Cancer Patients Missing Pain Score Information: Application with Imputation Techniques

Gajendra K Vishwakarma [1], Atanu Bhattacharjee [2], Jesna Jose [1], Ramesh Kumar V [3]

(1) Department of Applied Mathematics, Indian Institute of Technology (ISM) Dhanbad - 826004, Jharkhand, India
(2) Division of Clinical Research and Biostatistics, Malabar Cancer Centre, Thalassery - 670103, Kerala, India
(3) Department of Biostatistics, ICON clinical Research pvt. Ltd. Bangalore, India

CORRESPONDING AUTHOR: *Atanu Bhattacharjee, Division of Clinical Research and Biostatistics, Malabar Cancer Centre, Thalassery-670103, Kerala, India; email: atanustat@gmail.com*

DOI: ***10.2427/11916***
Accepted on September 27, 2016

ABSTRACT

Background: Methods for handling missing data in clinical research are getting more attention since last few years. Contemplation of missing data in any study is crucial as they may lead to considerable biases. It can be handled by simply excluding any patients with missing values from the analysis; this will result in a diminution in the number of cases available for analysis and hence have the impact on statistical power. Hence every attempt should be made to minimise the amount of missing data. The missing data handling technique such as single imputation methods are attractive, but do not reflect the uncertainty about the predictions of the unknown missing values, and hence estimated variance of the parameter will be biased toward zero. The palliative care treatment is a specialised medical care for people with serious illness and it focuses on providing relief from symptoms, and can be used at any stage of an illness if there are troubling symptoms, such as pain or sickness.
Objective: This manuscript presents different imputation techniques to handle missing observations obtained from a repeatedly measured pain score data on palliative cancer.
Methods: Imputation methods such as Regression, Predictive Mean matching, Propensity Score, EM algorithm and Markov Chain Monte Carlo (MCMC) methods were adopted and compared to find out the appropriate imputation method on pain score data. The appropriate imputation method is decided based on the lowest standard error (SE) calculated during the Regression analysis.
Results: The mean (SD) of observed data was 3.638 (3.175) whereas the imputed mean (SD) values were 3.356 (2.6603), 3.502 (2.6100), 3.406 (2.4334), 3.474 (2.6285) and 3.264 (2.6336) respectively, for the methods with Regression, Predictive Mean matching, Propensity Score, EM algorithm and MCMC methods for pain score values at visit three. The mean (SD) of observed data was 3.528 (3.1112) whereas the imputed mean (SD) were 3.231 (2.8715), 3.253 (2.8691), 3.278 (2.7935), 3.268 (2.8725) and 3.227 (2.8952) respectively, for the methods Regression, Predictive Mean Matching, Propensity Score, EM algorithm and MCMC methods for pain score values at visit two.
Conclusion: Accordingly, to our methodology, the Propensity Score Method has appeared to be the most appropriate imputation method for pain score data. The multiple imputation techniques have few advantages; the imputed values are drawn from a distribution, so they inherently contain some variation by introducing an additional form of error in the parameter estimates across the imputation.

Key words: EM algorithm, Regression method, Imputation, Handling Missing Data.

INTRODUCTION

Palliative treatment is designed to relieve symptoms and stress of any serious illness and can be used at any stage of an illness if there are troubling symptoms, such as pain. It may help someone to live longer and to live comfortably, even if they cannot be cured [1]. The pain scores measured at repeated time point of non-cured patient is the main concern in any palliative care. One of the most frequently encountered problems while conducting trials on palliative care patients is that they will drop out (or withdraw) before study completion. There are many probable reasons for missing data (e.g., patient refusal to continue the study, patient withdrawals due to treatment failure, treatment success or adverse events, patients moving), only some of which are related to study treatment. Dropouts can produce biased treatment comparisons and also reduce the overall statistical power. Interpretation of the results of a study is always challenging when the proportion of missing values is extensive.

There are various degrees of data incompleteness that might occur in the study, i.e. measurements may be available only at baseline, measurements may be missing at baseline, or may be missing for one, several or all follow-up assessments. Even if a patient completes the study, some data may remain simply unreported or uncollected. This paper will focus on the observation where missing data occur as a result of patients dropping out of the study.

