

Heterogeneous Phenotypic-Specific Developmental Trajectory and Stage-Specific Symptom Pattern Among Rheumatoid Arthritis Patients: *Who is Suffering, When, and How?*

Student: Ai Ye Clients: Leigh F. Callahan, PhD Mary Link Briggs Distinguished Professor of Medicine leigh_Callahan@med.unc.edu Becki Cleveland, PhD Assistant Professor of Medicine becki@unc.edu Thurston Arthritis Research Center University of North Carolina at Chapel Hill

04/09/2019

ABSTRACT

Background. Rheumatoid arthritis (RA) is a systemic autoimmune disorder primarily characterized by the inflammation and degradation of synovial joints, affecting an estimated 0.6% of all adults in the US. RA is a heterogeneous disease, representing the end result of many pathophysiologic mechanisms.

Objective. Goal 1. To group RA patients based on their developmental trajectories of phenotypes throughout the early onset of the disease; and Goal 2. To identify latent profile of RA patients with respect to their response pattern of all phenotype at each stage, as well as the transition pattern amongst profiles from one stage to the next. Ultimately, the goal is to identify demographic, clinical, and health behavior factors that are associated with the patient's memberships in the identified categories.

Methods. We analyzed longitudinal data from 354 patients over the age of 18 recruited in the longitudinal studies across five participating institutions in the US. Six phenotype outcomes were repeated measured at baseline, 36-month and 60-month bench marks. For Goal 1. we performed a series of univariate and multivariate Latent Growth Curve models (LGCM) and Growth Mixture models (using the curve factor structure from corresponding LGCM) for each phenotype; for Goal 2, we conducted a series of Latent Profile Analyses at each wave, followed by Latent Transition Analysis using all three waves of data.

Results. For each phenotype, we identified two clusters of people with distinct phenotype-specific developmental trajectories. Specifically, we found “recovering” vs. “progressing” groups for QOL and patient-reported health outcome, “stable” vs. “recovering” groups for functional disability, and finally, “slow-progression” vs. “fast-progression” groups for both joint damage phenotypes. In addition, at each wave, we established three latent profiles of patients with stage-specific phenotypic patterns, i.e., a large “low-risk”, a small “severe joint damage”, and a substantial “poor mental health” profiles. The transition pattern was characterized by increased in the counts of both “low-risk” and “severe joint damage” while a decrease in that of “poor mental health”. Specifically, some patients with “poor mental health” were likely to progressively transition to “low-risk” profile throughout the development of RA, and a few patients with “severe joint damage” transitioned to “low-risk” at early onset of the disease. In contrast, some patients developed “severe joint damage” or “poor mental health” at later stage.

Conclusion. Our findings suggested qualitatively distinct experience for RA patients. The two clustering approaches complementarily unraveled heterogeneous experience of patients from two different perspectives, and revealed latent group with phenotypic-specific developmental trajectory and stage-specific phenotypic pattern. We found that although the majority of patients tend to get worse in joint damage, some may recover in their symptoms and such recovery are likely to happen at early onset of the disease. On average, patients are likely to become more tolerant about the pain and fatigue they suffer throughout their first five year of disease, even though their symptoms persist. Such findings are important to improve clinical trial design, to aid the screening of patients for targeted treatment, as well as to inform the most effective treatment at the most critical stage.

INTRODUCTION

Rheumatoid arthritis (RA) is an autoimmune disease in which the body's own immune system mistakenly attacks the joints. RA affects about 1.5 million Americans, is about 3 times higher in women and it typically diagnosed between the ages of 30 and 60. It is a disease that is characterized by chronic inflammation and structural damage of cartilage, bone, and ligaments. Further, extra-articular disease can have effects on a variety of body systems, including the cardiovascular and respiratory, and also increases morbidity and mortality.

Severity of RA is classified by erosions on radiographs, counts of affected joints, disability, and health-related quality of life (QOL) such as pain and fatigue. RA is a heterogeneous disease, however, representing the end result of many pathophysiologic mechanisms. Identifying specific RA phenotypes could improve clinical trial design, and identify treatments that could be better targeted to those likely to benefit, or exclude those unlikely to benefit. Therefore our goal is to apply innovative strategies for improved phenotyping by leveraging data collected from the Consortium for the Longitudinal Evaluation of African Americans with Early Rheumatoid Arthritis Registry (CLEAR). Our goal is to cluster the key features that best define and separate RA phenotypes and subgroups within phenotypes over time.

METHOD

Data Collection Procedure

Study Population

The CLEAR registry was established by the National Institute of Arthritis and Musculoskeletal and Skin Disease (NIAMS) in order to provide the research community with thorough data on traditionally under-represented African Americans. Additional information about the CLEAR consortium can be found in the literature. But briefly, starting in 2002, 354 individuals over the age of 18 were recruited into the longitudinal CLEAR cohort at five participating institutions. Enrollment criteria included self-identification as African American, meeting the American College of Rheumatology criteria for RA, disease duration less than two years, providing informed consent, willingness to regularly participate in follow-ups, and no concurrent diagnosis of other rheumatic diseases aside from osteoarthritis. Baseline and follow-up data were collected by trained research staff. Follow-up assessments took place 36 months after disease diagnosis and again 60 months after disease diagnosis. Measures collected at baseline included demographics, comorbid conditions and SES measures including education, income, occupation and home ownership. Data also collected at baseline and at subsequent follow-ups included measures of QOL (pain and fatigue), functional disability (Health Assessment Questionnaire [HAQ]) and Joint Alignment and Motion Scale [JAM]) and joint destruction (radiographic erosions and joint space narrowing [JSN]).

Phenotype Outcome Measures

Health Outcomes Protocol. Data were collected upon participant enrollment by trained research staff, and included radiographs, dual x-ray absorptiometry data, genomic DNA, serum and plasma measurements, joint count evaluations, self-reported health status measures, comorbid conditions, demographics, education level, individual and household income, occupation, and homeownership. All study materials and methods were approved by the institutional review boards of the participating institutions. Joint damage. Radiographs of the hands and feet were taken according to a standard protocol. A single posteroanterior view of each hand, including the wrist, and radiographs evaluating the forefoot of each foot separately were taken. Radiographs were scored on the modified Sharp/van der Heijde scale by an experienced reader, blinded for patient identity and clinical information: joint erosion and joint space narrowing (JSN) were cumulatively assessed in 32 joints in the hands and 12 joints in the feet, with a maximum erosion score of 5 per hand joint and 10 per foot joint, and a maximum JSN score of 4 per joint (22). A total joint malalignment (JAM) score was created

from clinical measurements of joint normal range of motion (ROM): for the hips and knees, 10–25% decreased ROM increased the JAM score by 3 and 25% by 4; for every other joint, 0–5% decreased ROM increased JAM score by 1, 6–25% by 2, and 26–75% by 3.

Patient-reported outcomes. Self-reported health outcomes included self-rated pain and fatigue from visual analog scales (VAS), the Health Assessment Questionnaire (HAQ) disability index. Pain and fatigue symptoms were collected visual analog scales (VAS) for the previous week on anchored with “no pain/fatigue” (0) and “pain/fatigue as bad as could be” (10) resulting in a continuous variable for pain ranging from 0–10. Self-rated fatigue was evaluated following a similar procedure. VAS measures are considered to be valid representations of actual symptoms (32).

The HAQ scale is a widely used, validated measure of self-reported disability. Participants rate their ability to perform daily tasks with responses on a scale of 0–3, where higher scores indicate greater difficulty. Twenty questions were grouped into 8 categories representing functional activities for which mean scores are computed and then averaged (29).

Joint Damage. A total JAM score was created from clinical measurements of joint normal range of motion (ROM) as follows: for the hips and knees, 10–25% decreased ROM increased JAM score by 3 and >25% decreased ROM increased JAM score by 4; for every other joint, 0–5% decreased ROM increased JAM score by 1, 6–25% by 2, and 26–75% by 3. Radiographs were scored on the modified Sharp/van der Heijde scale by an experienced reader, blinded for patient identity and clinical information: joint erosion and joint space narrowing (JSN) were cumulatively assessed in 32 joints in the hands and 12 joints in the feet, with a maximum erosion score of five per hand joint and 10 per foot joint, and a maximum JSN score of four per joint.

Covariates. The following covariates were included in our analyses: age at enrollment (continuous years), sex, a clinical comorbidity index based on 23 doctor-diagnosed conditions (33), use of methotrexate and/or leflunomide (yes or no), use of biologic agents (yes or no), RA disease duration (continuous years), body mass index (BMI; kg/m²), and cigarette smoking, according to the self-reported number of pack-years smoked.

Data Analytic Strategy

Latent Growth Curve Model (LGCM) and Growth Mixture Model (GMM)

Growth Mixture Model (GMM) based on latent growth curve aims to categorize individual based on the characteristics of the individual trajectory as a function of time. Statistically speaking, this is a clustering technique on the time vector. A univariate GMM clusters a single repeatedly measured variable (time vector), a multivariate GMM clusters multiple repeatedly measured variables simultaneously (time matrix).

The LGCM is a technique to model individual trajectories as a function of time for repeated measures that is predominantly developed in social science under the latent variable framework (Bollen & Curran, 2006; Browne & du Toit, 1991; Rao, 1958; Tucker, 1958; Meredith, 1960; Meredith & Tisak, 1984). Each case in the sample is modeled as an individual trend of change dictated by latent growth factors such as an intercept and a slope in a linear case. LGCM has been increasingly popular in longitudinal studies as the attractive property to study the inter-individual differences in their intra-individual trajectories over time (Meredith and Tisak, 1990; Preacher et al., 2008; Duncan et al., 2013). Three major types of parameters are estimated to realize that goal: the mean of growth factors that dictates the overall characteristic of an averaged shape of trajectory from the sample, the variance of growth factors represents the variability of individual trajectories surrounding that mean curve, and the covariance of the growth factors that further unravels the relationship amongst growth factors. An added advantage of LGCM is the capability to easily incorporate covariates that might explain between-person variabilities in their growth trajectories. Readers interested in fitting an unconditional and conditional LGCM please refer to Bollen and Curran (2006) for more illustrations.

Under the framework of structural equation models (SEM), a LGCM can be specified as a polynomial function of repeatedly measured outcome regressed on a time variable. With the initial level and the rate of change (i.e., the intercept and slope of the growth line, respectively) being the latent factor of the SEM, the

factor loadings with fixed values represent the value of time. Such linear LGCM can be easily extended to a nonlinear LGCM by incorporating higher-order power functions (e.g., squared, cubic, etc.) to represent the nonlinear characteristic of the trajectory over time. For example, in a quadratic term, a curvilinear trend is modeled as a product of a quadratic latent factor and a second-order time metric as its factor loading (see Figure 3).

