

## REVIEWS

## Mixed effect frameworks in the analysis of longitudinal data

Anupama KR<sup>1\*</sup>, Chandrashekara S<sup>2</sup><sup>1</sup>Research Associate, ChanRe Rheumatology and Immunology Center, Basaweswaranagar, Bangalore, India<sup>2</sup>Managing Director and Consultant Rheumatologist, ChanRe Rheumatology and Immunology Center, Basaweswaranagar, Bangalore, India

Longitudinal research generates data with correlated measurements and clustered data structures. The main interest in longitudinal studies is to find the change in outcome over time and to analyze the influence of predictor variables on the outcome. Model-based approach is a powerful tool to integrate the multiple measurements and complexity of such data. Mixed effect models provide the framework to identify the role of individual differences in responses, while incorporating information from multifactorial measures at the individual and group levels to advance our understanding of the underlying components influencing response. This methodology is flexible, could be extended to different data patterns and computes more accurate and stable estimates. Using a multilevel modeling approach, the hierarchical structure can be explicitly modeled. With the availability of standard statistical software for both classical and Bayesian approaches, its applicability has increased in various fields such as genome-wide association studies, understanding of disease/recovery process, disease-marker association and behavioral studies to understand personality traits and health outcomes. This review seeks to focus on the methodological approaches to model multiple failure time data, conceptual issues, arising due to correlation, heterogeneity, and clustering and estimation procedures. The article also emphasizes on their application in data analysis especially in immunology and rheumatology studies.

### Introduction

Most of the observational studies found in literature are cross-sectional studies. However, in clinical studies, researchers may confront scenarios where it is necessary to follow-up patients at multiple time points and assess outcome variables and covariates. Such study designs involving multiple observations over a period of time are categorized under the domain of longitudinal studies, in contrast to cross-sectional studies.<sup>1</sup> In a cross-sectional study, observations for each sample are collected only once. Such observations may all occur at the same time point or at different time points across individuals. There is a growing interest in longitudinal studies for estimating and drawing inferences about changes brought about by targeted interventions in clinical trials and where patients are monitored for prolonged period such as in transplantation, cancer, and autoimmune diseases.

Multiple failure time (MFT) data are frequently encountered in subjects at risk of repeated occurrence of the same or different types of events and in clustered

data structures (e.g. patients within the same hospital in a multicenter trial, animals within the same litter). It is also seen routinely in disease processes where a single primary endpoint may not be sufficient to measure clinical outcome and multiple failure measurements need to be considered, for example, patients with repeated events requiring multiple visits to the clinician. It is also seen in patients on prolonged treatment or in clinical trials where patients are monitored for the risk of occurrence of events during the recovery process, in spite of the treatment.<sup>2</sup>

In medical applications, failure time is attributed to the occurrence of a pre-specified event such as time to disease or symptom occurrence, recurrence, remission or death. An important feature of longitudinal studies is to distinguish datasets of events with distinct ordering from those with unordered datasets. Ordered multivariate failure time data or recurrent event time data are commonly observed in recurrence of tumors, epileptic seizures, asthma attacks, migraines, repeated episodes of infections, multiple infections after surgery, myocardial infarctions, injuries,

and repeated admissions to the hospital. Symptoms developed by patients with connective tissue diseases are an example for unordered failure time data.<sup>3, 4</sup>

When a given subject contributes multiple events, the failure times are correlated within the subject. These events are not independent, as the occurrence of one may increase or decrease the risk of another event, conditional on given measured covariates.<sup>2</sup> Similarly, individuals may be grouped at multiple levels and each level may contribute variables. Example of data structures with multiple levels are given in figure 1. Individuals within a cluster are similar compared to individuals from other clusters. Multiple repeated events and grouping at multiple levels induces statistical dependence between observations within individuals and/or clusters that may not be captured by covariates, violating a key assumption of independence of all observations. Thus, models, which compensate for the biases introduced by correlated data structure, should be used for the analysis of longitudinal data.<sup>5, 6, 7</sup> Inferences based on conventional statistical methods are derived from survival analysis for time to the first event or overall survival time. These methods do not utilize the subsequent events, and hence, do not consider the correlation structure between multiple failure times in the data. The subsequent events provide information on the disease/recovery process, and are worth modelling to get a more precise understanding about disease management strategies. The methods of choice are statistical tools that analyze the association of outcome, treatment, and covariates with the length of time that account for dependency in the data. Established correlation models are available to fit such type of data.<sup>8, 9</sup>

The purpose of this paper is to review different approaches for data analysis of the subjects experiencing multiple failures and having a clustered data structure.

