Comparison of Methods Dealing with Missing Data in a Longitudinal Rheumatologic Study

Boylamsal Bir Romatoloji Çalışmasında Kayıp Verilerin Üstesinden Gelen Yöntemlerin Karşılaştırılması

Hasan Fatih ÇAY, Mehmet Ziya FIRAT, Cahit KAÇAR

[1]Antalya Education and Research Hospital, Physical Medicine Rehabilitation and Rheumatology Department, Antalya, Turkey

[2]Akdeniz University, Faculty of Agriculture, Department of Animal Science, Biometry and Genetics Unit, Antalya, Turkey

[3]Akdeniz University, Faculty of Medicine, Physical Medicine Rehabilitation and Rheumatology Department, Antalya, Turkey

Correspondence Address
Yazışma Adresi

Mehmet Ziya FIRAT

Antalya Education and Research Hospital, Physical Medicine Rehabilitation and Rheumatology Department, Antalya, Turkey

E-mail: mziyafirat@gmail.com

Hasan Fatih ÇAY
ORCID ID: 0000-0001-5466-7482

Mehmet Ziya FIRAT
ORCID ID: 0000-0002-0091-4713

Cahit KAÇAR
ORCID ID: 0000-0002-1632-1621

ABSTRACT

Missing data are unavoidable in longitudinal studies and can lead to serious problems, such as loss of power and biased estimates, which should be solved in the statistical analysis of clinical studies. In this paper, three different techniques for handling missing data are shown using an example from a rheumatologic study. It is also shown how sensitive the conclusions of the study can be in terms of how the incomplete data are analyzed. The missing data process is studied in the framework of longitudinal data. The common approaches to handling missing longitudinal clinical trial data because of dropout are complete case (CC) and last observation carried forward (LOCF) analyses. These methods, while intuitively appealing, require tough assumptions to reach valid statistical conclusions. A relatively new and up to date statistical method for analyzing data with incomplete repeated measures is "likelihood-based ignorable method" which has less constraints and fewer tough assumptions than those required for CC and LOCF. We apply these three methods to data set of a rheumatologic trial comparing disease groups in terms of the joint pain scores using a mixed model. No significant differences were found between the methods of analysis. It can be concluded that attention to the mechanisms of missing data should be very important part of the analysis of rheumatologic data.

Keywords: Last observation carried forward, Complete case analysis, Longitudinal data, Rheumatology, Missing data process, Mixed models

ÖZ

Boylamsal çalışmalarda eksik veriler kaçınılmazdır ve klinik çalışmaların istatistiksel analizinde çözülmeleri gereken yanlı tahminler ve güç kaybı gibi ciddi sorunlara yol açabilirler. Bu makalede, romatolojik bir çalışmadan alınan örnek kullanılarak eksik verilerin üstesinden gelen üç farklı teknik gösterilmiştir. Ayrıca çalışmanın sonuçlarının, eksik verilerin analiz edilme şekli açısından ne kadar hassas olabileceği gösterilmiştir. Eksik veri süreci boylamsal veriler için incelenmiştir. Ayrılmalar nedeniyle eksik boylamsal klinik çalışma verilerinin üstesinden gelen yaygın yaklaşımlar, tam vaka (CC) ve son gözlemi ileri taşıma (LOCF) analizleridir. Sezgisel olarak çekici olmalarına rağmen bu yöntemler, geçerli istatistiksel sonuçlar üretmek için kısıtlayıcı varsayımlar gerektirirler. Eksik tekrarlanan ölçümleri içeren verileri analiz etmek için nispeten yeni ve modern bir istatistiksel yöntem, CC ve LOCF'ye göre daha az sınırlamaya ve kısıtlayıcı varsayımlara sahip olan "olabilirlik tabanlı" yöntemdir. Bu üç yöntemi, karışık bir model kullanarak eklem ağrı skorları açısından hastalık gruplarını karşılaştıran romatolojik bir çalışma setine uyguladık. Analiz yöntemleri arasında anlamlı bir fark bulunmamıştır. Eksik veri mekanizmalarına dikkatin romatolojik verilerin analizinde ayrılmaz bir parçası olması gerektiği sonucuna vardık.

Anahtar Sözcükler: Tam vaka analizi, Son gözlem ileriye taşıma, boylamsal veriler, Romatoloji, Eksik veri süreci, Karışık modeler

INTRODUCTION

Rheumatic diseases are of major concern for both individual patients and for community. Although there are so many sophisticated laboratory and technical tools for evaluation of the severity of disease, the value of patient's complaints –in other words the symptomatic evaluation- is still the cornerstone for clinical assessment. Intensity of joint pain is an important parameter to determine the severity of disease. Visual analogue scale (VAS) is considered one of the most frequently used methods to quantify the intensity of pain. It is considered as a sensitive method to alterations but difficult to apply especially in elderly and illiterate patients (1). Most of the clinical trials involve the follow up of patients over a period of times in order to describe the response and/or adverse effect of treatment. Rheumatologic assessments are performed regularly with periodic intervals which change from 3-4 weeks to several months. In such longitudinal rheumatologic studies, in actuality, it is not uncommon for some measurement sequences to end early due to circumstances outside the investigator's control (i.e. vacation, moving, death etc.), and a dropout or missing value is a unit that has been affected in this way. As a result, it may be required to account for dropout in the modeling process by considering the mean and covariance structure.