In a univariate unconditional LGCM with a Quadratic term

$$y_{it} = \Lambda_{t1}\alpha_i + \Lambda_{t2}\beta_i + \Lambda_{t3}\beta_i^2 + \epsilon_{it}$$

where Λ_{tp} is the factor loading for each moment of latent variable, the values in Λ_{tp} allow the incorporation of the shape (linear or nonlinear) of trajectories, α_i is the random intercept for person i , β_i is the random slope for person i , and β_i^2 is the random quadratic term for person i , and that

$$\alpha_i = \mu_\alpha + \zeta_{\alpha i}$$

$$\beta_i = \mu_\beta + \zeta_{\beta i}$$

$$\beta_i^2 = \mu_{\beta^2} + \zeta_{\beta^2 i}$$

Note that the trajectory model can be extended in several ways to allow for alternative form of nonlinear patterns of change. For example, Meredith and Tisak (1984, 1990) proposed the latent-basis LGC model where the curvilinear trajectories modeled by freeing one or more factor loadings. This way, nonlinear patterns are accommodated by retaining the basic form of the model above but estimating all but two of the factor loadings for the slope factor and quadratic factor, rather than fixing these to linearly or nonlinearly increasing values (see Figure 4). The free loadings allows for flexible forms as a type of nonlinear “spline” that best fits the data between any two time points. A more general form of using the concept of “spline” is what is known as piecewise linear growth models using two or more linear piecewise splines, developed under the mixed modeling framework (Bryk & Raudenbush, 1992). The general idea is to break the full range of the time period into a number of pieces using deterministic knots, and fit a regular linear LGCM to each piece of trajectory. The selection of knots often carry subjective, transitional meaning and can sometimes vary randomly across individuals.

GMM is special type of the general finite mixture models, that instead of imposing the assumption that all individual response follow one distribution, now the joint distribution can be modeled as a mixture of multiple distributions that may or may not be the same type (although the former is more common). Specifically, in GMM, parameters of growth factors (i.e., mean, variance, and covariance of intercept, slope, quadratic) are allow to vary across cluster (mixtures) while responses of individuals in the same cluster follow one distribution. Assumptions can be made by adding constraints to the cluster variable whether the distributions between clusters share some level of similarities (e.g., same variance, covariance, etc.).

Latent Class or Profile Analysis (LCA or LPA) and Latent Transition Analysis (LTA)

Latent class analysis (LCA) and Latent profile analysis (LPA), as well as their longitudinal extension latent transition analysis (LTA), are a multivariate statistical clustering technique based on measurement theory in which unobserved, underlying grouping variables (i.e., latent classes) are inferred from a set of categorical (for LCA) or continuous indicators (for LPA; Goodman, 1974; Lazarsfeld, Henry, & Anderson, 1968). Conventional LTA simultaneously defines the latent class structure and models individual-level changes in class membership over time (Collins & Flaherty, 2002). Contrasting to GMM that cluster the time vector (or matrix), LCA or LPA clusters individual based on response patterns to variables (response vector), the longitudinal extension LTA that models cluster membership at each time simultaneously (response matrix) as well as the transition pattern and probability from one cluster to another (or stay in the same cluster) over time.

Prior to conducting our LTA, it was necessary to decide how many latent classes were needed to model the indicator variables, forming the measurement model. A common procedure is to fit a series of LCA or LPA models with increasing number of classes at each time point in order to determine the number of classes to

be used (e.g., whether they are the same at each wave, etc.) in the subsequent LTA based on a combination of statistical, theoretical, and practical grounds. Researchers usually first specified a 2-class LCA or LPA model and increased the number of classes until models no longer converged or fit the data more poorly. Fit indices were then considered along with interpretations of class structures to choose the number of classes that yielded the most meaningful, parsimonious, and statistically sound model (Collins & Lanza, 2010). The Bayesian Information Criterion (BIC) and Bootstrapped Likelihood Ratio Test (BLRT) have both been recommended as the most reliable measures of fit for determining the appropriate number of latent classes (Nylund, Asparouhov, & Muthén, 2007). Strong latent class separation—manifested in highly distinct latent classes and posterior class membership probabilities near 0 and 1—is also critical for LCA models. The degree of difference among included classes is measured by the Entropy statistic, which ranges from 0 to 1, with larger values indicating better latent class separation (Collins & Lanza, 2010). If the measurement model has a low entropy value (e.g., $< .6$), this suggests that the latent class structure is a poor representation of the dependent variable, and analyses incorporating latent class membership will be unreliable.

After conducting our preliminary LPAs (since all the CLEAR outcomes are assumed continuous), we constructed an autoregressive model to estimate latent transition probabilities—the probabilities of particular patterns of profile change across time points (Collins & Lanza, 2010). Although it is possible to allow for different numbers of latent profiles at each time point, doing so can make it difficult to interpret transitions among profiles across time points, because the meaning of membership in a particular profile may fail to remain the same. Thus, we adopted the same number of latent classes with time-invariant structure at each time point to simplify interpretations of transitions, as has been previously recommended in LTA research (e.g., Lanza, Patrick, & Maggs, 2010; Nylund, 2007).

Tests of measurement invariance. Prior to estimating latent transition probabilities, it is important to investigate the common assumption of full measurement invariance—that is, the assumption that latent profile indicators produce the same classifications across time points. Although we assumed that the number of classes was constant over time, we directly tested the assumption that the parameters of the measurement model remain constant, meaning that conditional item-response probabilities underlying the latent profile structure are equivalent regardless of test time. Although this assumption has the benefit of simplicity and increases the degrees of freedom, it may be questionable more often than many researchers realize (Nylund, 2007). Rather than lifting the assumption of measurement invariance completely, however, which would lead to a full non-invariant model and the need to estimate a large number of parameters, we chose to test for partial measurement invariance—that is, a scenario in which some parameters of the measurement model are held constant, while others are allowed to vary freely (Nylund, 2007). Tests for measurement invariance were conducted prior to including longitudinal relationships between latent variables or covariate relationships in the model. We imposed the measurement invariance constraint on each indicator variable one by one, at each stage conducting a likelihood ratio test to see whether adding the constraint significantly reduced the predictive power of the model. After testing measurement invariance for the first transition (i.e., T1 to T2), we repeated this procedure for the second transition (i.e., T2 to T3).

One downside to allowing measurement non-invariance is that even if the number of classes is held constant, interpretations of class structure could potentially change if items tend to function differently at different time points. Differential item functioning that could threaten interpretations of class structure would be accompanied by significant changes in conditional response probabilities (CRP)—the probability of a given indicator response conditioned on class membership. Thus, if large changes in CRPs showed that the class structures differ qualitatively across time points, we would then impose measurement invariance on the model regardless of invariance tests in order to ensure results maintain interpretability.

The inclusion of auxiliary variables (e.g., covariates) in LCA and LTA models can be problematic if latent class membership, latent transition probabilities, and effects of covariates on each of these are all estimated simultaneously in a single step (Asparouhov & Muthén, 2014; Vermunt, 2010). The primary issue is that covariate relationships can have a marked influence on the estimation of the measurement model. Our goal was to observe a latent class structure, and then investigate how covariate values relate to both the latent profile structure at particular time points and the probabilities of transitioning between latent profiles across time points by logistic regression (for two-profile scenario) or multinomial regression (for more than two profiles scenario).

All mixture or cluster types of models (GMM, LPA, and LTA) were performed in Mplus 7.4 using full information maximum likelihood (FIML) estimation, which maximizes power by assigning values to certain kinds of cells with missing data based on all of the other available data for a given participant. Since FIML estimation is unable to assign missing values to covariates, separate the LTA and logistic or multivariate regression prevent loss of data for the LTA. Data manipulation, cleaning, management, as well as all other analyses were conducted in R studio (version X). Functions from multiple packages were used for model estimation, and result presentation, and visualization. This report was generated by Rstudio markdown.

RESULTS

This section contains the following five Parts.

Part I Descriptives

Part II Dimension Exploration (PCA)

Part III Latent Growth Curve and Growth Mixture Models (LGCM and GMM)

Part IV Latent Profile and Latent Transition Analysis

Part V Covariate Effect (TO BE ADDED)

Part I Descriptive Statistics

Boxplots (Figure 1) and summary statistics (Table 1) are provided below. The distributions of the two QOL measures at each time point were similar (slightly skewed with heavy tails, suggesting large variabilities). The baseline means for Pain and Fatigue were 6.04 and 5.73, respectively, which decreased slightly at 36 months (5.68 and 5.21, respectively) and picked up again at 60 months (5.84 and 5.62, respectively). The descriptives of the two functional disability measures behave differently. Specifically, the distribution of HAQ was approximately normal with a positive skewness, while that of JAM was extremely zero-inflated. The same decreased-first-and-increased-again pattern was seen for this two outcomes. The distributions of the two Radiographic measures were also extremely zero-inflated, but the means of which both increased over time. Note that there were larger portion of missing data in the Radiographics measures (12% at baseline alone), but the Little's test indicated that the missing pattern was "missing at random". Thus, no multiple imputation was performed prior to the analyses. Instead, Full-Information Maximum Likelihood estimation method was used to account for missingness in the analyses, and Robust Maximum Likelihood was used to correct for non-normal distribution and achieve asymptotically unbiased point estimate and efficient standard errors,

Figure 1 Boxplots for the Six Outcome Variables.