### **Basic aspects of longitudinal data analysis**

An important aspect of failure time data is that each individual under study may not experience the anticipated event during the study period, that is, the exact failure time is not known. The study subject for whom no failure time is available is censored. If the study ends before the observation of the event, it is referred as right censored data. If the events are known to have occurred before the study entry, but the times of these events are unknown, then it is taken as left censored data.

Subjects in the study will either have an event or will be right censored (those who contribute to the model until the

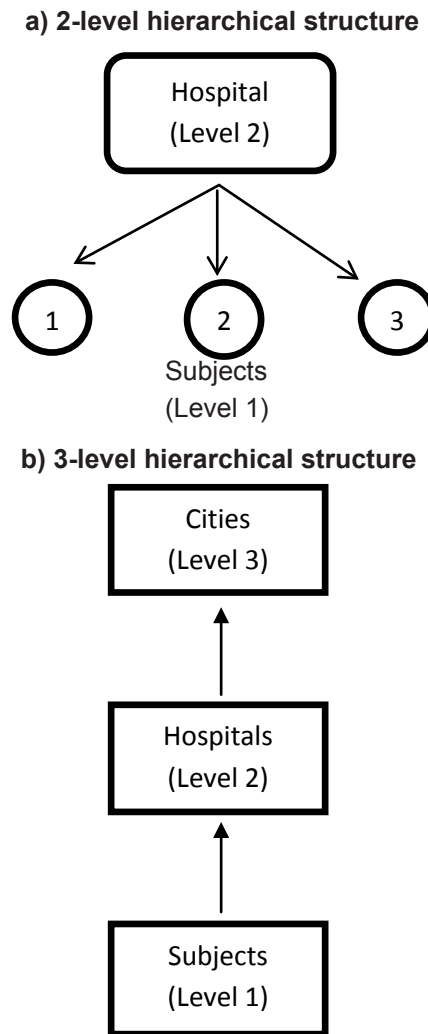
end of the study or withdrawal). Excluding the subjects, who do not experience the anticipated failure events during the study period, leads to loss of data and introduce bias due to reduced sample size. Hence, these observations are retained by assuming that censoring is non-informative, that is, event times are independent of censoring mechanism. However, censoring is informative when individuals are selectively withdrawn from the sample, because they are more or less likely to experience an event, such as dropout or death due to event process. Then the non-informative censoring assumption is violated, and subjects in the sample no longer represent the study population. The non-informative censoring is an important assumption required for the validity of the statistical analysis. Suitable modification of the methodologies is done to account for informative censoring mechanisms. In some patients, only intermittent follow-up data is available and the complete information from the periods between visits is not observed, such observations are said to be interval censored.

Another aspect is that observations are truncated. In truncation, subjects do not appear in the data because they have not been observed. Left truncation refers to a process when a subject is not observed from the beginning of the study, but only conditionally, on having survived until a later point in time. Exclusion of patients with short treatment history is considered as a left-truncated data, as inclusion is conditioned on length of compliance. If the entry into study depends on event of interest, then it is termed as right truncation. For example, in an AIDS outbreak study, patients with AIDS are recruited to assess the time from infection with HIV to development of AIDS, where only information on subjects with AIDS is available. Other examples include study of patients based on the registries like a cancer registry, where entire study population has already experienced the event of interest.<sup>10</sup>

In most cases, the truncation and censoring mechanisms are assumed to be independent of the event process.<sup>1</sup> The presence of these features creates bias in data analysis and needs to be carefully considered and compensated, when constructing likelihood functions.