There are a lot of studies that are incomplete (2-3). In recent years, To deal with missing data sets, a number of strategies have been developed. However, the validity of the various approaches is dependent on the structure of the missing data and there are few studies dealing with missing data from rheumatologic studies. Complete case analysis (CC), in which only cases with complete data for all collected variables are analyzed, last observation carried forward analysis (LOCF), in which every missing value is replaced by the last observed value from the same subject, and likelihood-based ignorable analysis, developed under the missing at random assumption (MAR) which uses all available data without the need to delete or impute measurements or complete subjects, are three of these strategies

Among the above mentioned methods, the most prevalent methods for analyzing incomplete longitudinal clinical observations are complete case analysis (CC) and last observation carried forward (LOCF). These approaches have the benefit of being computationally simple and not requiring a full longitudinal model (4). Such methods, on the other hand, are based on strong assumptions, such as missing completely at random (MCAR) for CC and maintaining a constant profile after dropout for LOCF. The impact of these assumptions on the final results is sometimes overlooked when analyzing incomplete longitudinal data. In recent years, there are now a variety of full longitudinal data analysis approaches accessible, like as the likelihood-based ignorable analysis suggested by Molenberghs *et al*. (5) based

on the mixed-effects linear model for Gaussian outcomes. This approach uses all data, obviating the requirement for both removing and filling in data, and it requires MAR rather than the much stronger assumptions that underpin CC and LOCF.

Nevertheless, analysis of missing data has to be done carefully when subjects are discontinued for causes beyond the investigator's control. This implies the need for sensitivity analysis that was discussed by Fitzmaurice (6). Molenberghs *et al*. (5) also demonstrate that the missing sequences contribute to estimates of interest. Moreover, they demonstrate that sensitivity analysis is probable, without any further data modification.

The purpose of this study was to assess certain new statistical advancements in the context of rheumatological investigations where patients are measured repeatedly and dropout is a concern and to investigate the impact of three commonly used approaches, CC, LOCF and MAR, to dealing with missing data on the disease group and interaction effects of rheumatologic data analysis. Using empirical data, we compared these three mixed-model techniques without providing any mathematical specifics. To arrive at a trustworthy conclusion about disease group effects, the sensitivity analysis contrasts the final point estimates from the three statistical methodologies. Repeated measures analyses of variance using a mixed model were used to explore the effect of time on joint pain score among the patients participated in the study and to identify factors that were longitudinally associated with joint pain score. Furthermore, the sensitivity of results to model specification and alternative assumptions with different mean and covariance structures under three different methods were explored. This study will teach rheumatologist researchers how to handle missing data from clinical studies using statistical tools.

MATERIAL and METHODS

Patients

The template study, of which the data set used, is concerned about the relationship of rheumatologic symptoms and meteorological variables (7). It was used from December 2005 to July 2006. Enrollment took place between the 1st of December 2005 and the 28th of February 2006. Patients with rheumatoid arthritis (RA), spondyloarthropathy (SPA), and osteoarthritis (OA) of the knee were included in the study, as defined by the 1988 American College of Rheumatology (ACR) criteria for RA, European Spondyloarthropathy Study Group criteria, and ACR criteria for hand, knee, and hip OA, respectively (8–10). Patients were asked to complete questions in diaries that were handed to them. Each diary page –which were asked to be filled very day-has a question regarding rheumatic symptoms. The

evaluation used a 10-cm visual analogue scale (VAS). "How severe is your joint pain today?" is a question that needs to be answered. is known as the "joint pain score (JPS)" and is employed as a response variable in this research. The measurement on the answers of the question was employed by calculating the distance of marking applied by patient from point of zero. Although the measurements were taken from December 2005, the data used in this study included only the JPS taken between March 2006 and July 2006. This design was originally vulnerable to have plenty amount of missing data. Because we ask our patients to fill the VAS-based questionnaire of pain every day unless they are out of city or forgot to fill. No back-up recording was allowed. So there are many dropout observations due to natural design of investigation.

Table I presents the descriptive statistics of our rheumatologic data and sums up the measurements that can be repeated at each time point by disease groups individually. We can see that the frquency of missingness differs in different disease groups. It is clear from this table that the dropouts do not occur for SPA disease group. In general, all the fifty-four patients were observed at the first two time points, whereas only two, four and eleven patients were not seen at the third, fourth and fifth time points, respectively.

Statistical analyses were performed using SAS v8 software (11). The package allows users to model several covariance structures that can be modeled via MIXED procedure (12). The form of these covariance structures and more details are given in SAS/STAT 8.2 User's Guide.