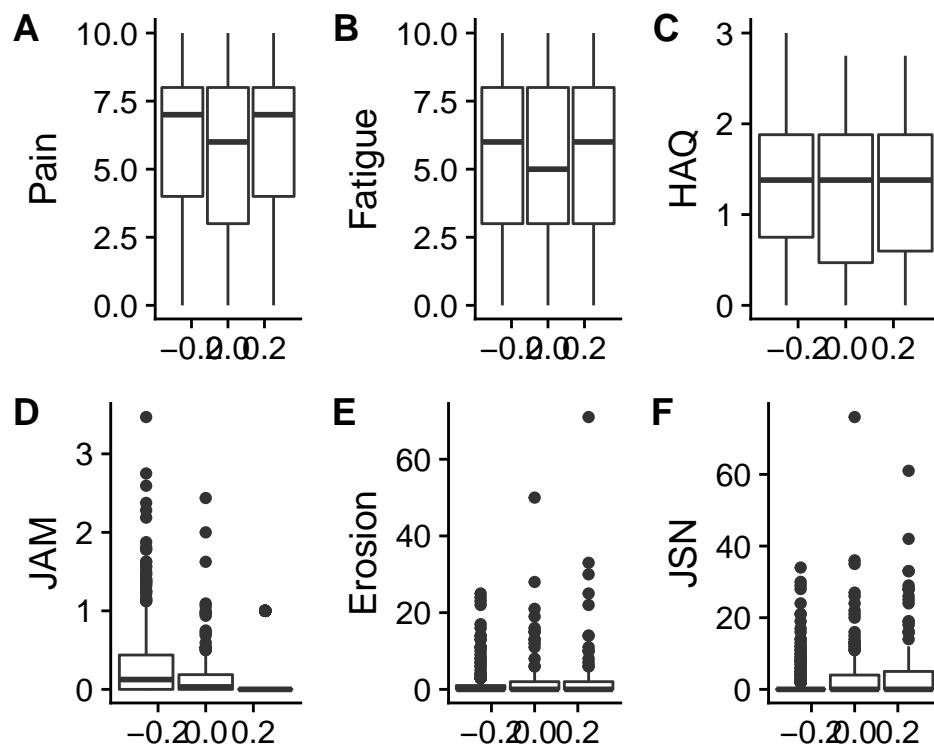


Table 1 Summary Statistics for the Six Outcome Variables

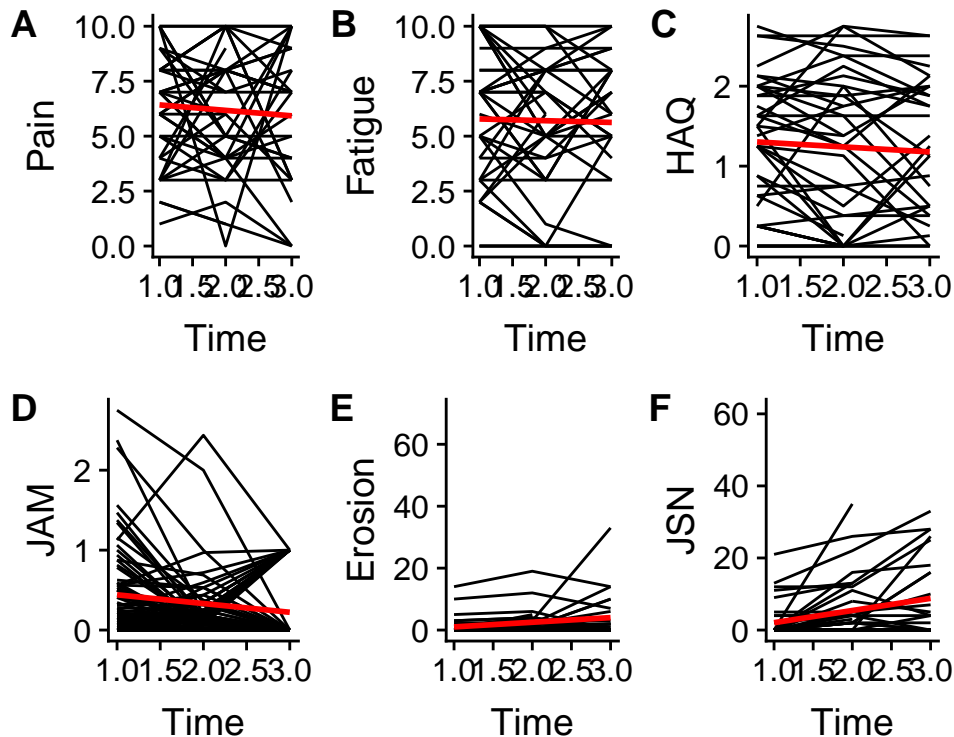
	Time: 1 (N = 354)	Time: 2 (N = 354)	Time: 3 (N = 354)
Pain			
min	0	0	0
median	7	6	7
mean (sd)	351; 6.04 (2.99)	222; 5.68 (3.12)	171; 5.84 (3.02)
max	10	10	10
Fatigue			
min	0	0	0
median	6	5	6
mean (sd)	351; 5.73 (3.18)	221; 5.21 (3.24)	172; 5.62 (3.31)
max	10	10	10
HAQ			
min	0	0	0
median	1.38	1.38	1.38
mean (sd)	353; 1.35 (0.76)	224; 1.22 (0.81)	176; 1.25 (0.78)
max	3.00	2.75	2.75
JAM			
min	0	0	0
median	0.12500	0.03125	0.00000
mean (sd)	353; 0.33 (0.51)	227; 0.15 (0.30)	226; 0.24 (0.43)
max	3.46875	2.43750	1.00000
Erosion			

 & min	0	0	0	
 & median	0	0	0	
 & mean (sd)	302; 1.28 (3.69)	170; 2.07 (5.52)	106; 3.09 (8.84)	
 & max	3.46875	2.43750	1.00000	
JSN	 & ;	 & ;	 & ;	
 & min	0	0	0	
 & median	0	0	0	
 & mean (sd)	302; 1.77 (4.99)	170; 3.72 (8.57)	106; 5.88 (10.78)	
 & max	34	76	61	

Individual Trajectories

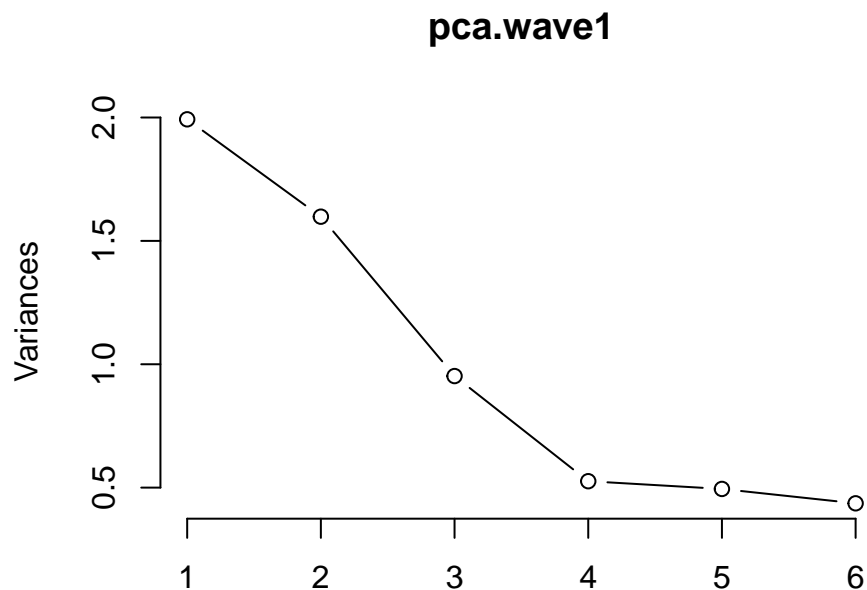
Individual trajectories were plotted by each outcome (Figure 2), in which the y-axis represents the corresponding score, and the x-axis represents the three time points (note that the time intervals were unequal, with 36 months and 24 months, respectively). The red line represents observed mean trajectory in each plot, which provides a sense of the overall trend. We found that there was a large amount of variabilities in the individual trajectories of all the outcomes. Specifically, for the two QOL measures as well as HAQ, the mean trajectories seem to represent an overall linear with a slight curvilinear trend. In the next section, Latent Growth Curve Models (LGCM) were used to model the repeated measures to determine the characteristics of the trajectories (e.g., intercepts, slopes, curvilinearity, etc.). Although individuals showed distinct patterns (e.g., some increased first and then decreased, others did the opposite), it is unclear from the plots whether there existed clusters of individuals with distinguishable types of trajectories. The variabilities in the individual trajectories were even more pronounced for JAM, Erosion, and JSN, where there seemed to be a large group that had low scores at first which changed slightly over time, and another small group that had high initial scores which changed even dramatically across time. This might suggest that there were clusters of people with distinguishable patterns of trajectories. Growth Mixture Models (GMMM) were used to determine the existence and number of clusters with respect to the characteristics of the trajectories.

Figure 2 Individual Trajectory for the Six Outcome Variables



Part II Dimension Exploration (PCA)

PCA (with log transformed observed outcomes) were conducted on the six RA phenotype measures at each time point, the results of those at the baseline were presented in this section (results at each time point were extremely similar). Before performing LGCM and GMM, it is important to understand the correlation of amongst the outcome measures. Such understanding helps to decide the factor structure of the outcomes that might be used in the longitudinal models, that is, how we might group variables that represent the same underlying factor together. By modeling the trajectory of the factor(s) instead of the observed variables, we might overcome the issues such that models can be hardly identified or the lack of power for parameter estimation due to the limited time point and large missing data.



Importance of components:

	PC1	PC2	PC3	PC4	PC5	PC6
Standard deviation	1.4116	1.2642	0.9758	0.72526	0.70331	0.66059
Proportion of Variance	0.3321	0.2664	0.1587	0.08767	0.08244	0.07273
Cumulative Proportion	0.3321	0.5985	0.7572	0.84483	0.92727	1.00000

Standard deviations (1, ..., p=6):

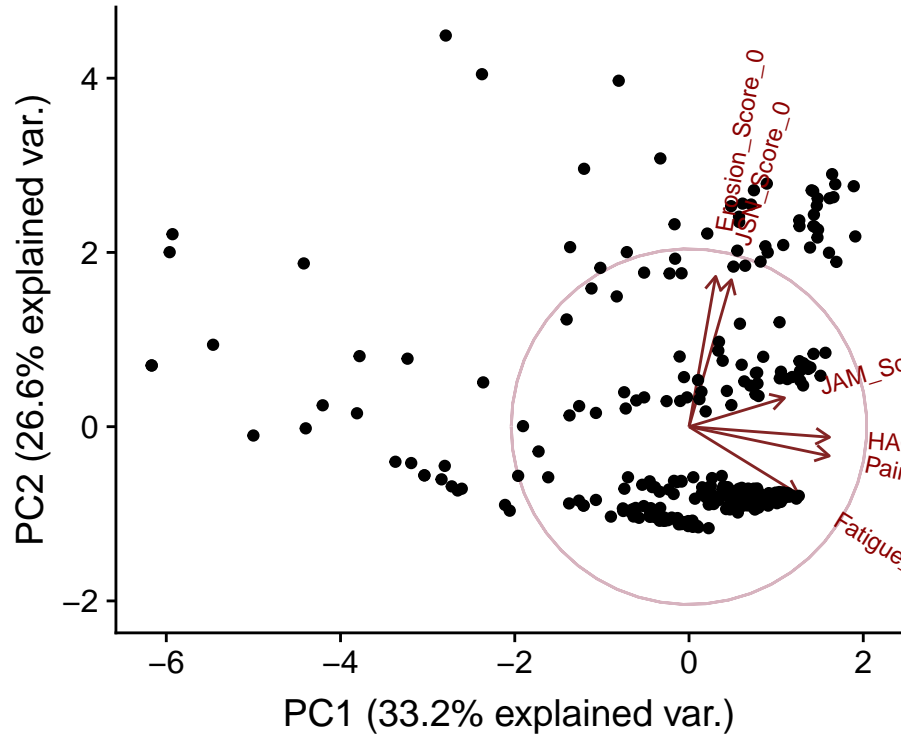
```
[1] 1.4116254 1.2641767 0.9757760 0.7252597 0.7033144 0.6605903
```