The goal of longitudinal studies is not only to examine the effects on the time until an event occurs, but also to assess the relationship between failure time and explanatory variables. Explanatory variables (also referred as covariates or independent variables) assess the impact of certain characteristics on the dependent variable, that is, covariates influence on the time-specific outcomes. Due

**Figure1: Examples of hierarchical structures**



Patients visiting the hospital are grouped at level 1 and the hospitals (multiple centers) at level 2. The hospitals that are in different cities are grouped at level 3.

to the longitudinal feature of the data gathering process, some covariates may serve as time-invariant while others as time-varying. Time-invariant variables, which do not change across time, include gender, race, and other personality traits. Time-varying variables have values that change over time such as age, level of a biomarker, etc.

Event history data can be considered as a series of time measurements. In each interval, we observe a response indicating whether an event has occurred. The underlying data distribution of the dependent variable has to be considered for appropriate model selection in model-based approaches. Event times are measured as a continuous time series or at discrete intervals. Continuous-time

approach is adopted in the former case and discrete-time in the latter.<sup>1</sup>

In MFT data, some individuals have a higher risk of experiencing an event than the others, and this cannot be explained by the observed covariates alone. The presence of unobserved, individual-specific risk factors leads to frailty or unobserved heterogeneity. This needs to be accounted for, as it introduces systemic variability or randomness to each unit and violates the assumption that all subjects are homogenous.<sup>1,11</sup>

### **Methodological approaches**

Model strategies that are commonly used to analyze

longitudinal data include: marginal models, mixed effects models and transition models.<sup>11</sup> Transition models are used when events in the disease/recovery process involve a finite number of states.<sup>12</sup> When the correlation among the observations is not of interest, marginal models can be utilized. It provides estimates for the fixed parameters in the model, and treats the existence of any random parameters as a necessary 'nuisance', without providing explicit estimates for them.<sup>13</sup> In mixed effects framework, both the fixed and random effects are modeled in the same analysis for the multiple failure data.<sup>14</sup> As this type of data has both correlated observations and hierarchy along with variability, the model allows each factor to have fixed and random effect parameters.<sup>15</sup>

which indicates the differences in the response among the levels or units and the random effects, their variances; allow both marginal and subject-specific inferences. The relationship between an independent and a dependent variable may be at a single level or multiple levels. It may differ between levels (random intercept) and/or by levels (random slope). Thus quantifying the heterogeneity at each level, in turn explains the hierarchical grouping of the data. In addition, interaction effects among the variables may be specified. Finally, correlations among measurements that are not fully described by the random intercepts and slopes are indicated by an error term. Thus, outcomes at the individual level are modeled in terms of both individual- and cluster-level variables, and simultaneously estimating and adjusting for the amount of correlation (non-independence due to repeated measurement on the same individual) present in the data. Multilevel modeling is ideal for modeling processes where variables at one level influence those at another level; it models the effects of group-level predictors while accounting for unobserved group and individual characteristics. Furthermore, these models allow varying number of subjects within each cluster as they make no assumption regarding cluster sample size.<sup>5, 6</sup>

### Classical approach

Mixed effects framework is closely related to a variety of model types, like the random coefficients models, random effects model, hierarchical or multilevel models (MLM), variance components models, nested models and split-plot designs. Mixed models are used to model both linear and non-linear relationships between dependent and independent variables, and can handle both balanced and unbalanced design data. Continuous data are analyzed by linear regression model with mixed effects, and binary or ordinal or categorical data by logistic or multinomial regression model with mixed effects.<sup>5, 9, 16</sup>

A mixed effects model is an extension of the general linear model that can specify additional random effects terms (Table 1). It estimates both fixed and random effect features. Fixed effects give the mean of each factor level,

To run a mixed effects model, the user must specify the variables with fixed and random effects, the nature of the hierarchy, and model selection. For each factor or explanatory variables, fixed effects are considered if it affects the average responses of all the subjects. Presence of unobserved heterogeneity in the variables contributes to the random effects.<sup>17</sup> Hence, differentiating the fixed and random explanatory variables is important.