Method

To assess the association between age and each of the JPS metrics, month and disease groups (RA, SPA and OA), mixed effect models for repeated measurements (RMMEM) (13-14) were fitted with month (linear, quadratic and cubic), disease groups and the interaction of linear, quadratic and cubic month effect with disease groups as the fixed effects, and month only (marginal models) as the random effect. We fitted eight different models that belong to the family

of the general linear mixed models. Let $\mathbf{Y}_i$ indicate the vector of JPS measures for the i th patient. RMMEM can be written in a generalized manner as

$$Y_i = X_i\alpha + Z_i\gamma_i + \varepsilon_i \quad \text{for } i = 1, \ldots, N \quad (1)$$

where N denotes the total number of patients; a $(\mathbf{n}_i \times 1)$ vector of dependent variable for the i th patient is denoted by $\mathbf{Y}_i$; the number of measurements for the i th patient is shown as $\mathbf{n}_i$; $\mathbf{X}_i$ denotes a $(\mathbf{n}_i \times p)$ design matrix for the fixed effects; a $(p \times 1)$ vector of fixed regression coefficients is denoted by α; $\mathbf{Z}_i$ idenotes a $(\mathbf{n}_i \times q)$ matrix of covariates related with the random effect; γ_i $(q \times 1)$ denotes a vector of random effects parameters and γ_i is normally distributed as N($\mathbf{0}$, $\mathbf{D}$); and ε_i $(n_i \times 1)$ indicates a vector of error terms and ε_i is normally distributed as N($\mathbf{0}$,Σ) and γ_i and ε_i are independent of each other. Both measurement error and serial correlation are included in the random error, ε_i. In this research, the design will be based on age (as a covariate), disease groups, month and/or interaction between disease groups and month. The marginal distribution of the dependent variable $\mathbf{Y}_i$ is used to make inferences. After integrating over random effects, $\mathbf{Y}_i$ can be re-written as follows

$$Y_i \sim N(X_i\alpha, Z_iDZ_i' + \textstyle\sum_i).$$

Here $\sum_i = \sigma^2 I_{n_i} + \tau^2 H_i$ is a variance-covariance matrix that groups the measurement error and serial components and $V_i = Z_i DZ_i' + \sum_i$ is defined as the general variance-covariance matrix of $\mathbf{Y}_i$

Variance-Covariance Structures

In this section, a parsimonious description of eight models used in this study is given. The models are defined according to their mean structure and the variance-covariance structure. Model 1 implies that each disease group x month combinations has its own mean, together with an unstructured covariance, yielding additional parameters. It is assumed that the error term's ε_i variance-covariance matrix $\sum$ is a positive definite matrix and no random effects are included. Model 2 implies that each illness group has a linear trend,

Table I: The descriptive statistics of our rheumatologic data and summary of the number of available repeated measurements at each time point by disease groups separately.

		Disease Group								All			
		RA			SPA			OA			All		N
	N	Mean	Stdev	N	Mean	Stdev	N	Mean	Stdev	N	Mean	Stdev	Missing
March	31	2.83	2.32	13	2.83	2.84	10	3.72	2.51	54	3.00	2.46	0
April	31	2.53	1.94	13	3.20	2.80	10	3.97	2.25	54	2.96	2.26	0
May	30	2.54	2.06	13	3.26	2.76	9	2.40	1.67	52	2.69	2.18	2
June	28	2.77	2.27	13	3.13	2.85	9	2.18	2.29	50	2.76	2.40	4
July	24	2.58	1.98	13	2.97	2.91	6	1.79	1.32	43	2.59	2.22	11

while Model 3 assumes that the mean profiles are parallel straight lines. Toeplitz variance-covariance structure and first-order autoregressive (AR(1)) covariance matrix are assumed in Models 4 and 5. Model 6, which is a simplified version of the unstructured covariance Model 2, allows for random intercept and slope parameters. Model 7 is a hierarchical random intercepts model implying a compound symmetry model at the marginal level. Finally, Model 8 is an independence model in which measurement error is the only source of variability.

Variance–covariance structure selection and evaluation

There are various candidate variance–covariance structures to choose from, as previously indicated. A lot of analytical criteria must be examined while making a proper selection. These are known as Information Criteria, and they are based on likelihood estimations (15-16). The Akaike information criterion (AIC) was used to determine the best covariance structure of the models for the data (17) and the Schwarz Bayesian Information Criterion (BIC) (18). Let l be the maximum value of the model's log likelihood and d be the number of the parameters of the variance–covariance structure, then when comparing the variance–covariance structures of two identical expectation structures, the AIC can be stated as

$$AIC = -2l + 2d \qquad (2)$$

BIC is almost the same, instead of doublingthe number of covariance parameters d, the penalty is obtained as $\log(N)$

$$BIC = -2l + \log(N)d \qquad (3)$$

With the rule that smaller is better, the corresponding values of these two criteria are compared across various covariance structures (i.e., the structure with the least criteria value is the best covariance structure for the data, and the fixed-effect tests associated with this structure should be interpreted accordingly). Because the BIC has a higher penalty, which is a function of the number of unknown parameters and sample size, the two criteria may not always agree on the optimum structure. We will use the BIC criterion rather than the AIC criterion because our goal is to model the covariance structure as efficiently as possible.