There is a substantial amount of literature available about handling missing data technique for the repeated measurement [2-5] and types of missing data [6-8]. However, all of these works are dedicated and developed through consideration of the therapeutic arms effect comparison. In this study, we tried to bring statistical methodologies as extensions to impute the missing values of repeated measured pain scores, without considering therapeutic arms for modelling and data imputation. The supporting covariates to impute the repeated measured pain score were considered as time independent variable measured at baseline measurement. This work is particularly dedicated to provide the statistical inference to overcome the presence of missing observations obtained from palliative treated patients.

DATA METHODOLOGY

The different multiple imputation techniques were considered from pain score for palliative treated patients. A total of 326 palliative care subjects suffering from cancer attending outpatient department were included in this analysis. The data considered in this analysis are pain score observed at visit one, two and three respectively. The missing observations were observed at visit two and three for few subjects. The data set has a monotone missing data pattern and it is assumed that the missing data are missing at random (MAR), that is, the probability that an observation is missing may depend on Y_{obs} but not on Y_{miss} [9,10]. If any palliative care subject has the missing observation, the reason for this missing observation might only depend on their observed pain score values and not on unobserved pain score values. The purpose of this analysis was to compare the estimates of different imputation techniques. The PROC MI statement is the only required statement in the MI procedure for imputation. Available options in the PROC MI statement are considered: METHOD=REGRESSION, METHOD=PROPENSITY, METHOD=MCMC and impute=10. The chain=multiple had been used with method=MCMC for each imputation. The option INITIAL=EM was adapted to EM algorithm. The means and standard deviations from available cases were the initial estimates for the EM algorithm using proc MI. The correlations are set to zero. The resulting estimates are used to begin the MCMC process. Pain scores at previous visits and baseline characteristics such as age, sex, diagnosis and stage were used to impute the missing pain score. Regression analysis was carried out to examine the relationship between change in pain score at visit three (visit three - visit one) and baseline characteristics for the available case (complete case analysis) patients. Similarly, regression analysis was carried out for the data sets which were imputed using imputation methods; Regression, Predictive Means Matching, Propensity Score, EM algorithm and MCMC methods. Results of the regression analysis on 10 imputed datasets were combined to derive an overall result in each method: i.e., SAS procedure proc mianalyse was used to estimates the parameters in each method. We have not used any kind of transformation, though data violates from normality assumptions.

The data characteristics considered for the palliative care patient study are pain score, which ranges from 0 to 10, age of patients and gender. The variable "subject" is the patient's code; the age is the patient's age captured at the time of study initiation; the pain scores measured for each patient ranges from 0 to 10, with the pain score of 0 being no pain and 10 being unbearable pain; the variable "stage" indicates patient's cancer stages and ranges from stage one to stage four; the variable "diagnosis" is patient diagnosis with cancer (Head & Neck, Breast, Female genito-urinary, GIT, Lung, Male genito-urinary, Melanoma, Unknown, Haematological and Sarcoma). The study includes both male and female patients.

METHODS

Even though using only complete cases has its simplicity, you lose information. This approach also ignores the possible systematic difference between the complete cases and incomplete cases, and the resulting inference may not be applicable to the population of

all cases, especially with a smaller number of complete cases. Another strategy is single imputation, in which you substitute a value for each missing value. For example, each missing value can be imputed from the variable mean of the complete cases, or it can be imputed from the mean conditional on observed values of other variables. This approach treats missing values as if they were known in the complete case analyses. Single imputation does not reflect the uncertainty about the predictions of the unknown missing values and the resulting estimated variances of the parameter estimates will be biased toward zero [11]. Limitations of these imputation techniques in general lead to an underestimation of standard errors and, thus, overestimation of test statistics. The main reason is that the imputed values are completely determined by a model applied to the observed data, in other words, they contain no error [12]. Instead of filling in a single value for each missing value, a multiple imputation procedure replaces each missing value with a set of plausible values that represent the uncertainty about the right value to impute. These multiply imputed data sets are then analysed using standard procedures for complete data and combining the results from these analyses [13, 14].

Regression Method

In this method, the posterior predictive distribution of the parameters is used to impute the missing values for each variable. Let, continuous variable Y_j, is the response value of *Jth* patient with missing values, and the model is defined as

$$Y_j = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k \quad (1)$$

The variable is Y_j and its covariates are $X_1, X_2, \ldots X_k$. The fitted model includes the regression parameter estimates $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k)$ and the associated covariance matrix $\sigma_j^2 V_j$, where V_j is the usual $X'X$ inverse matrix derived from the intercept and covariates $X_1, X_2, \ldots X_k$. Where $\hat{\sigma}_j^2$ is the estimated variance of J^{th} patient.