**Table 1: A general linear mixed model equation**

<p>The linear mixed model for level 1 can be represented by the equation:  <math>Y = X\beta_0 + Z\beta_1 + \epsilon</math></p> <p>where, Y= dependent variable  X= fixed effect regressors  <math>\beta_0</math>= fixed effect parameter that is assumed to be the same for all individuals  Z= random effect regressors  <math>\beta_1</math>= random effects parameter that is assumed to vary from individual to individual  <math>\epsilon</math>= error term</p>
<p>This equation could be extended to categorical data, multiple explanatory factors, and to multiple levels by inclusion of additional terms.</p>

Fixed effect parameters for a mixed model are interpreted based on the nature of the explanatory variables. Fixed coefficients that have a corresponding random effect represent the mean of all the subjects, and each individual subject will have their own value for that coefficient. A random effect parameter estimate is considered as the magnitude of the variability of personal coefficients from the mean fixed effects coefficient. The fixed effect coefficients represent an average around which individual units may or may not vary. If it varies it suggests that there are unmeasured explanatory variables with random effects for each subject influencing the disease/recovery process or differences in coefficient could also be due to variables at the higher level being modeled. Random effect estimates are variances that combine the between-subject and within-subject variance, and are used to draw conclusions about the population from which the subjects were drawn and are not limited to subjects in the study.

The parameter estimates are needed for the interpretation of results. Parameters of the mixed model can be estimated using maximum likelihood estimation (MLE) or restricted maximum likelihood estimation (RMLE). The Akaike Information Criteria (AIC) and the Bayesian Information Criteria (BIC) are used as measures of 'goodness of fit' for model selection, and smaller values are preferred for both the criteria. These are based on penalized likelihood method and provide a trade-off between 'goodness of fit' and complexity of the model. The likelihood increases with the complexity of the model, hence, a penalty is included to discourage over-fitting. When comparing models that differ on both fixed and random effects, both AIC and BIC based on MLE can be used. However, AIC and BIC with restricted likelihood can be used only for comparing models with random effects and those having same fixed effects.

For complex models, likelihood ratio tests can also be used to check the validity of the model fit. This is done by comparing the model selected with a null model. The test compares a null model that excludes the factors of interest to the model that includes them. It determines the significance of including additional terms into the statistical model of the data. It gives chi-square value, the associated degrees of freedom and the P value. A null model serves as a baseline model and needs to have the same random effects structure as the one being compared, when verifying with a random slope model.<sup>14, 16, 18</sup>

If the distribution of the observed data is not known, or if

the binomial or Poisson data has over dispersion or under dispersion, quasi likelihood estimation can be used.<sup>19</sup>

MLE approaches constitute a large family of commonly used methods for the analysis. Continuous responses are usually estimated via maximum likelihood, and binary (and other discrete) responses, by direct MLE via numerical quadrature, quasi-likelihood and Bayesian estimates. Multinomial responses could be estimated by adaptive Gauss-Hermite quadrature, penalized quasi likelihood (PQL), Monte Carlo expectation maximization (MCEM) algorithms and non-parametric maximum likelihood (NPMLE) methods.

With several extensions of the standard mixed model available, estimation and inference of the model may differ according to specific procedures. Measuring fit is further complicated when different estimation methods are used.<sup>16, 18, 20, 21</sup> Statistical software have assisted in the application of these estimation procedures and studies/literature comparing the statistical software available to perform multilevel modeling have made it easier to choose the appropriate software.<sup>8, 9</sup>

Missing values further complicate the analysis. In MLE, not only patients who attended all the visits, but also patients with missing visits contribute information. All the available information, including incomplete data, could be used, in contrast to ordinary regression analysis. In the context of likelihood inference, data with missing completely at random (MCAR) and missing at random (MAR) are ignorable. That means, we can ignore the missing data and obtain valid estimate. While missing not at random (MNAR) are non-ignorable. Finally, the sample size at various levels in MLM could influence precision of the variance terms and large sample sizes are required to conduct research with sufficient power. In multilevel analysis, the number of individual observations in groups is not as important as the number of groups in a study.<sup>5, 16, 18</sup>