The likelihood ratio chi-square test can be used to see if there is a statistical difference between alternative models. Because a model's variance–covariance structure (model B, for example) is a reduced (nested) structure of mixed models (model A or nested model, for example), the likelihood-ratio test (LRT) can be used to determine whether the chosen variance–covariance structure is significantly more appropriate than model B. The LRT statistic is derived from the following equation:

$$LRT = -2(\log l_B - \log l_A) \sim \chi^2(d.f.),$$

where loglB and loglA signify the log of the model B's restricted maximum likelihood and the mixed model A's variance–covariance structures, respectively. The LRT statistic follows a χ^2 distribution with the degrees of freedom (d.f.) equal to the difference in the number of parameters of both models. Under the null hypothesis that the mixed model with selected variance–covariance structures is not different from the model B. The d.f. are in practice equal to the difference in the variance–covariance parameters of both models because their expectation structures are all the same. When compared to model B, the model fit of the mixed model with the chosen variance–covariance structure is regarded superior if $LRT > \chi^2$, with a significance level of α.

Methods for Dealing With Missing Data

In the literature, there are a variety of approaches for dealing with missing data that were developed to deal with missing data in longitudinal clinical trials. The goal of this research is to exemine three of the most commonly used approaches for dealing with missing data, which are outlined below.

Complete Case (CC) Analysis: After removing all cases with missing data, this approach does statistical analyses on the smaller data set. There is no missing data problem to deal with now that all patients with missing data have been eliminated. The most important advantage of this method that it is easy to use. However, the method may be preferable for the large sample sizes, and MCAR is the mechanism for missing data (19). On the other hand, the biggest disadvantage of the method is that there could possibly the loss of statistical power because of the reduction of the sample size (20).

Last Observation Carried Forward (LOCF): This approach has frequently been employed in dealing with missing data problems and every missing value is replaced with the most recent observed value. It is generally used in continuous longitudinal data under MCAR. There is no temporal effect since the last observed data because the LOCF approach believes that the outcome will not change after the last observed value. It has been demonstrated that LOCF can produce biased findings, resulting in overestimation or underestimating of parameter estimates (5), (21-24).

Missing at Random (MAR) : A value of a clinical outcome variable is known to be missing at random (MAR) if, conditional on the observed data, the missingness is independent of the unobserved outcomes (25). Because it can be predicted from the observed data in the model, the missingness mechanism owing to prior lack of efficacy can be MAR. Therefore, the analysis model is closely linked to

MAR. We will be operating under MAR if we include all of the factors that affect missingness in our model; otherwise, the assumptions of the MAR would not apply to our analysis.

RESULTS

Complete case analysis includes analyses of the cases for which all n_i measurements are complete. In the complete case analysis of our data, the eleven patients that lack measurements are removed, as a result, a functional data set of 43 patients was created. The model fit summary of the eight models described previously is given in Table II. The deviance (minus twice the log-likelihood at maximum) of Model 1 equals 577.3, and there are 30 model parameters. This deviation will be used as a benchmark to determine the goodness-of-fit of much simpler models. All the deviances are listed in Table II. As can be noted from this table, Models 6, 7 and 8 are strongly rejected. The common slope Model 3 is also rejected at the %5 significance level. The banded correlation structure Model 4 and the first-order autoregressive Model 5 are not rejected when compared to Model 2. When the Model 5's likelihood is compared to the likelihood of the reference Model 4, it becomes clear that Model 5 is consistent with the data. From this analysis it can be concluded that Among the eight models investigated in this study, Model 5 provides the most concise description that is compatible with the data.

In the last observation carried forward analysis of our data, a new data set is obtained by completing in the missing values, instead of removing patients with missing observations. The principle of imputation is that whenever a value is missing, the last observation on the same patient is substituted. A summary of the model fit of the eight models is presented in Table III. Again, there are several important

features of the results of LOCF analysis. It is clear from Table III that Models 4, 6, 7 and 8 are strongly rejected. Models 2 and 3 are also rejected at the nominal %5 and %1 levels, respectively. Model 5 is highly rejected when compared to Model 2, but it is not rejected when compared to Model 4, as shown in the CC analysis. In summary, to summarize the results, Model 5 is recommended among the models offered.

A summary of the model fit of the eight models is presented for likelihood-based ignorable analysis in Table IV. The findings of this investigation are quite similar with the results of LOCF analysis. It is clear from Table IV that there is clear evidence of lack of fit in the unstructured models (Models 2 and 3), Toeplitz (banded), AR(1) in comparison to Model 2, compound symmetry and simple models. Once again, the first order autoregressive model when compared with Model 4 does exhibit the best fit among all the eight models considered.

