=op;
var CDR01-CDR03; indep t01-t03;
model CNORM;  MAX 100;  ngroups 3; order 1 1 1 ;
%TRAJTEST('interc1=interc2') /* intercept equality test between 1&2*/
risk T1LOT  T1EPQ T1PANASposaff
T1PANASnegaff ms edu; *converges problem if we include more variable; id idnr; run;
%trajplot(OP, OS," conditional LCGM", "CDR cancer", "time in year ");
%TRAJTEST('interc1=interc3');
************************************************************;
*For CDR control group we found three groups without and with covariates ;
proc traj data=traj out=cdrcontrol outstat=os outplot=op;
var CDR01-CDR03; indep t01-t03;
model CNORM;  MAX 100; ngroups 3; order   1 1 1 ;
risk  T1LOT T1LES T1EPQ T1PANASposaff  T1PANASnegaff age edu ms ;  id idnr; run;
%trajplot(OP, OS," conditional LCGM", "CDR control",
"time in year ");
************************************************************;
* For Anxiety cancer group we found 3 groups without and 2 groups with covariates ;
proc traj data=trajS out=Acancer outstat=os outplot=op;
var HADSA01-HADSA02; indep t1-t2;
model CNORM;  MAX 20;  ngroups 2; order  1 1;
risk   ms edu Age T1LES  T1LOT   T1EPQ T1PANASposaff T1PANASnegaff ; id idnr; run;
%trajplot (OP, OS, "conditional LCGM", "anxiety", " Anxiety  ", "time in year ");
************************************************************;
*For Anxiety control group we found 3 groups without and 2 groups with covariates ;
proc traj data=traj out=Acontrol outstat=os outplot=op;
var HADSA01-HADSA02; indep t1-t2;
model CNORM; min 0;MAX 19;  ngroups 2; order  1 1;
```

```
risk  ms edu age  T1CDR T1LES T1LOT T1EPQ T1PANASposaff  T1PANASnegaff ; id idnr; run;
%trajplot (OP, OS, "Conditional LCGM", "Control", " Anxiety", "time in year ");
************************************************************;
*For depression cancer and control group we
 found 3 groups without and 2 groups with covariates ;
proc traj data=trajS out=Dcancer outstat=os outplot=op;
var HADSD01-HADSD02; indep t1-t2;
model CNORM; min 0; MAX 20;  ngroups 2; order  1 1;
risk ms edu Age T1CDR T1LES  T1LOT   T1EPQ T1PANASposaff T1PANASnegaff ; id idnr; run;
%trajplot (OP, OS, "Conditional LCGM"," Depression ", "time in year ");
************************************************************;
* for depression control group***;
proc traj data=traj out=Dcontrol outstat=os outplot=op;
var HADSD01-HADSD02; indep t1-t2;
model CNORM;  min 0;MAX 15; ngroups 2; order 1  1;
risk ms edu Age T1CDR T1LES T1LOT   T1EPQ T1PANASposaff T1PANASnegaff ; id idnr; run;
%trajplot (OP, OS, "Conditional LCGM","Depression ", "time in year ");
************************************************************;
*** Multinomial logistic Regression only for CDR  cancer and control group;
PROC LOGISTIC DATA=cdrcancergroup ;
*CLASS ;
MODEL group = T1CDR  T1LOT  T1EPQ T1PANASposaff T1PANASnegaff  T1LES
/link=glogit aggregate scale=none expb; RUN;
************************************************************;
proc logistic data=cdrcontrolgroup;
*class  ;
model group = T1CDR T1LOT T1EPQ T1PANASposaff T1PANASnegaff  T1LES/ link=glogit
aggregate scale=none expb; run;
************************************************************;
* The macro for sensetivity analysis under MNAR;
```

```
************************************************************;
%macro midata( data=, smin=, smax=, sinc=, out=);
data &out;
set _null_;
run;
/*------------ # of shift values ------------*/
%let ncase= %sysevalf( (&smax-&smin)/&sinc, ceil );
/*------- Imputed data for each shift -------*/
%do jc=0 %to &ncase;
%let sj= %sysevalf( &smin + &jc * &sinc);
proc mi data=&data seed=123458 nimpute=15 out=outmi;
class groupnumber; fcs reg;
mnar adjust(T1CDR/ shift=&sj adjustobs=(groupnumber='2') );
mnar adjust(T2CDR/ shift=&sj adjustobs=(groupnumber='2') );
mnar adjust(T3CDR/ shift=&sj adjustobs=(groupnumber='2') );
var groupnumber AGE MaritalStatus Education  T1CDR T2CDR T3CDR T1HADSA T2HADSA  T1HADSD
T2HADSD T1LOT T1EPQ T1LES T3LES  T1PANASposaff T1PANASnegaff;
run; data outmi; set outmi; Shift= &sj; run;
data &out; set &out outmi; run; %end; %mend midata;
ods listing close; %midata( data=mydata, smin=-4, smax=5, sinc=3, out=sens);

data ex1;set sens; array y(3) T1CDR T2CDR T3CDR;
do j=1 to 3; CDR=y(j); t=j; output;end;
keep CDR idnr t groupnumber age MaritalStatus Education T1LOT T1EPQ T1LES T3LES
T1PANASposaff T1PANASnegaff _Imputation_ shift; run;
data last;set ex1;if t=1 then time=0; if t=2 then time=0.5;
if t=3 then time=1.5; if Education=1 or Education=2 then edu=1;
else if Education=3 or Education=4 then edu=0;
if MaritalStatus=1 then ms=1;
else if MaritalStatus=2 or MaritalStatus=3 or MaritalStatus=4 then ms=0; drop t;
```

```
timeclass=time;run;
proc sort data=last; by shift _imputation_; run; data last1; set last; if groupnumber=1
then groups=1; else groups=0; run;
proc mixed data=last1 method=ml ;
title 'SENSITIVITY: LMM'; by shift _Imputation_;
class idnr timeclass ;
model CDR=groups time T1LOT T1EPQ T1PANASposaff T1PANASnegaff T3LES  groups*time
/ covb solution  ddfm=satterth;
random intercept time/ subject=idnr;
repeated timeclass/ type=simple subject=idnr;
ods output SolutionF=mixparms Covb=mixcovb;
run; quit;
/* Make inference */
PROC MIANALYZE parms=mixparms Covb(effectvar=rowcol)=mixcovb wcov bcov tcov;
by shift ;
modeleffects Intercept groups time T1LOT T1EPQ
T1PANASposaff T1PANASnegaff T3LES  groups*time;
ods output parameterestimates=miparm1; run;
```

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
The trajectories of resilience, anxiety and depression in breast cancer patients and control groups

Richting: **Master of Statistics-Biostatistics**
Jaar: **2015**

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,

Mesfin, Negasi Asres

Datum: **1/09/2015**