Rotation (n x k) = (6 x 6):

	PC1	PC2	PC3	PC4	PC5
Pain_Scale_0	0.5586411	-0.13013497	0.2146500	-0.1749295	0.7381231
Fatigue_Scale_0	0.4361134	-0.30043655	0.5070265	-0.3234190	-0.5973368
HAQ_Score_0	0.5597281	-0.04691145	-0.1918209	0.7717831	-0.1385793
JAM_Score_0	0.3803568	0.12805334	-0.7599318	-0.4676757	-0.1836854
Erosion_Score_0	0.1061087	0.66870066	0.2101295	0.1377053	-0.1912225
JSN_Score_0	0.1688176	0.65348592	0.1959679	-0.1773903	0.0940877

	PC6
Pain_Scale_0	0.22245061
Fatigue_Scale_0	-0.03249658
HAQ_Score_0	-0.18125991
JAM_Score_0	0.09473025
Erosion_Score_0	0.66475347
JSN_Score_0	-0.68244674

Figure 3 Scatterplot by A Two-dimensional PCA



Finding: Three to four PCs seemed to be appropriate and sufficient to represent the total variance of the outcomes (75.7% or 84.5%, respectively). The first PC was mainly composed of the two QOL measures as well as HAQ (55.9%, 43.6%, and 56.0%, respectively), the second PC was mainly composed of the two Radiographic progression measures (66.9% and 65.3%, respectively). The two Functional disability measures did not behave consistently, as HAQ loaded partly on to the first PC and primarily on the fourth PC, while JAM be the major component of the third PC (76.0%). This result was later confirmed by Confirmatory Factor analyses (see Appendix for all CFA results). The situation of the two Functional disability measures made it challenging to decide whether to group the two together or remain separate in the longitudinal analyses. I thus tried multiple approaches in the LGCM and GMM regarding these two measures.

Part III Latent Growth Curve and Growth Mixture Models (LGCM and GMM)

This section is organized as follows. First, results are broke down by each RA phenotype outcome measure: A. QOL, B. Functional disability, and C. Radiographic progression. Within each part, linear and/or nonlinear LGCMs are presented first. These series of models were tested to determine the shape (linear, quadratic, latent basis, or piecewise) of the trajectory. Although models using the observed variables and the factor were both conducted, part A and C only present those with factor, because the two measures in corresponding categories were previously confirmed to represent one underlying factor. Whereas part B contains models with the observed variables seperately, because the measures of that category did not consistently represent one underlying factor. In what is followed, GMM with one or two classes based on the final LGCM model is presented, which were conducted to investigate the existence of clusters w.r.t, the characteristics of the shape.

A. Quality of Life

Preliminary LGC Models for Pain and Fatigue

Linear, Quadratic, and latent basis LGCM were conducted for pain and fatigue, respectively. The Quadratic LGCM was under-identified due to exceeding number of parameters. The results were similar for the two measures, in which latent basis LGCM suggested that a curvilinear shape of trajectory, i.e., a slight decelerating decrease trend, was the most appropriate for both measures. CFA confirmed PCA result that the two measures represented by one factor fitted consistently well over time. With more variables for a LGCM on the factor, the quadratic LGCM on the factor can be identified, this is because there was more degree of freedom with more variables in the model (six instead of three). Hence, the final model for QOL outcome is a nonlinear LGCM with a quadratic term on the QOL outcome (see Path Diagram 1 below).

Path Diagram 1 Final LGCM for Pain and Fatigue: A Quadratic LGCM on the QOL Factor

Interpretation: Overall, the mean initial status (i.e., intercept) of the QOL was 6.015, the rate of change (i.e., slope) was -0.348 and the change in the rate of change (i.e., the quadratic term) was 0.056. This means that on average patients experienced a 0.348 units decrease per year in their pain and fatigue, but their average score was not decreasing at a constant rate, but instead had a deceleration rate of 0.056 per year over the course of the five years of study. In other words, the decreasing trend was fading as unfolded in the passage of time.

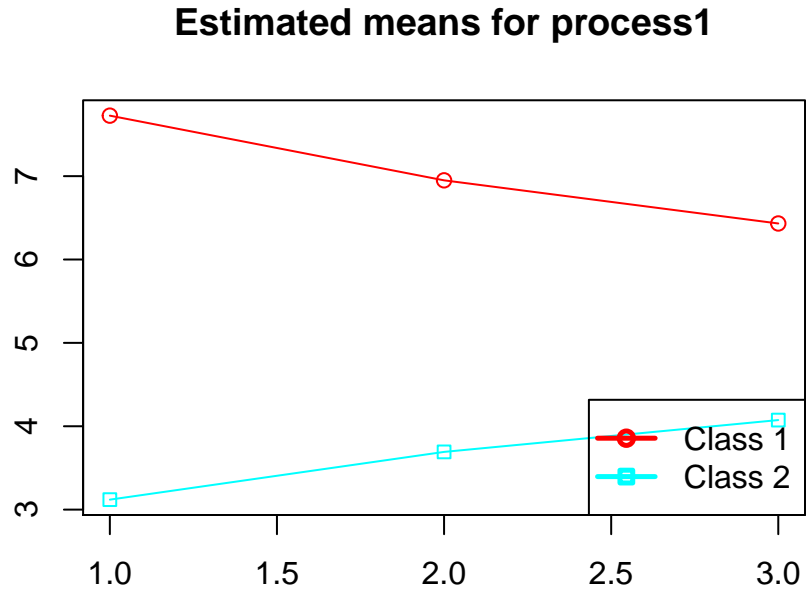
GMM Models for Quality of Life

In order to investigate potential clusters of distinct mean trajectories, linear GMM was performed on the QOL factor (a quadratic or latent basis GMM was way under-identified). A two-class linear GMM with some assumptions (i.e., homogenous class structure) fit very well (see Path Diagram 2 below). Since this is already a nearly saturated models (using only three time points), any more complex models, such as a latent basis growth curve or one with a quadratic term, or adding a third class of growth curve, would not be identified even with strict assumptions. Hence, there was no way of fitting a nonlinear GMM or reaching maybe an optimal number of class for GMMs.

Path Diagram 2 Final Two-Class Linear GMM for QOL

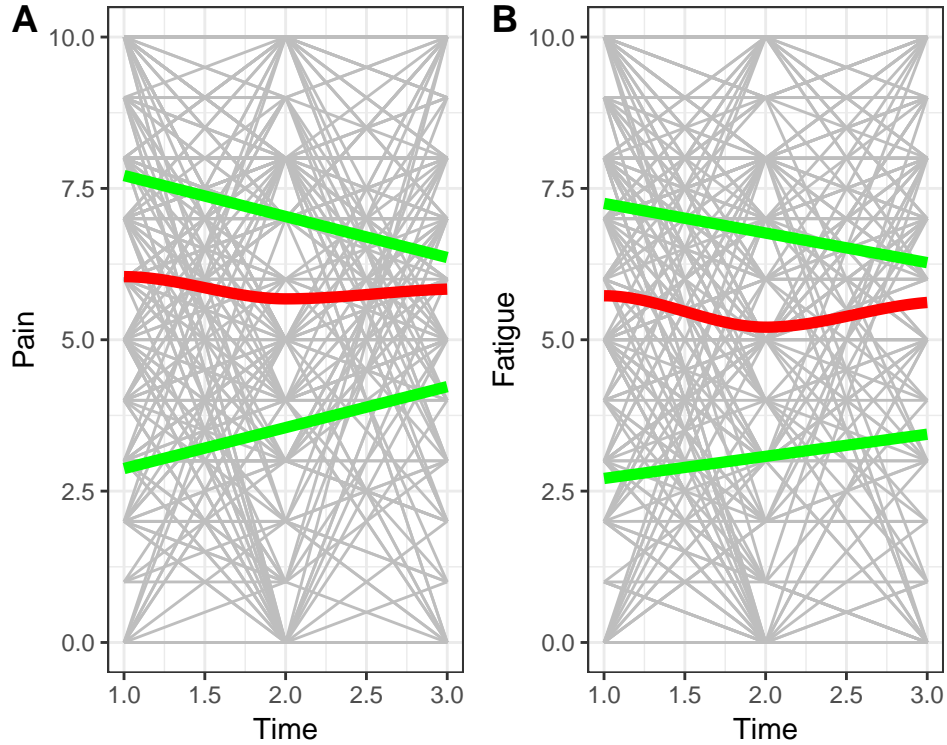
Interestingly, the two-class GMM revealed two distinct trajectories (see Figure 4 below): Class one (225, 64%) represented a “recovering group” that had higher means of the QOL factor, with a mean initial score of 7.726 (S.E. =.208, $p = .001$, within-class variance =1.369), which decreased at a rate of .259 (S.E. =.072, $p = .001$, within-class variance =0.181) per year over the five years of study; in comparison, Class 2 (126, 36%) represented a “progressing group” had lower means of the QOL factor, with a mean initial score of 3.120 (S.E. =.276, $p = .001$, within-class variance =1.369), which increased at a rate of 0.191 (S.E. =.092, $p = .039$, within-class variance =.181) per year (for more details, refer to the Appendix B. Model 2. The 2-Class Linear GMM of QOL). As a comparison, while the nonlinear LGCM provide a more accurate overall curvilinear trajectory for all individuals combined (see the red curve in Figure 5), it failed to distinguish the two patterns of trajectory. The final GMM was able to disentangle the two distinct trajectories, but the trade-off was to adopt more parsimonious within-class trajectories (see the green lines in Figure 5).

Figure 4 Estimated Means (Y-axis) Over Time (X-axis, row in Table) for Each QOL Latent Class (column in Table)



	Class 1 (225, 64%)	Class 2 (126, 36%)
Baseline	7.725864	3.119531
36 months	6.949774	3.692157
60 months	6.432381	4.073907

Figure 5 Contrasting A Two-class Linear GMM (green lines) vs. a Quadratic LGCM for Pain and Fatigue (red curve)



B. Functional disabilities

Preliminary LGC Models for HAQ and JAM

The same types of previous models for QOL measures (i.e., linear, quadratic, and latent basis LGCM) were conducted for HAQ and JAM, respectively, but only the linear LGCMs were converged (see Appendix B Diagram 1 for HAQ and Diagram 2 for JAM and corresponding details). This might be due to their zero-inflated nature of distributions. As a consequence, this made it impossible to test the curvilinearity in the trajectory. Although LGCM on factor of the two variables helped the models identified and converged, according to both PCA and CFA, the two measures can not be represented equally well by one factor. This is also consistent with what we observed from the individual trajectories for the outcomes, that the HAQ plot (similar to those of Pain and Fatigue) showed large between-person variabilities yet no easily distinguishable groups, whereas there seemed to be even larger variabilities and seemingly two types of trajectory for JAM. That is, a large group of people started low but showed a slight increase in their JAM scores across the study (especially during the second time interval), the other small group, in comparison, started with a high JAM score yet it immediately decreased over time. We suspected two clusters of people in terms of the JAM trajectory, and it would also be interesting to investigate potential clusters of HAQ trajectories. Due to the distinct trend, the LGCM and GMM analyses were done separately for these two measures.