The indices of relative goodness-of-fit Akaike's information criteria and Schwarz's Bayesian criterion can be used to evaluate models with the same fixed effects but different covariance structures. Both of these criteria are applicable to model selection and hypothesis testing in general. AIC and BIC values for the eight covariance structures are shown for the CC analysis in Table II. Since AR(1) has the smallest AIC and BIC values, we can conclude that the first order autoregressive model is the best choice of covariance structure for the CC analysis. Table III shows AIC and BIC values for the eight models from the LOCF analysis. 'Unstructured' (Model 1) has the smallest AIC, but AR(1), 'first order autoregressive model', has the smallest BIC. The penalty for the large number of parameters in the UN covariance matrix is reflected in the difference between AIC and BIC for the UN structure. As our objec-

Table II: Model fit summary for complete case analysis.

Model Mean	Covariance	Par.	-2l	Ref.	G²	d.f.	p-value	AIC	BIC
M1 unstr.	unstructured	30	577.3					639.3	693.9
M2 ≠ slopes	unstructured	28	582.6	1	5.3	2	0.0706	638.6	687.9
M3 = slopes	unstructured	26	590.2	2	7.6	2	0.0223	642.2	688.0
M4 ≠ slopes	Toeplitz	18	600.6	2	18.0	10	0.0549	636.6	668.3
M5 ≠ slopes	AR(1)	15	603.8	2	21.2	13	0.0690	633.8	660.2
				4	3.2	3	0.3618		
M6 ≠ slopes	random	17	618.3	2	35.7	11	0.0001	652.3	682.2
M7 ≠ slopes	CS	15	671.0	2	88.4	13	0.0000	701.0	727.4
				4	70.4	3	0.0000		
				6	52.7	2	0.0000		
M8 ≠ slopes	simple	14	949.8	7	278.8	1	0.0000	977.8	1002.5

Par: number of model parameters; **−2l:** minus twice log-likelihood; **Ref:** reference model for likelihood ratio test; **G²:** likelihood ratio test statistic value; **df:** corresponding number of degrees-of-freedom).

Table III: Model fit summary for last observation carried forward analysis.

Mean	Covariance	Par.	-2l	Ref.	G²	d.f.	p-value	AIC	BIC
M1 unstr.	unstructured	30	757.4					819.4	881.0
M2 ≠ slopes	unstructured	28	768.4	1	11.0	2	0.0040	824.4	880.1
M3 = slopes	unstructured	26	775.9	2	7.5	2	0.0235	827.9	879.6
M4 ≠ slopes	Toeplitz	18	815.6	2	47.2	10	0.0000	851.6	887.4
M5 ≠ slopes	AR(1)	15	817.5	2	49.1	13	0.0000	847.5	877.4
				4	1.9	3	0.5934		
M6 ≠ slopes	random	17	835.9	2	67.5	11	0.0000	869.9	903.7
M7 ≠ slopes	CS	15	897.7	2	129.3	13	0.0000	927.7	957.6
				4	82.1	3	0.0000		
				6	61.8	2	0.0000		
M8 ≠ slopes	simple	14	1207.5	7	309.8	1	0.0000	1235.5	1263.4

Par: number of model parameters; **−2l:** minus twice log-likelihood; **Ref:** reference model for likelihood ratio test; **G²:** likelihood ratio test statistic value; **df:** corresponding number of degrees-of-freedom.

Table IV: Model fit summary for likelihood-based ignorable analysis.

Mean	Covariance	Par.	-2l	Ref.	G²	d.f.	p-value	AIC	BIC
M1 unstr.	unstructured	30	726.4					788.4	850.1
M2 ≠ slopes	unstructured	28	739.5	1	13.1	2	0.0014	795.5	851.2
M3 = slopes	unstructured	26	748.6	2	9.1	2	0.0105	800.6	852.3
M4 ≠ slopes	Toeplitz	18	782.3	2	42.8	10	0.0000	818.3	854.1
M5 ≠ slopes	AR(1)	15	785.4	2	45.9	13	0.0000	815.4	845.2
				4	3.1	3	0.3764		
M6 ≠ slopes	random	17	793.5	2	54.0	11	0.0000	827.5	861.3
M7 ≠ slopes	CS	15	845.4	2	105.9	13	0.0000	875.4	905.3
				4	63.1	3	0.0000		
				6	51.9	2	0.0000		
M8 ≠ slopes	simple	14	1123.1	7	277.7	1	0.0000	1151.1	1178.9

Par: number of model parameters; **−2l:** minus twice log-likelihood; **Ref:** reference model for likelihood ratio test; **G²:** likelihood ratio test statistic value; **df:** corresponding number of degrees-of-freedom.

tive is parsimonious modeling of the covariance structure, we conclude that AR(1) provides the most parsimonious description among the eight models for the LOCF analysis. AIC and BIC values of the eight models are illustrated for the likelihood-based ignorable analysis in Table IV. These results are similar with the results of LOCF analysis and once again AR(1) model exhibit the best fit compared to other models.