Based on above approach the model for imputed values is defined as follows:

$$Pain\_score(Visit_2) = \beta_0 + \beta_1 \cdot Age + \beta_2 \cdot Sex + \beta_3 \cdot Diagnosis + \beta_4 \cdot Stage + \beta_5 \cdot Painscore(Visit_1) \quad (2)$$

$$Pain\_score(Visit_3) = \beta_0 + \beta_1 \cdot Age + \beta_2 \cdot Sex + \beta_3 \cdot Diagnosis + \beta_4 \cdot Stage + \beta_5 \cdot Painscore(Visit_1) + \beta_6 \cdot Painscore(Visit_2) \quad (3)$$

where β_0 is the intercept and $\beta_1, \beta_2, \ldots, \beta_6$ are regression coefficients.

Predictive Mean Matching Method

The continuous variables missing values are imputed through predictive mean matching procedure. The number of closest observations is specified in this method. A smaller set of observations (k_0) tends to increase the correlation among the multiple imputations for the missing observation and which results in a bigger variability of point estimators in repeated sampling. This procedure ensures that the imputed values might be more appropriate than the regression method if the normality assumption is violated.

Propensity Score Method

The propensity score method is appropriate for continuous variables when the data set is monotonically missing. Here, the conditional probability of assignment to a particular factor provides a vector of observed covariates. The score is generated for each observation to estimate the probability that the observation is missing. The observations are then grouped based on these propensity scores and on an approximated through bootstrap imputation technique.

The propensity score method uses the following steps to impute values for variable with missing values:

1. Let indicator variable R_j is *0* for observations with missing Y_j and 1 otherwise.
2. Fits a logistic regression model

$$\text{logit}(p_j) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k \quad (4)$$

where $X_1, X_2, \ldots X_k$ are the values of the covariates for Y_j, $p_j = Pr(R_j = 0 \mid X_1, X_2, \ldots, X_k)$, and $\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$

3. Creates a propensity score for each observation to estimate the probability that it is missing.
4. Divides the observations into a fixed number of groups (typically assumed to be five) based on these propensity scores.
5. Applies an approximate Bayesian bootstrap imputation to each group. In group *K*, suppose that Y_{obs} denotes the n_1 observations with non-missing Y_j values and Y_{miss} denotes the n_0 observations with missing Y_j. The approximate Bayesian bootstrap imputation first draws n_1 observations randomly with replacement from Y_{obs} to create a new data set Y^*_{obs}. This is a non-parametric analog of drawing parameters from the posterior predictive distribution of the parameters. The process then draws the n_0 values for Y_{miss} randomly with replacement from Y^*_{obs}.

Steps one through five are repeated sequentially for each variable with missing values.

It is effective for inferences about the distributions of individual imputed variables, such as a univariate analysis.

Markov Chain Monte Carlo (MCMC) Method

The MCMC is also used to generate pseudo random draws from multidimensional and otherwise intractable probability distributions via Markov chains. A sequence of random variables was generated for each element depends on the value of the previous one and information about unknown parameters through a posterior probability distribution assuming that the data follow the multivariate normal distribution.

(I) The missing values for observation *i* is denoted as $Y_{i(mis)}$ and the variables with observed values by $Y_{i(obs)}$, then the I-step draws values for $Y_{i(mis)}$ from a conditional distribution $Y_{i(mis)}$ given $Y_{i(obs)}$. (II) The P-step simulates the posterior population mean vector and covariance matrix of the complete sample estimates. These new estimates are then used in the I-step. Without prior information about the parameters, a non-informative prior is used. We can also use other informative priors.