Different approaches for modeling correlated data are available based on the inclusion of the predictors at the group level, multiple levels of groupings, and on the multivariate responses of the dependent variable (Table 2). When the response variable is continuous and normal, standard Gaussian model for continuous responses is useful. However, when the response variable is dichotomous or multinomial, the distribution is not normal and conclusions based on normal assumptions do not hold. The non-linear relation between the covariates and

**Table 2: Different approaches for analysis of correlated data structures**

Model type	Description
Fixed effects	Models with only fixed factors and fixed covariates (if present) as predictors. Variables are only representative of the study units. Intercepts are modeled as fixed effects. Fixed coefficient does not vary across individuals.
Random effects	Models with one or more random factors and random covariates (if present) as predictors. The variables could be random sample of all values and not just the complete set under study. Effects that influence covariance structure are also modeled as random factors. If only intercept is modeled is known as random intercept model. If a slope is also modeled is known as random co-efficient model.
Mixed effects	Mixed models with both fixed and random effects. Contributes both as an intercept and covariance structure to the model. Predictors at any level can be fixed factors. Grouping variables can be random factors.
Hierarchical or multilevel models	Mixed model with hierarchical structure where the variables are measured at more than one level. Level 1 data is nested in Level 2, where focus is on explaining the differences between groups in relation to differences within group. Can have more than 2 levels.
Random intercept model	Intercept of the level 1 dependent variable is modeled as an effect of level 2 grouping variable & level 1 or 2 or both covariates. It will only answer why the mean values of the dependent variable vary.
Random coefficient model	Along with the intercept, slope of the lower level covariate is also modeled as random effect of the grouping variable at level 2. Each group at higher level has an intercept and a slope to predict level 1 dependent variable. Random co-efficient vary across individuals, conditionally on the subject-specific random intercept and covariates. It explains the relative effects of predictor variables.
Multivariate Models	Enables simultaneous analysis of multiple dependent variables at level 1 in a multilevel setup in mixed model format
Cross-classified models and multiple membership models	Data that do not meet the nesting assumptions of hierarchical data can be analyzed. For example, if the subjects in the study are seen by doctor at different hospitals. Subjects are cross-classified by doctor effect.
Bayesian methods	Introduce prior probability distributions to the unknown parameters of the available data and draw inferences using Markov chain Monte Carlo (MCMC) methods by adding new data; this is taken as posterior distribution. Updating the prior beliefs with the new data. This gives exact inferences where likelihood based methods give approximations. The MCMC methods are particularly useful in models like the cross-classified and multiple membership models.



the response variable affects the covariance structure and adjusting the same is an important aspect of mixed modeling frameworks. Hence, identifying the proper distribution structure of the data and using an appropriate link function for the analysis is crucial.

The link function describes the non-linear transformation between the linear predictor and the assumed distribution structure. The most common distribution used for a binary outcome is the binary distribution, and in special cases Bernoulli distribution is preferred. For multinomial data, an extension of Bernoulli distribution is considered as multinomial distribution for outcome variable, ordinal data has an ordered categorical distribution and if the occurrence is very rare, the Poisson distribution is used. The most commonly used link functions are the log, logit, probit and complementary log-log link functions. The logit, log-log and probit models can be specified to model proportions and binary outcomes, logit for multinomial and ordered multinomial models to model multiple categories, and log link generally is used for Poisson and negative binomial distribution models to model counts. Complementary log-log link function is used for interval censored models with dependent variable having binomial distribution. McCulloch and Neuhaus (2012) have assessed the impact of misspecification on various aspects of data modeling.<sup>22</sup>

### **Bayesian approach**

An alternative to the classical approach discussed above is the Bayesian approach. Unlike classical approach that assumes parameters have fixed but unknown values, Bayesian method treats parameters as random variables and assigns probability distribution to unknown values. It requires formulation of prior distribution for any unknown parameter, a previous estimate, which is updated with probability distribution of the observed data. This updated distribution is called the posterior distribution.<sup>5</sup> It combines our prior beliefs with the new evidence, allowing the full use of all the information. Computation techniques for application of Bayesian are required, as it involves extensive calculations, and the most commonly used software is based on Markov Chain Monte Carlo (MCMC) method. MCMC gives a chain of correlated parameter estimates from the full posterior distributions and can be used to provide exact inferences, while classical statistics gives optimum point estimate of the parameter. Two main MCMC procedures used are Gibbs sampling and Metropolis-Hastings sampling.