We have so far fitted eight covariance structures to our data, and compared the covariance structures of three different methods of analysis using -2l, AIC and BIC values. Model comparisons performed in Tables I, II, and III quantitatively produce the same inferences. In all cases, we determined that the first order autoregressive model (AR(1), or Model 5) provided the best fit and was the most concise explanation of the data among the eight models

considered for all three methods of analysis, CC, LOCF and MAR. Let us now turn attention to the performance of the three different methods of analysis and compare the results with each other. We will first study the effects of the selected variance-covariance structure, AR(1), on tests and fixed effects's estimates for different methods of analysis. In this type of longitudinal analysis, only the disease groups that are of main interest will be considered. This means we estimate the main effects of disease groups as well as the interaction effect of disease groups by month. We treated month as a continuous variable and modeled month effects in third degree polynomials. Then we used the polynomial model to obtain the parameter estimates, and standard errors for different methods of analysis. The general linear mixed model was fitted using the selected variance-covariance structure to model fixed effects of disease groups and month, and the parameter estimates, standard errors

and *p* values from three different methods of dealing with missing data are shown for age, disease group main effects, and disease group by month interaction in Table V. Disease group SPA is considered as the reference group in this table. In either the illness group main effects or the disease group by month interaction, there is not much difference between the estimates of the three approaches when comparing CC, LOCF, and MAR. In terms of other model characteristics, however, there will be some significant discrepancies between the solutions for addressing missing data.

Table V contains the fitted correlations. Clearly, the correlations produced by LOCF and MAR are nearly identical. On the other hand, the correlation estimated using CC is slightly stronger than the other two methods.

Table VI contains values of *F* tests for fixed effects for the selected covariance structure, AR(1), using the three different methods of analysis. A significant disease group effect and the interaction between disease group and month were found using all three methods. The largest *F* values and the smallest *p* values are seen with the MAR analysis. Note also that the *F* values are substantially smaller for CC from those of LOCF and MAR.

The results of fitting the selected Model 5 to our data using three different methods of analysis, CC, LOCF, and MAR, are compared in Table VII. The effect of age, disease groups, month, and the interaction between disease groups and month were all included in the model. It is worth noting that the results are identical to those in Table V, where the inference is based on the same model definition. Note also that the χ^2 results demonstrate a different picture in terms of evidence for methods of analysis, with extremely strong evidence for CC when comparing LOCF with CC and MAR with CC and very strong evidence for MAR when comparing LOCF with MAR. However, when

Table V: Estimated parameters, standard errors and p-values for fixed effects of the selected Model 5 and correlations using different methods of analysis (CC, LOCF and MAR).

Effect	DG*	CC Estimate	CC S.E.	CC p	LOCF Estimate	LOCF S.E.	LOCF p	MAR Estimate	MAR S.E.	MAR p
Intercept		-6.757	7.750	0.388	-6.459	8.822	0.467	-6.534	9.162	0.479
Age		0.043	0.032	0.181	0.036	0.029	0.222	0.038	0.029	0.197
DG	OA	-17.610	13.591	0.203	-14.978	13.256	0.264	-18.165	14.259	0.209
DG	RA	15.050	9.478	0.120	18.511	10.406	0.081	19.485	10.923	0.081
DG	SPA									
Month*DG	OA	14.182	5.968	0.018	12.715	5.293	0.017	14.485	5.893	0.015
Month*DG	RA	-4.334	2.984	0.148	-5.790	3.006	0.056	-6.353	3.243	0.052
Month*DG	SPA	3.670	4.054	0.367	3.670	4.643	0.430	3.670	4.826	0.448
Month²*DG	OA	-2.531	1.023	0.014	-2.248	0.908	0.014	-2.565	1.022	0.013
Month²*DG	RA	0.792	0.512	0.124	0.970	0.515	0.061	1.074	0.559	0.056
Month²*DG	SPA	-0.533	0.695	0.444	-0.533	0.796	0.504	-0.533	0.828	0.520
Month³*DG	OA	0.141	0.057	0.014	0.125	0.050	0.014	0.142	0.057	0.014
Month³*DG	RA	-0.047	0.028	0.101	-0.054	0.029	0.062	-0.059	0.031	0.057
Month³*DG	SPA	0.025	0.038	0.523	0.025	0.044	0.576	0.025	0.046	0.591
r		0.931	0.015	0.000	0.917	0.016	0.000	0.909	0.017	0.000

***DG:** Disease groups.

Table VI: Type III test of fixed effects for the selected model 5 using different methods of analysis (CC, LOCF and MAR).

Effect	CC F value	CC p value	LOCF F value	LOCF p value	MAR F value	MAR p value
Age	1.86	0.1806	1.53	0.2224	1.71	0.1968
DG	3.82	0.0306	4.81	0.0123	5.02	0.0104
Month*DG	2.86	0.0387	3.37	0.0195	3.49	0.0169
Month²*DG	3.03	0.0309	3.38	0.0193	3.47	0.0173
Month³*DG	3.10	0.0284	3.33	0.0205	3.38	0.0195

Table VII: Comparison of fitted CC, LOCF and MAR models to our data.