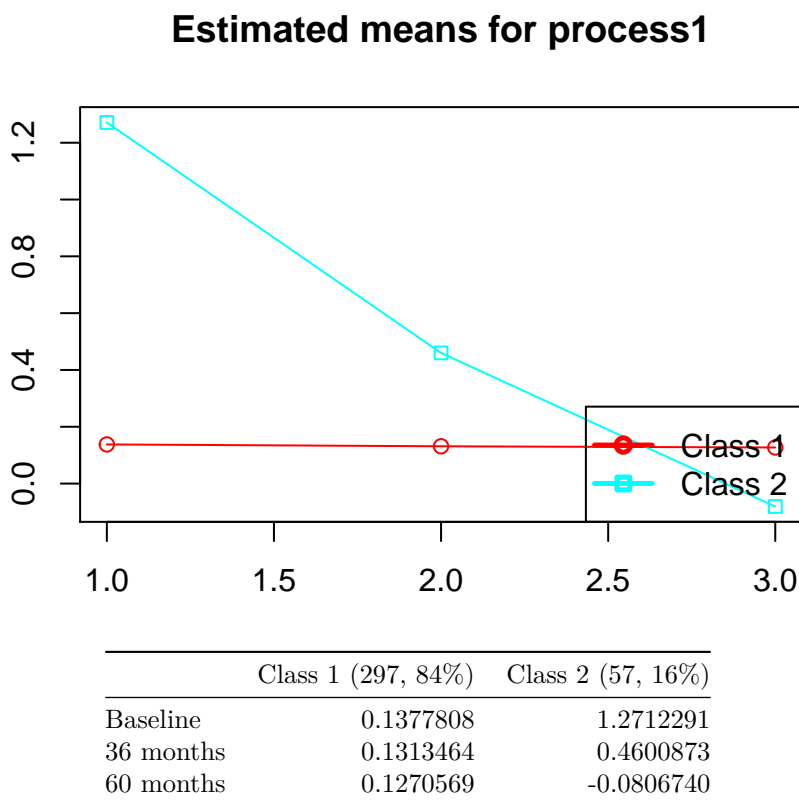
GMM Models for JAM

Not surprisingly, a two-class linear GMM of JAM fitted very well (see Path Diagram 3 for the model structure and Figure 6 for estimated means by class), revealing that most patients were alike and belonged to a “stable group” (297, 84%) with a mean initial score of .138 (S.E. = .017, $p = .001$) and a flat slope of $-.002$ (S.E. = .006, $p = .715$); in comparison, the remaining patients were “recovering group” (57, 16%) with a higher

mean initial score of 1.271 (S.E. =.106, $p = .001$) that decreased at an average rate of $-.270$ (S.E. =.029, $p = .001$) per year (for more details, refer to the Appendix B. The 2-Class Linear GMM of JAM).

Path Diagram 3 Final Two-Class Linear GMM for JAM

Figure 6 Estimated Means (Y-axis) Over Time (X-axis, row in Table) for Each JAM Latent Class (column in Table)



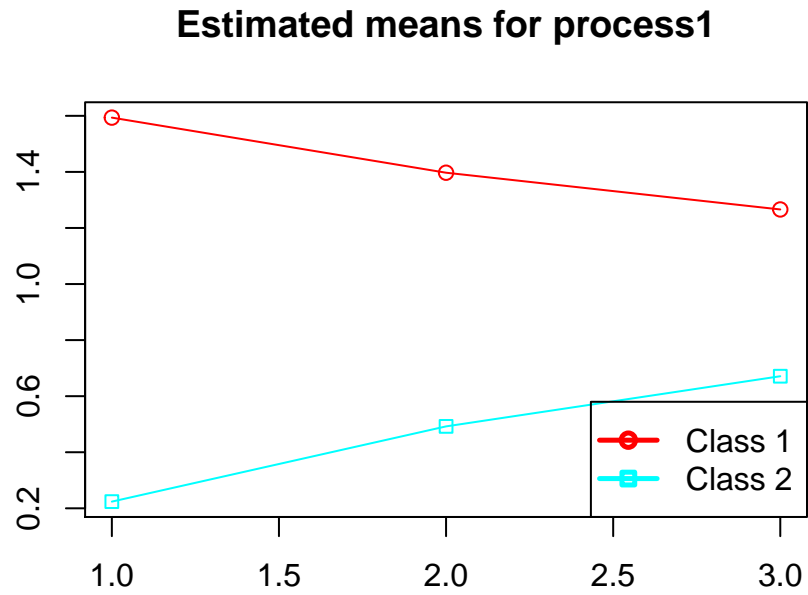
GMM Models for HAQ

Similarly, a two-class linear GMM of HAQ fitted very well (see Path Diagram 4 for the model structure and Figure 7 for estimated means by class), but the group-specific patterns were quite different (nearly reversed from those of JAM). Specifically, four in five (284) patients were “recovering group” with a mean initial score of 1.594 (S.E. =.077, $p = .001$, a large within-class variance of .357) and a slope of $-.066$ (S.E. =.011, $p = .001$, within-class variance of .015); the remaining one in five (69) patients were “progressing group” with a mean initial score of .224 (S.E. =.111, $p = .044$) that increase at an average rate of .089 (S.E. =.043, $p = .036$) per year (for more details, refer to the Appendix B. The 2-Class Linear GMM of HAQ).

As a comparison, Figure 8 provided a contrast of the two-class GMMs vs. linear LGCM for both functional disability outcomes plotted on the individual trajectories, respectively. The distinct classes validated the separate approach for this type of phenotype outcomes.

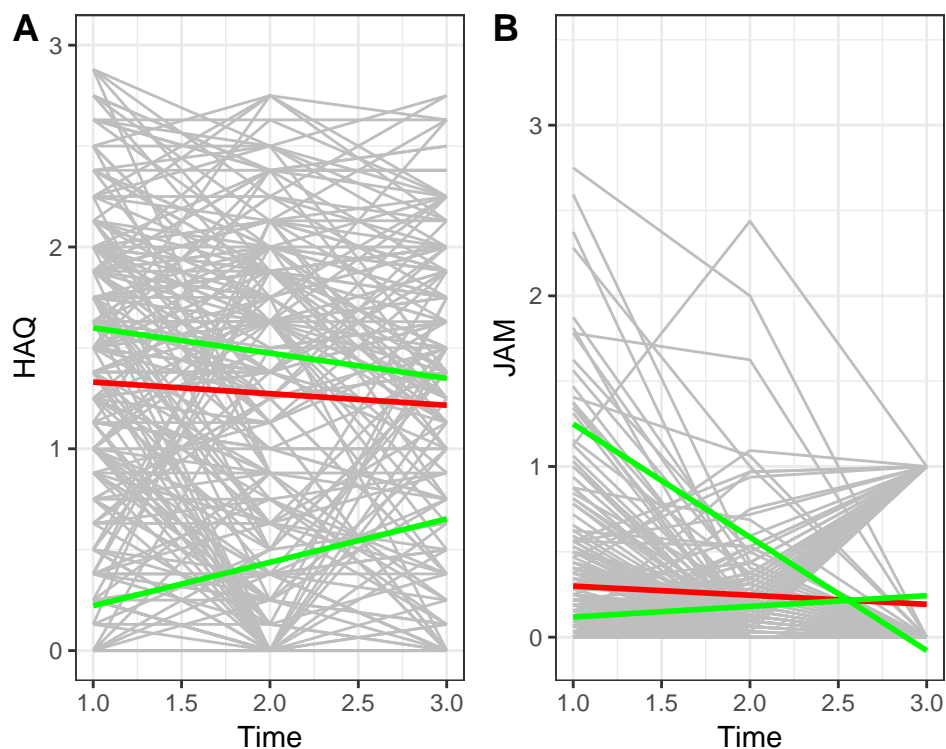
Path Diagram 4 Final Two-Class Linear GMM for HAQ

Figure 7 Estimated Means (Y-axis) Over Time (X-axis, row in Table) for Each HAQ Latent Class (column in Table)



	Class 1 (284, 80%)	Class 2 (69, 20%)
Baseline	1.593625	0.2238819
36 months	1.397058	0.4922567
60 months	1.266013	0.6711733

Figure 8 Contrast A Two-class Linear GMM (green lines) vs. A Linear LGCM (red line)



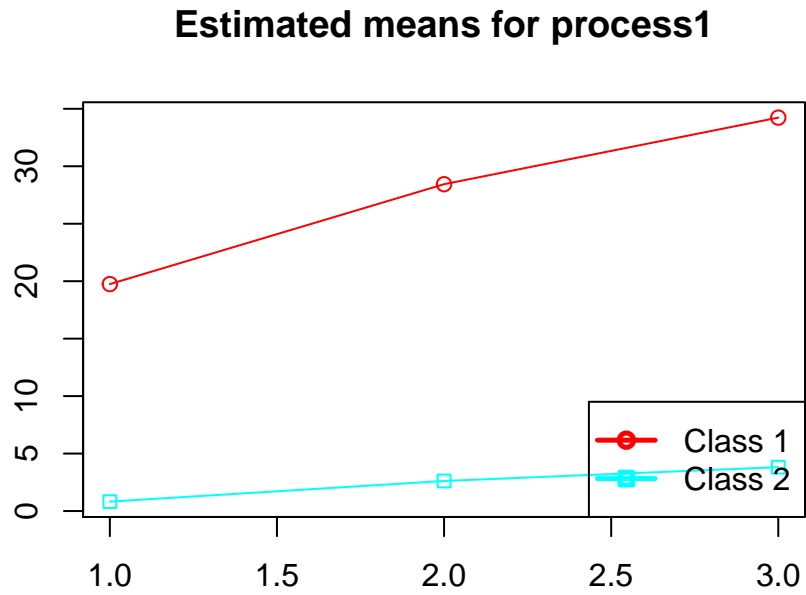
C. Radiographic of Joint Damage

LCM and GMM of Erosion and JSN

Similar to the QOL measures, these two measures behave very similarly, and a CFA confirmed that they represented the same factor. LCM recovered a linear trajectory for either measure (see Appendix for preliminary models). GMM revealed two types of the trajectories of the Radiographic factor (see Path Diagram 5 for the model structure and Figure 9 for estimated means by class). The majority of patients fell into the first group (293, 94%) who started low and experienced a slight increase in joint damage, called “slow progressors”; while the other small group (18, 6%) was characterized by very high initial scores and deteriorated at an even steeper rate, hence called “high and fast progressors”. Again, as a comparison, we provided in Figure 10 a contrast of this two-class GMMs vs. linear LGCM for both Radiographic progression outcomes plotted on the individual trajectories, respectively.