### **Applications in biomedical studies**

The correlated modeling methodology has been used widely in genome-based studies to map copy number variations (CNV), particularly in over-dispersed and skewed coverage data to detect low-coverage and high-coverage regions in the genome. Sepulveda *et al.* (2013) have demonstrated the use of Poisson hierarchical models to detect CNV in seven *Plasmodium falciparum* malaria genomes.<sup>23</sup> They have selected Poisson-gamma and Poisson-log normal distributions for estimations using Bayesian approach. Hierarchical methodologies have been incorporated for the selection of more promising single nucleotide polymorphism (SNP) and allele-specific copy number variants estimation.<sup>24</sup> <sup>25</sup> Conventional techniques for the detection ignore the correlated data structure like association with phenotype and conservation across the species, which can be included in the model-based analysis. Association studies on autoimmunity and lymphoma has been assessed by multivariate hierarchical models.<sup>26</sup> Autoimmunity data presents a complex data structure with multiple factors at subject level and family level, along with multiple autoimmune conditions to be included to increase the precision of the estimates. Another area where this approach has been widely used is in studies on disease activity and pain, and coping abilities in autoimmune patients and personality traits.<sup>27</sup> Finan *et al.* (2010) have measured the influence of genetic variants on pain in fibromyalgia patients.<sup>28</sup> Further, a random effects model has been fit to study the increased risk of developing allergic disease in children with parental autoimmune condition.<sup>29</sup> Treatment efficacy studies have also been reported.<sup>30</sup> A Bayesian approach to calculate age-specific reference limits for serum cytokine levels in children has been demonstrated by Delezuch *et al.* (2012).<sup>31</sup>

Use of this methodology for research in clinical field has been restricted to very few studies, particularly in immunological and rheumatology. As most of the data structures used in these fields involve longitudinal data with correlated and clustered data sets, the applicability of mixed effect methodologies will be beneficial and essential. Patients with autoimmune diseases, such as systemic lupus erythematosus (SLE), polymyositis/ dermatomyositis (PM/DM), Sjogren's syndrome, rheumatoid arthritis (RA), etc. come up with multiple spells of clinical symptoms (outcome variable), with a prolonged disease/recovery process leading to correlation of observations and it is further complicated by multiple outcome variables. In addition, each disease has multiple covariates such as age, gender, autoantibodies, C-reactive protein and other biomarkers influencing the

process. The complexity of the data is increased with the involvement of multiple organs in some of the diseases. The disease can be clustered for different autoantibody groups (PM/DM), biomarkers (sero-positive or sero-negative RA), organ involvement (SLE) or disease activity (RA). Separate model strategies can be built to analyze the disease/recovery process for each autoimmune disease by accounting for clustering, confounding and correlation. This review tries to fill the gap in knowledge about the methodologies involved in statistical modeling of such data, with citations from available literature.<sup>10, 26-31</sup> It also highlights the significance of implementing such methodologies by researchers in immunologic and rheumatologic studies.

Overall, mixed modeling or multilevel modeling is used in both experimental and observational studies for data reduction, prediction and causal inference.<sup>32</sup> It plots an individual's dependent variable over time and explains the variability in the intercepts and slopes due to covariates at multiple levels to specify the relationship among variables. It can be used to assess whether the average rate of change is a function of grouping or due to time-invariant/time-variant explanatory variables and can explicitly model heterogeneity when differences between individuals cannot be explained by 'averages'. The effects could be generalized to wider population, if the subject selection process is random. The error term in the model indicates variance, which could not be accounted for by explanatory variables in the data structure. The model basically answers how each unit in the study change over time and predicts differences among units in their change. It models the data into two components: primarily, it analyzes factors that operate at the individual level, which is called 'compositional effect' like age, gender, habits, biomarkers etc. Secondly, it helps to gauge the 'contextual effects' with factors like physician, hospital set-up or the area/city in which the subject is undergoing treatment, socio-economic background, and literacy rate influencing the outcome. In addition, the model estimates the contribution of variables at intra-subject and inter-subject level. Different approaches are being reported in literature to suit the data structure requirements in various fields. Consequently the availability of standard software with updated versions is necessary for the analysis. A thorough knowledge of the modeling process and computational skills are essential; this is the only limitation for its application.