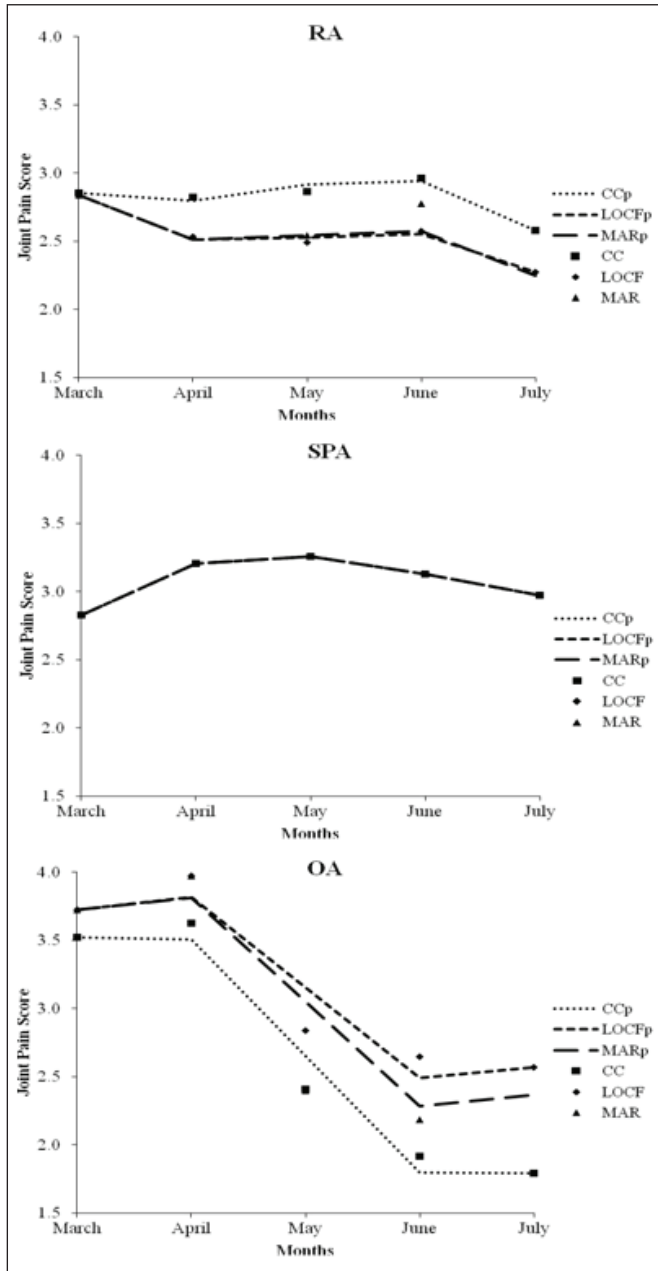| Methods	-2l	Comparison		χ^2	p value
CC	603.8	LOCF - CC	=	213.7	<0.0001
LOCF	817.5	LOCF − MAR	=	32.1	<0.0001
MAR	785.4	MAR - CC	=	181.6	<0.0001

Figure 1: Polynomial curves of the three different methods of analysis for each disease group, RA, SPA, and OA. (The large dots are the observed means for different methods (CC, LOCF, MAR) and the dotted lines are the corresponding quantities for the fitted values (CCp, LOCFp, MARp)).

assessing evidence for or against CC, utmost caution should be exercised (26-27).

The polynomial curves resulting from the three different methods of analysis, CC, LOCF and MAR, are plotted for disease groups, RA, SPA, and OA, in Figure I. In this figure, the observed (CC, LOCF and MAR) and fitted (CCp, LOCFp and MARp) mean structures of the selected Model 5 are overlaid on the same graph. The graph illustrates that means for the two disease groups, RA and SPA are basically the same in March. However, the mean for disease group OA is larger than the mean value for RA and SPA. Means for disease group OA decrease sharply between months April and June whereas the means for disease group RA and SPA seem to remain almost the same after the month of April. However, the magnitudes of the differences between disease groups decrease dramatically with time. As can be seen from Figure I, the observed and fitted means coincide for the disease groups RA and SPA, but they do not coincide for the disease group OA. While there is absolutely no difference between the three methods in SPA, there is a clear distinction between LOCF and CC, as well as MAR and CC in RA and OA. In RA, there is clearly no difference between LOCF and MAR over time, but the difference between these methods with CC is consistent over time. In OA, the difference between LOCF and MAR is relatively mild and a strong separation is observed between these methods and CC.

DISCUSSION

In the current study considered here, we have shown that when evaluating incomplete longitudinal data from rheumatologic investigations, a range of techniques such as CC, LOCF, and MAR can be used. The main mode of analysis is the linear mixed model. These strategies are now within reach thanks to advances in statistical computer power, and we've used them in a real-world rheumatologic study.