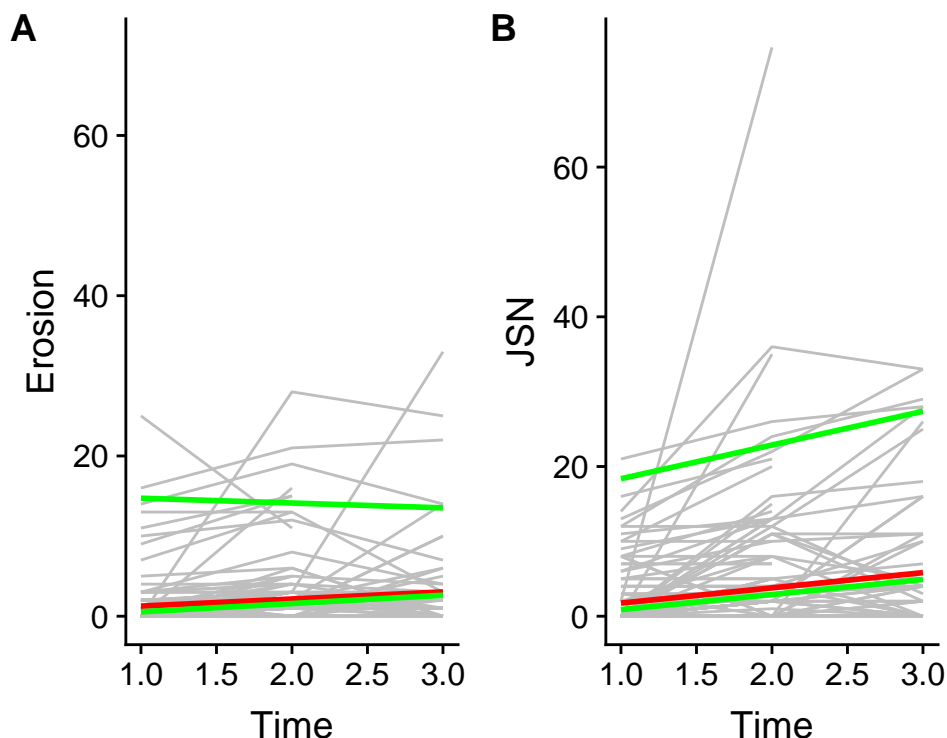
Path Diagram 5 Final Two-Class Linear GMM for Radiographic of Joint Damage

Figure 9 Estimated Means (Y-axis) Over Time (X-axis, row in Table) for Each Latent Class (column in Table)



	Class 1	Class 2
Baseline	19.74671	0.8192309
36 months	28.43772	2.6162388
60 months	34.23172	3.8142440

Figure 10 Contrast A Two-class Linear GMM (green lines) vs. A Linear LGCM (red line)



Part IV Latent Profile and Latent Transition Analysis (LPA and LTA)

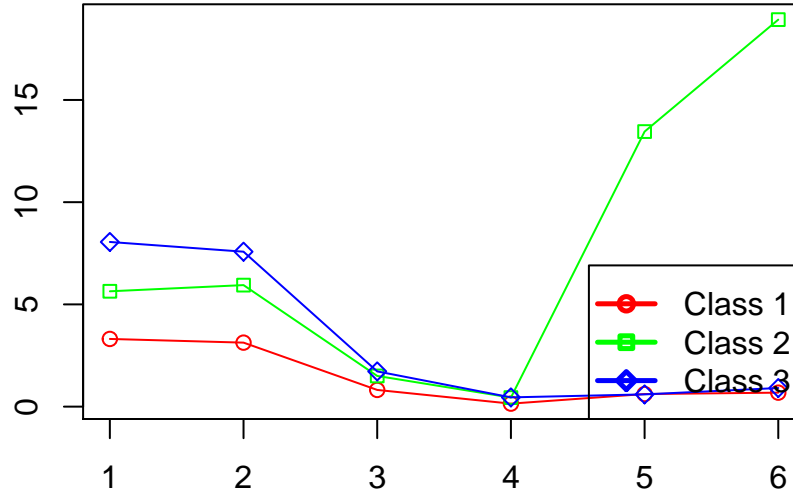
Previous section using GMM to cluster patients based on characteristics of their univariate trajectory on each phenotype measure (i.e., clustering by time vector), which answered the question as to whether there were qualitatively different progression rate in QOL, Functional Disabilities, and Radiographic of Joint Damage, separately. However, little is known whether there is distinct multivariate phenotypic patterns of experience at a given stage (i.e., clustering by response vector). As research has showed that RA experience is largely heterogeneous, we suspected that there might be distinguishable clusters of patients with respect to their manifest levels and response patterns of these phenotypes at a given time. We used LPA for such purpose, which can identify latent classes or profiles of patients in the way that those within each class or profile share large similarity in their experiences (i.e., manifest levels and response patterns of phenotypes) yet as a whole they differ substantially from those in another cluster. After latent profiles were established for each measured stage, we used the longitudinal model (i.e., LTA) to model transition probability among profiles from one stage to the next.

A. Latent Profile Analysis at Baseline

Using information criteria (i.e., AIC, BIC) as well as cluster separation index (entropy value), the LPA model at each wave of multivariate data recovered a three-profile best-fitting model. Here we presented the baseline-specific latent profiles (see Figure 11 for profile means by phenotype for each class): Profile One (143, 40%) represented the “Low Risk” group that had the lowest mean scores across all phenotypes; Profile Two (16, 5%) was characterized by moderate QOL and HAQ mean scores, high mean JAM and extremely high mean Radiographic scores, hence called “High Joint Damage” profile; and Profile Three (195, 55%) by high QOL and HAQ mean scores and moderate JAM and Radiographic scores, we called “Poor QOL” profile.

Figure 11 Estimated Means (Y-axis) by Phenotypes (X-axis, row in Table) for Each Latent Profile (column in Table)

Estimated means for process1



	Class 1 (143,40%)	Class 2 (16,5%)	Class 3 (195,55%)
Pain(0-10)	3.3102496	5.6407876	8.0539837
Fatigue(0-10)	3.1299410	5.9430108	7.5784559
HAQ(0-3)	0.8154331	1.5002719	1.7228335
JAM(0-4)	0.1454875	0.4422979	0.4515201
Erosion(0-140)	0.6145935	13.4527540	0.5980450
JSN(0-84)	0.6824146	18.9287643	0.9099796

B. Latent Transition Analysis

Further, we conducted LPA analyses on phenotype outcomes measured at 36 months and 60 months in the same manner, the corresponding three-class LPAs were the best fitting model (see Appendix for results of stage-specific latent profiles at later two benchmarks. We then performed a LTA analysis on phenotype outcomes across all waves, and examined the transitioning pattern. Figure 12 shows the estimated expected count in each latent class across three waves (this is calculated based on Bayesian rule of mostly likely class membership), and we observed an immediate increase in the number of Class 1, the “low-risk” group at early stage of development (i.e., from 132 at baseline to 150 at 36 month), as well as a progressive increase in the number of patients in the “High Joint Damage” group. On the contrary, there was a progressive decrease number of patients with “poor mental health”. On the one hand, this observation is broadly consistent with the previous findings from phenotypic trajectory groups, e.g., the majority of “recovering” QOL and HAQ patients as well as “slow progressor” and “fast progressor” in joint damage phenotypes. On the other hand, this new findings on time-specific grouping were able to pinpoint at exact what stage did this change or progression actually happen.

Note that the count were not exact what we observed in the sample, but instead estimated by FIML which counted for missing patients in a way that each missing patient was assigned into the most likely class based on the posterior probability estimation. But we returned to this in our limitation section in the Discussion section.

Figure 12 Counts by Class (color coded) in Each Wave (panels from left to right)

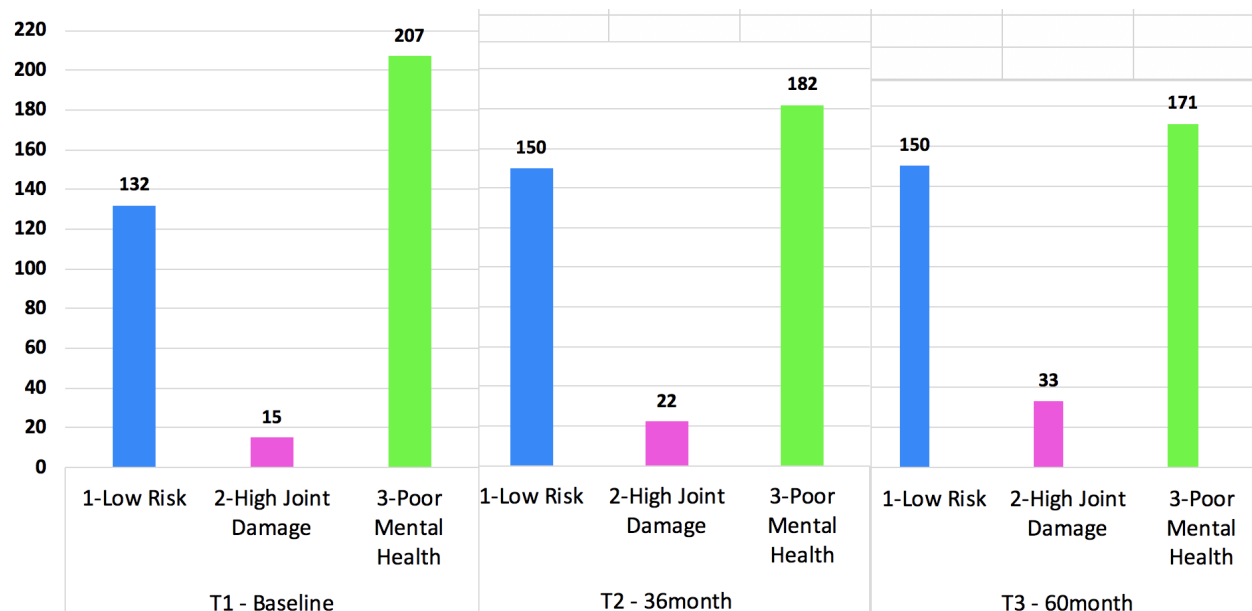


Figure 1: transition figure

variable	class	count	proportion
C1	1	132	0.37288
C1	2	15	0.04237
C1	3	207	0.58475
C2	1	150	0.42373
C2	2	22	0.06215
C2	3	182	0.51412
C3	1	150	0.42373
C3	2	33	0.09322
C3	3	171	0.48305