## Conclusion

In the disease/recovery process, often more than one type of event plays a role and a better understanding of these

processes help in disease management by developing new patient management strategies. In summary, use of mixed models allows dependency/clustering and heterogeneity at the within subject and between subject levels to be analyzed along with an unbalanced data structures commonly seen in biomedical studies. Statistical approaches for the analysis of the data are now available as standard software. The choice of a suitable model type should be done by considering the underlying data structure, as this will help in proper parameter estimates. In particular, it is important to clarify the distribution of dependent variable, choice of fixed and random factors, and the selection aspects of the model fit when performing the analysis, as it can be used in predicting the effect of variables and can influence the research question.

## Competing interests

The authors declare that they have no competing interests.

## Citation

Anupama K R, Chandrashekar S. Mixed effect frameworks in the analysis of longitudinal data IJRCI. 2014;2(1):R1.

**Received:** 5 March 2014, **Accepted:** 28 May 2014, **Published:** 10 June 2014

\*Correspondence: Anupama K R, ChanRe Rheumatology and Immunology Center, Basaweswaranagar, Bangalore, India. anupama\_kr789@yahoo.co.in

## References

1. Mills M. The fundamentals of survival and event history analysis, In: *Introducing Survival Analysis and Event History Analysis*. London: SAGE Publications; 2011. p. 1-17.
2. Wei LJ, Glidden DV. An overview of statistical methods for multiple failure time data in clinical trials. *Statist Med*. 1997;16(8):833-9.
3. Hamerle A. Multiple-splines Regression Models for Duration Data. *Appl Statist*. 1989;38(1):127-138.
4. Lim HJ, Zhang X. Additive and multiplicative hazards modeling for recurrent event data analysis. *BMC Medical Research Methodology*. 2011 Jun 27;11(1):101.
5. Goldstein H, Browne W, Rasbash J. Multilevel modelling of medical data. *Statist. Med*. 2002;21:3291-3315
6. Guo G, Zhao H. Multilevel modeling for binary data. *Annu. Rev. Sociol*. 2002;26:441-462.
7. Glidden D, Vittinghoff E. Modelling clustered survival data from multicentre clinical trials. *Statist. Med*. 2004;23(3):369-88.
8. Zhou X, Perkins AJ, Hui SL. Comparisons of software packages for generalized linear multilevel models. *The American Statistician* 1999;53:282-290.
9. Li B, Lingsma HF, Steyerberg EW, Lesaffre E. Logistic random effects regression models: a comparison of statistical packages for binary and ordinal outcomes. *BMC medical research methodology*. 2011;11:77.
10. Cole SR, Hudgens MG. Survival analysis in infectious disease research: describing events in time. *AIDS*. 2010;24(16):2423-31.
11. Hubbard A. Chapter 5 Longitudinal Data Models [Internet]. Berkeley: Environment health sciences, school of public health, University of