We looked at the impact of several techniques for dealing with missing data. The results of the three methods showed some slight differences with respect to the parameter estimates, standard errors, p values and F values of the main effects and interaction terms. However the differences were not extreme between the various analyses conducted and caution should be used when deciding on the method of analysis.

Finally, there is no "gold" standard available in the context of this empirical study. Simulation studies are needed to better understand the relationship between the amount and nature of missing values, as well as to compare the results to a gold standard. In conclusion, only minor variations were discovered between the three strategies of managing missing data regarding the illness group in our empirical example from a rheumatologic study.

ACKNOWLEDGEMENTS

REFERENCES

1. van Riel PCLM, van Gestel AM, Welsing PMC. Evaluation and outcome of the patient with established rheumatoid arthritis. In: Hochberg MC et al (eds) Rheumatology, 3rd edn. Mosby Elsevier: Philadelphia, 2003; 893-905.

2. Little RJA, Rubin DB. Statistical Analysis with Missing Data, 2nd ed. Wiley: New York, 2002.

3. Schafer JL, Graham JW. Missing data: Our view of the state of the art. Psychological Methods 2002; 7 (2): 147-177.

4. Jansen I, Beunckens C, Molenberghs G, Verbeke G, Mallinckrodt C. Analyzing incomplete discrete longitudinal clinical trial data. Statistical Science 2006; 21: 52–69.

5. Molenberghs G, Thijs H, Jansen I, Beunckens C, Kenward M G, Mallinckrodt C. Carroll R J. Analyzing incomplete longitudinal clinical trial data. Biostatistics 2004; 5: 445–464.

6. Fitzmaurice GM. Methods for handling dropouts in longitudinal clinical trials. Statistica Neerlandica 2003; 57: 75–99

7. Cay HF, Sezer I, Firat MZ, Kaçar C. Which is the dominant factor for perception of rheumatic pain: meteorology or psychology? Rheumatol Int 2011;31(3):377-85.

8. Harris ED. Clinical features of rheumatoid arthritis. In: Harris ED et al (eds) Kelley's textbook of rheumatology, 7th edn. Elsevier Saunders: Amsterdam, 2005; 1043–1078.

9. Olivieri I, van Tubergen A, Salvarani C, van der Linden S. Seronegative spondyloarthritides. Best Pract Res Clin Rheumatol 2002; 16(5):723–739

10. Dennison E, Cooper C. Osteoarthritis: epidemiology and classification. In: Hochberg MC et al (eds) Rheumatology, 3rd edn. Mosby Elsevier: Philadelphia, 2003; 1781–1791.

11. SAS Institute. SAS/STAT Software, Version 8. Author, Cary, NC, 2008.

12. Kincaid C. Guidelines for selecting the covariance structure in mixed model analysis. SUGI 30 Proceedings: Paper 198-3, 2005.

13. Larid NM, Ware JH. Random-effect models for longitudinal data. Biometrics 1982; 38: 963–974.

14. Diggle PJ, Liang KY, Zeger SL. The Analysis of Longitudinal Data. Oxford University Press: Oxford, 1994.

15. Oman SD. Multiplicative effects in mixed model analysis of variance. Biometrika 1991; 78: 729–739.

16. Wolfinger R. Covariance structure selection in general mixed models. Commun.Stat-Simul. C 1993; 22: 1079–1106.

17. Akaike H. A new look at the statistical model identification. IEEE Transaction on Automatic Control 1974; AC-19: 716-723.

18. Schwarz G. Estimating the dimension of a model. Annals of Statistics 1978; 6: 461-464.

19. Kim JO, Curry J. The Treatment of Missing Data in Multivariate Analysis. Sociological Methods Research,1977; 6: 215-240.

20. Allison, PD. Multiple Regression: A Primer. Pine Forge Press, Thousand Oaks, 1999.

21. Shao J, Zhong B. Last Observation Carry-Forward and Last Observation Analysis. Statistics in Medicine 2003; 22: 2429-2441.

22. Mallinckrodt CH, Clark WS, David SR. Accounting for Dropout Bias Using Mixed-Effects Models. Journal of Biopharmaceutical Statistics 2001; 11: 9-21.

23. Mallinckrodt CH, Kaiser CJ, Watkin JG, Detke MJ, Molenberghs G, Carroll RJ. Type I Error Rates from Likelihood-Based Repeated Measures Analyses of Incomplete Longitudinal Data. Pharmaceutical Statistics 2004; 3: 171-186.

24. Gadbury GL, Coffey CS, Allison DB. Modern Statistical Methods for Handling Missing Repeated Measurements in Obesity Trials: Beyond LOCF. Obesity Reviews 2003; 4: 175-184.

25. Little RJA, Rubin DB. Statistical Analysis with Missing Data. New York: Wiley, 1987.

26. Diggle PJ, Kenward MG. Informative drop-out in longitudinal data analysis (with discussion). Applied Statistics 1994; 43: 49–93.

27. Verbeke G, Molenberghs G. Linear Mixed Models for Longitudinal Data. New York: Springer, 2000.