The count results provided the general change in number of patients in each profile, but how about the transition pattern and transition probability that explain the the change in number of patients? That is, from which group patients transitioned to which group or stayed in the same group, when did that happen? Similarly, if a patient was categorized in one profile, what the odd would be for him or her to transition to a certain profile versus stay at the current profile? Table 2 and table 3 were provided to answer these questions. For example, we found that from baseline to 36 month benchmark, there were around 30 patients made the transition from Class 3 “Poor QOL” to Class 1 “low-risk” (3-1-, the probability being 15.3%) although a few of them (6) transitioned to Class 2 “High Joint Damage” later (3-1-2), compared with other 10 who transitioned directly to Class 2 (probability of 5%) and stayed (3-2-2). Additionally, Class 2 “High Joint Damage” patients were likely to transition to “low risk” Class (2-1-, 3, probability = 19.1%). In comparison, neither were patients likely to transition from “low-risk” group to other higher risk groups at this period, which is very consistent with the increase number of patient in this group (equavilently, decreased number of patients in the other two groups). Further, from 36 month to 60 month benchmark, patients in Class 3 “Poor QOL” continued transition to either “low-risk” class (3-3-1, 16, 9.8%) or “High Joint Damage” class (3-3-2, 9, 5.5%). Interestingly, some patients previously remained in “low-risk” class now made their transition either to “High Joint Damage” class (1-1-2, 14, 12.8%) or to “Poor QOL” (1-1-3, 6, 5.1%), this may reflect the “progressing” group previous found in QOL and HAQ GMM groups and those “slow progressor” joint damage

GMM group.

Table 2 Counts in Each Transition Pattern Group (Class crossed by wave)

class.C1	class.C2	class.C3	count	proportion
1	1	1	101.54939	0.28686
1	1	2	14.11935	0.03989
1	1	3	6.48449	0.01832
1	2	1	0.16586	0.00047
1	2	2	3.86065	0.01091
1	2	3	0.00000	0.00000
1	3	1	0.34442	0.00097
1	3	2	0.26848	0.00076
1	3	3	3.99538	0.01129
2	1	1	2.76930	0.00782
2	1	2	0.52516	0.00148
2	1	3	0.18388	0.00052
2	2	1	0.37610	0.00106
2	2	2	14.38138	0.04063
2	2	3	0.00000	0.00000
2	3	1	0.00000	0.00000
2	3	2	0.00000	0.00000
2	3	3	0.00000	0.00000
3	1	1	24.51283	0.06925
3	1	2	5.52904	0.01562
3	1	3	1.35567	0.00383
3	2	1	1.03854	0.00293
3	2	2	9.21764	0.02604
3	2	3	0.00000	0.00000
3	3	1	16.18085	0.04571
3	3	2	8.99712	0.02542
3	3	3	138.14448	0.39024

Note: the count is not integer number because these are counts based on posterior probabilities The first three columns (class.C1, class.C2, class.C3) are class membership at each time point

Table 3 Transitional Probability by Class Across Waves

from	to	probability
C1.1	C2.1	0.934
C1.2	C2.1	0.191
C1.3	C2.1	0.153
C1.1	C2.2	0.031
C1.2	C2.2	0.809
C1.3	C2.2	0.050
C1.1	C2.3	0.035
C1.2	C2.3	0.000
C1.3	C2.3	0.797
C2.1	C3.1	0.820
C2.2	C3.1	0.054

from	to	probability
C2.3	C3.1	0.098
C2.1	C3.2	0.128
C2.2	C3.2	0.946
C2.3	C3.2	0.055
C2.1	C3.3	0.051
C2.2	C3.3	0.000
C2.3	C3.3	0.846

Discussion

Conclusion

Our findings suggested qualitatively distinct experience for RA patients. The two clustering approaches complementarily unraveled heterogeneous experience of patience from two different perspectives, and revealed latent group with phenotypic-specific developmental trajectory and stage-specific phenotypic pattern. We found that although the majority of patients tend to get worse in joint damage, some may recover in their symptoms and such recovery are likely to happen at early onset of the disease. On average, patients are likely to became more tolerant about the pain and fatigue they suffer throughout their first five year of disease. Such tolerance may carry over to make them more likely to feel less hopeless or stressful about general health condition, even though their sympotoms persist. Such finding is important to aid the screening of patients for targeted treatment, as well as to inform the most effective treatment at the most critical stage.

Limitation.

Missing data, time points. [TO BE ADDED]

Future Direction.

Covariates. [TO BE ADDED]

References

1. Bridges SL, Jr., Hughes LB, Mikuls TR, et al. Early rheumatoid arthritis in African-Americans: the CLEAR Registry. *Clin Exp Rheumatol.* 2003; 21: S138-45.
2. Grant S, Aitchison T, Henderson E, et al. A comparison of the reproducibility and the sensitivity to change of visual analogue scales, Borg scales, and Likert scales in normal subjects during submaximal exercise. *Chest.* 1999; 116: 1208-17.
3. Ramey DR, Raynauld JP and Fries JF. The health assessment questionnaire 1992: status and review. *Arthritis Care Res.* 1992; 5: 119-29.
4. Boini S and Guillemin F. Radiographic scoring methods as outcome measures in rheumatoid arthritis: properties and advantages. *Ann Rheum Dis.* 2001; 60: 817-27.

APPENDIX

The sections in Appendix is organized in corresponding order to the analysis reported in the result section.

Model 2. The 2-Class Linear GMM of QOL

Preliminary Linear LGCMs for HAQ and JAM (separately)

lavaan 0.6-3 ended normally after 31 iterations

Optimization method	NLMINB		
Number of free parameters	8		
	Used	Total	
Number of observations	353	354	
Number of missing patterns	4		
Estimator	ML		
Model Fit Test Statistic	2.596		
Degrees of freedom	1		
P-value (Chi-square)	0.107		

Parameter Estimates:

Information	Observed
Observed information based on	Hessian
Standard Errors	Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
i =~				
HAQ_Score_0	1.000			
HAQ_Score_36	1.000			
HAQ_Score_60	1.000			
s =~				
HAQ_Score_0	0.000			
HAQ_Score_36	3.000			
HAQ_Score_60	5.000			

Covariances:

	Estimate	Std.Err	z-value	P(> z)
i ~~				
s	-0.034	0.017	-1.976	0.048

Intercepts:

	Estimate	Std.Err	z-value	P(> z)
.HAQ_Score_0	0.000			
.HAQ_Score_36	0.000			
.HAQ_Score_60	0.000			
i	1.345	0.041	33.098	0.000
s	-0.037	0.010	-3.750	0.000

Variances:

	Estimate	Std.Err	z-value	P(> z)
.HAQ_Score_0	0.074	0.070	1.061	0.288
.HAQ_Score_36	0.223	0.033	6.699	0.000
.HAQ_Score_60	0.082	0.049	1.657	0.097
i	0.507	0.079	6.390	0.000
s	0.014	0.005	2.943	0.003

lavaan 0.6-3 ended normally after 48 iterations

Optimization method	NLMINB
Number of free parameters	8
Number of observations	354
Number of missing patterns	5
Estimator	ML
Model Fit Test Statistic	10.439
Degrees of freedom	1
P-value (Chi-square)	0.001

Parameter Estimates:

Information	Observed
Observed information based on	Hessian
Standard Errors	Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
i =~				
JAM_Score_0	1.000			
JAM_Score_36	1.000			
JAM_Score_60	1.000			
s =~				
JAM_Score_0	0.000			
JAM_Score_36	3.000			
JAM_Score_60	5.000			

Covariances:

	Estimate	Std.Err	z-value	P(> z)
i ~~				
s	-0.021	0.007	-3.001	0.003

Intercepts:

	Estimate	Std.Err	z-value	P(> z)
.JAM_Score_0	0.000			
.JAM_Score_36	0.000			
.JAM_Score_60	0.000			
i	0.307	0.027	11.429	0.000
s	-0.039	0.007	-5.521	0.000

Variances:

	Estimate	Std.Err	z-value	P(> z)
.JAM_Score_0	0.133	0.029	4.590	0.000

.JAM_Score_36	0.050	0.007	6.780	0.000
.JAM_Score_60	0.079	0.015	5.412	0.000
i	0.131	0.029	4.557	0.000
s	0.005	0.002	2.309	0.021

lavaan 0.6-3 ended normally after 62 iterations

Optimization method	NLMINB
Number of free parameters	20
Number of observations	354
Number of missing patterns	11
Estimator	ML
Model Fit Test Statistic	56.991
Degrees of freedom	7
P-value (Chi-square)	0.000

Parameter Estimates:

Information	Observed
Observed information based on	Hessian
Standard Errors	Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
i =~				
FUNC_0	1.000			
FUNC_36	1.000			
FUNC_60	1.000			
s =~				
FUNC_0	0.000			
FUNC_36	3.000			
FUNC_60	5.000			
FUNC_0 =~				
HAQ_Score_0	1.000			
JAM_Score_0	0.255	NA		
FUNC_36 =~				
HAQ_Score_36	1.000			
JAM_Score_36	0.124	NA		
FUNC_60 =~				
HAQ_Score_60	1.000			
JAM_Score_60	0.130	NA		

Covariances:

	Estimate	Std.Err	z-value	P(> z)
i ~~				
s	-0.038	NA		

Intercepts:

	Estimate	Std.Err	z-value	P(> z)
.HAQ_Score_0	0.000			
.JAM_Score_0	0.000			
.HAQ_Score_36	0.000			

.JAM_Score_36	0.000	
.HAQ_Score_60	0.000	
.JAM_Score_60	0.000	
i	0.716	NA
s	0.107	NA
FUNC_0	0.633	NA
FUNC_36	0.153	NA
FUNC_60	-0.070	NA

Variances:

	Estimate	Std.Err	z-value	P(> z)
.HAQ_Score_0	-0.032	NA		
.JAM_Score_0	0.214	NA		
.HAQ_Score_36	0.074	NA		
.JAM_Score_36	0.084	NA		
.HAQ_Score_60	0.048	NA		
.JAM_Score_60	0.100	NA		
i	0.521	NA		
s	0.015	NA		
FUNC_0	0.091	NA		
FUNC_36	0.151	NA		
FUNC_60	0.030	NA		