- California Online resources. c2006 [updated 2006 March 8]. Available from: <http://ehs.sph.berkeley.edu/hubbard/longdata/>
12. Meira-Machado L, de Una-Alvarez J, Cadarso-Suarez C, Andersen PK. Multi-state models for the analysis of time-to-event data. *Stat Methods Med Res.* 2009;18(2):195–222.
  13. Molenberghs G, Lesaffre E. Marginal modelling of Multivariate categorical data. *Statist. Med.* 1999;18:2237-2255.
  14. Rodelo A. An Overview of Mixed Effects Models [Internet]. California: San Francisco State University Online resources. c2005 [Updated on 2008 Jun 3]. Available from: <http://userwww.sfsu.edu/efc/classes/biol710/mixedeffects/Mixed-Effects-Models.pdf>
  15. Kwok OM, Underhill AT, Berry JW, Luo W, Elliott TR, Yoon M. Analyzing Longitudinal Data with Multilevel Models: An Example with Individuals Living with Lower Extremity Intra-articular Fractures. *Rehabil Psychol.* 2008;53(3):370-386.
  16. Hox J. *Multilevel Analysis Techniques and Applications*. 2nd ed. New York: Routledge publisher; 2010.
  17. Snijders, Tom AB. 'Fixed and Random Effects'. In: B.S. Everitt and D.C. Howell (eds.), *Encyclopedia of Statistics in Behavioral Science*. Volume 2, 664-665. Chichester (etc.): Wiley; 2005.
  18. Gibbons RD, Hedeker D. Applications of mixed-effects models in biostatistics. *Sankhya: the Indian journal of statistics.* 2000;62:70-103.
  19. Altman RM. Lecture 20: Quasi-Likelihood Estimation [Internet]. Simon Fraser University online resources. c2012. Available from: <http://people.stat.sfu.ca/~raltman/stat851/851L20.pdf>
  20. Stiratelli R, Laird N, Ware JH. Random-Effects Models for Serial Observations with Binary Response. *Biometrics.* 1984;40(4):961-971.
  21. Carriere I, Bouyer J. Random-effects models for ordinal responses: Application to self-reported disability among older persons. *Rev EpidemiolSantePublique.* 2006;54:61-72.
  22. McCulloch CE, Neuhaus JM. Misspecifying the Shape of a Random Effects Distribution: Why Getting It Wrong May Not Matter. *Statistical Science,* 2011;26(3):388-402.
  23. Sepúlveda N, Campino SG, Assefa SA, Sutherland CJ, Pain A, Clark TG. A Poisson hierarchical modelling approach to detecting copy number variation in sequence coverage data. *BMC Genomics.*2013;14:128.
  24. Chen GK, Witte JS. Enriching the Analysis of Genomewide Association Studies with Hierarchical Modeling. *Am. J. Hum. Genet.*2007;81:397-404.
  25. Scharpf RB, Ruczinski I, Carvalho B, Doan B, Chakravarti A, Irizarry RA. A multilevel model to address batch effects in copy number estimation using SNP arrays. *Biostatistics (Oxford, England).* 2011;12:33-50.
  26. Goldin LR, Landgren O. Autoimmunity and Lymphomagenesis. *Int J Cancer.* 2009;124(7):1497-1502.
  27. Zautra AJ, Fasman R, Parish BP, Davis MC. Daily fatigue in women with osteoarthritis, rheumatoid arthritis, and fibromyalgia. *Pain.* 2007;128:128-135.
  28. Finan PH, Zautra AJ, Davis MC, Lemery-Chalfant K, Covault J, Tennen H. Genetic Influences on the Dynamics of Pain and Affect in Fibromyalgia. *Health Psychology.* 2010;29(2)134-142.
  29. Maas T, Nieuwhof C, Passos VL, Robertson C, BoonenA, Landewé RB, et al. Transgenerational occurrence of allergic disease and autoimmunity: general practice-based epidemiological research. *Prim Care Respir J.* 2014;23.
  30. Barcellini W, Zaja F, Zaninoni A, Imperiali FG, Battista ML, Bona ED, et al. Low dose rituximab in adult patients with idiopathic autoimmune hemolytic anemia: clinical efficacy and biological studies. *Blood.* 2012;119(16)3691-3697.
  31. Delezuch W, Marttinen P, Kokki H, Heikkinen M, Vanamo K, Pulkki K, et al. Serum and CSF soluble CD26 and CD30 concentrations in healthy pediatric surgical outpatients. *Tissue Antigens.* 2012;80:368–375.
  32. Gelman A. Multilevel (Hierarchical) Modeling: What It Can and Cannot Do. *Technometrics.* 2006;48(3):432-435.