Preliminary LGC Models for JSN and Erosion

lavaan 0.6-3 ended normally after 57 iterations

Optimization method	NLMINB	
Number of free parameters	8	
	Used	Total
Number of observations	311	354
Number of missing patterns	7	
Estimator	ML	
Model Fit Test Statistic	0.066	
Degrees of freedom	1	
P-value (Chi-square)	0.798	

Parameter Estimates:

Information	Observed
Observed information based on	Hessian
Standard Errors	Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
i =~				
Erosion_Scor_0	1.000			
Erosion_Scr_36	1.000			
Erosion_Scr_60	1.000			
s =~				
Erosion_Scor_0	0.000			

Erosion_Scr_36	3.000
Erosion_Scr_60	5.000

Covariances:

	Estimate	Std.Err	z-value	P(> z)
i ~~				
s	-3.798	1.841	-2.063	0.039

Intercepts:

	Estimate	Std.Err	z-value	P(> z)
.Erosion_Scor_0	0.000			
.Erosion_Scr_36	0.000			
.Erosion_Scr_60	0.000			
i	1.328	0.219	6.072	0.000
s	0.355	0.097	3.668	0.000

Variances:

	Estimate	Std.Err	z-value	P(> z)
.Erosion_Scor_0	-13.917	5.475	-2.542	0.011
.Erosion_Scr_36	7.142	2.935	2.433	0.015
.Erosion_Scr_60	26.985	7.831	3.446	0.001
i	29.192	5.919	4.932	0.000
s	2.445	0.406	6.019	0.000

lavaan 0.6-3 did NOT end normally after 486 iterations
 ** WARNING ** Estimates below are most likely unreliable

Optimization method	NLMINB	
Number of free parameters	9	
	Used	Total
Number of observations	311	354
Number of missing patterns	7	
Estimator	ML	
Model Fit Test Statistic	NA	
Degrees of freedom	NA	
P-value	NA	

Parameter Estimates:

Information	Observed
Observed information based on	Hessian
Standard Errors	Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
i =~				
Erosion_Scor_0	1.000			
Erosion_Scr_36	1.000			
Erosion_Scr_60	1.000			
s =~				
Erosion_Scor_0	0.000			
Erosion_Scr_36	1.000			

Erosion_Scr_60	1.000	NA
----------------	-------	----

Covariances:

	Estimate	Std.Err	z-value	P(> z)
i ~~				
s	75185.335	NA		

Intercepts:

	Estimate	Std.Err	z-value	P(> z)
.Erosion_Scor_0	0.000			
.Erosion_Scr_36	0.000			
.Erosion_Scr_60	0.000			
i	1.376	NA		
s	1.035	NA		

Variances:

	Estimate	Std.Err	z-value	P(> z)
.Erosion_Scor_0	75182.850	NA		
.Erosion_Scr_36	0.494	NA		
.Erosion_Scr_60	44.115	NA		
i	-75167.574	NA		
s	-75168.043	NA		

lavaan 0.6-3 ended normally after 62 iterations

Optimization method	NLMINB	
Number of free parameters	8	
	Used	Total
Number of observations	311	354
Number of missing patterns	7	
Estimator	ML	
Model Fit Test Statistic	0.057	
Degrees of freedom	1	
P-value (Chi-square)	0.811	

Parameter Estimates:

Information	Observed
Observed information based on	Hessian
Standard Errors	Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
i =~				
JSN_Score_0	1.000			
JSN_Score_36	1.000			
JSN_Score_60	1.000			
s =~				
JSN_Score_0	0.000			
JSN_Score_36	3.000			
JSN_Score_60	5.000			

Covariances:

	Estimate	Std.Err	z-value	P(> z)
i ~~				
s	2.609	1.769	1.475	0.140

Intercepts:

	Estimate	Std.Err	z-value	P(> z)
.JSN_Score_0	0.000			
.JSN_Score_36	0.000			
.JSN_Score_60	0.000			
i	1.807	0.286	6.325	0.000
s	0.867	0.163	5.304	0.000

Variances:

	Estimate	Std.Err	z-value	P(> z)
.JSN_Score_0	3.891	6.699	0.581	0.561
.JSN_Score_36	18.191	5.333	3.411	0.001
.JSN_Score_60	-4.605	9.779	-0.471	0.638
i	21.062	6.896	3.054	0.002
s	4.322	0.778	5.553	0.000

lavaan 0.6-3 ended normally after 97 iterations

Optimization method	MLMINB		
Number of free parameters	9		
	Used	Total	
Number of observations	311	354	
Number of missing patterns	7		
Estimator	ML		
Model Fit Test Statistic	0.000		
Degrees of freedom	0		

Parameter Estimates:

Information	Observed
Observed information based on	Hessian
Standard Errors	Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
i =~				
JSN_Score_0	1.000			
JSN_Score_36	1.000			
JSN_Score_60	1.000			
s =~				
JSN_Score_0	0.000			
JSN_Score_36	1.000			
JSN_Score_60	1.606	0.247	6.494	0.000

Covariances:

	Estimate	Std.Err	z-value	P(> z)
i ~~				

s	8.505	6.564	1.296	0.195
---	-------	-------	-------	-------

Intercepts:

	Estimate	Std.Err	z-value	P(> z)
.JSN_Score_0	0.000			
.JSN_Score_36	0.000			
.JSN_Score_60	0.000			
i	1.804	0.286	6.306	0.000
s	2.674	0.580	4.611	0.000

Variances:

	Estimate	Std.Err	z-value	P(> z)
.JSN_Score_0	4.517	7.670	0.589	0.556
.JSN_Score_36	16.558	8.714	1.900	0.057
.JSN_Score_60	-0.519	19.237	-0.027	0.978
i	20.436	7.829	2.610	0.009
s	40.013	8.695	4.602	0.000

lavaan 0.6-3 ended normally after 73 iterations

Optimization method	MLMINB	
Number of free parameters	14	
Number of equality constraints	2	
	Used	Total
Number of observations	311	354
Number of missing patterns	7	
Estimator	ML	
Model Fit Test Statistic	109.936	
Degrees of freedom	15	
P-value (Chi-square)	0.000	

Parameter Estimates:

Information	Observed
Observed information based on	Hessian
Standard Errors	Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
i =~				
Radio_0	1.000			
Radio_36	1.000			
Radio_60	1.000			
s =~				
Radio_0	0.000			
Radio_36	3.000			
Radio_60	5.000			
Radio_0 =~				
Ersn_Scr_0	1.000			
JSN_Scor_0 (j)	1.517	0.086	17.689	0.000
Radio_36 =~				
Ersn_Sc_36	1.000			

JSN_Scr_36 (j)	1.517	0.086	17.689	0.000
Radio_60 =~				
Ersn_Sc_60	1.000			
JSN_Scr_60 (j)	1.517	0.086	17.689	0.000

Covariances:

	Estimate	Std.Err	z-value	P(> z)
i ~~				
s	1.923	0.466	4.124	0.000

Intercepts:

	Estimate	Std.Err	z-value	P(> z)
Radio_0	0.000			
Radio_36	0.000			
Radio_60	0.000			
.Erosion_Scor_0	0.000			
.JSN_Score_0	0.000			
.Erosion_Scr_36	0.000			
.JSN_Score_36	0.000			
.Erosion_Scr_60	0.000			
.JSN_Score_60	0.000			
i	1.253	0.195	6.429	0.000
s	0.501	0.087	5.741	0.000

Variances:

	Estimate	Std.Err	z-value	P(> z)
Radio_0	0.000			
Radio_36	0.000			
.JSN_Score_60	0.000			
.Erosion_Scor_0	5.594	0.732	7.644	0.000
.JSN_Score_0	6.538	1.413	4.628	0.000
.Erosion_Scr_36	8.388	2.516	3.334	0.001
.JSN_Score_36	36.560	7.776	4.701	0.000
.Erosion_Scr_60	31.639	4.380	7.223	0.000
i	8.822	1.202	7.342	0.000
s	0.775	0.257	3.017	0.003
Radio_60	16.730	9.791	1.709	0.087

lavaan 0.6-3 ended normally after 84 iterations

Optimization method	NLMINB	
Number of free parameters	20	
Number of equality constraints	2	
	Used	Total
Number of observations	311	354
Number of missing patterns	7	
Estimator	ML	
Model Fit Test Statistic	105.222	
Degrees of freedom	9	
P-value (Chi-square)	0.000	

Parameter Estimates:

Information	Observed
Observed information based on	Hessian
Standard Errors	Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
i =~				
Radio_0	1.000			
Radio_36	1.000			
Radio_60	1.000			
s =~				
Radio_0	0.000			
Radio_36	3.000			
Radio_60	5.000			
q =~				
Radio_0	0.000			
Radio_36	9.000			
Radio_60	25.000			
Radio_0 =~				
Ersn_Scr_0	1.000			
JSN_Scor_0 (j)	1.508	0.092	16.379	0.000
Radio_36 =~				
Ersn_Sc_36	1.000			
JSN_Scr_36 (j)	1.508	0.092	16.379	0.000
Radio_60 =~				
Ersn_Sc_60	1.000			
JSN_Scr_60 (j)	1.508	0.092	16.379	0.000

Covariances:

	Estimate	Std.Err	z-value	P(> z)
i ~~				
s	2.441	0.912	2.676	0.007
q	-0.152	0.224	-0.676	0.499
s ~~				
q	0.020	0.272	0.074	0.941

Intercepts:

	Estimate	Std.Err	z-value	P(> z)
Radio_0	0.000			
Radio_36	0.000			
Radio_60	0.000			
.Erosion_Scor_0	0.000			
.JSN_Score_0	0.000			
.Erosion_Scr_36	0.000			
.JSN_Score_36	0.000			
.Erosion_Scr_60	0.000			
.JSN_Score_60	0.000			
i	1.264	0.195	6.473	0.000
s	0.325	0.227	1.429	0.153
q	0.048	0.057	0.837	0.403

Variances:

	Estimate	Std.Err	z-value	P(> z)
--	----------	---------	---------	---------

Radio_0	0.000			
.Erosion_Scor_0	5.553	0.731	7.592	0.000
.JSN_Score_0	6.795	1.388	4.896	0.000
.Erosion_Scr_36	10.135	2.537	3.995	0.000
.JSN_Score_36	35.687	6.380	5.594	0.000
.Erosion_Scr_60	30.519	5.921	5.155	0.000
.JSN_Score_60	2.679	10.025	0.267	0.789
i	8.721	1.159	7.524	0.000
s	0.118	0.896	0.132	0.895
q	0.048	0.078	0.612	0.540
Radio_36	0.095	0.659	0.144	0.886
Radio_60	0.056	0.089	0.624	0.532