EM Algorithm

The EM algorithm is found as an iterative procedure that finds the maximum likelihood estimate of the parameter vector by repeating the following steps:

(I) The E-step calculates the conditional expectation of the complete-data log likelihood given the observed data and the parameter estimates and (II) the M-step finds the parameter estimates to maximise the complete-data log likelihood from the E-step. The detailed steps are given below:

in the EM process, the observed-data log likelihood is non-decreasing at each iteration. For multivariate normal data, suppose there are *G* groups with distinct missing patterns. Then the observed-data log likelihood being maximised can be expressed as

$$\log L(\theta \mid Y_{obs}) = \sum_{g=1}^{G} \log L_g(\theta \mid Y_{obs}) \quad (5)$$

where $\log L_g(\theta \mid Y_{obs})$ is the observed-data log-likelihood from the g^{th} group,

$$\log L_g(\theta \mid Y_{obs}) = -\frac{n_g}{2} \log|\Sigma_g| - \frac{1}{2}\Sigma_{ig}(y_{ig} - \mu_{ig})' \Sigma_g^{-1}(y_{ig} - \mu_g) \quad (6)$$

where n_g is the number of observations in the g^{th} group, the summation is over observations in the g^{th} group, y_{ig} is a vector of observed values corresponding to observed variables, μ_g is the corresponding mean vector, and Σ_g is the associated covariance matrix.

Applications

To impute missing values for a continuous variable in data sets with monotone missing patterns, you should use either a parametric method that assumes multivariate normality or a nonparametric method that uses propensity scores [12, 15]. Parametric methods available include the regression method [12] and the predictive mean matching method [15-17]. Although the regression and MCMC methods assume multivariate normality, inferences based on multiple imputations can be robust to departures from multivariate normality if the amount of missing information is not large because the imputation model is effectively applied not to the entire data set but only to its missing part [18]. The EM algorithm is a technique that finds maximum likelihood estimates in parametric models for incomplete data [19].

RESULTS AND DISCUSSION

Table 1 depicts the baseline and disease characteristics of palliative cancer patients. Table 2 shows analgesic used by various stages of palliative cancer patients across the visits. There is no significant difference in various demographic, disease characteristics and pain score status between the two sub populations, i.e. cases with missing and non-missing respectively. Table 3 displays the relationship between change in pain score at visit three and other baseline, and disease

TABLE 1. Summary of Baseline Characteristics

CHARACTERISTICS	STATISTICS/ CATEGORY	N=326
Age	N	326
	Mean(SD)	60(12.2)
	Min	14
	Max	90
Diagnosis	Head & Neck	56(17.2)
	Female genito-urinary	82(25.2)
	Male genito-urinary	20(6.1)
	GIT	46(14.1)
	Sarcoma	7(2.1)
	Breast	27(8.3)
	Haematological	8(2.5)
	Lung	59(18.1)
	Unknown	18(5.5)
	Melanoma	3(0.9)
Sex	Male	157(48.2)
	Female	169(51.8)
Stage	1	3(0.9)
	2	23(7.1)
	3	113(34.7)
	4	187(57.4)

TABLE 2. Summary of Type of Analgesic used by stage of disease

VISIT	TYPE OF ANALGESIC	STAGE 1	STAGE 2	STAGE 3	STAGE 4
Visit 1	Hydromorphone	0	5	35	47
	Methadone	0	2	9	19
	Fentanyl	2	11	54	93
	Morphine	1	4	14	23
	Diamorphine	0	0	1	1
	Missing	0	1	0	4
Visit 2	Hydromorphone	0	2	13	20
	Methadone	0	0	7	14
	Fentanyl	2	6	48	62
	Morphine	1	8	19	34
	Diamorphine	0	0	0	1
	Missing	0	7	26	56
Visit 3	Hydromorphone	0	2	9	11
	Methadone	0	0	4	7
	Fentanyl	1	6	25	43
	Morphine	2	6	24	26
	Diamorphine	0	0	0	1
	Missing	0	9	51	99

characteristics using regression analysis by comparing various imputation techniques with Complete Case approach. The estimate (SE) obtained without any imputation technique (i.e. Complete Case method) for parameter age is 0.025791 (0.0216). Whereas, Predictive Mean Matching Method slightly over estimated age parameter with 0.034 (0.0181). The Propensity Score Method gives precise estimates with less variation 0.0102 (0.0179).

The similar trend can be found for most of the categories for diagnosis, i.e. the Propensity Score Method gives slightly low variation for parameter estimates as compared to other methods. When a complete case approach considered in regression analysis, it was found that parameter estimate was slightly overestimated -0.811 (0.6939) as compared to other imputation techniques. However, estimates obtained through MCMC method give less variation for gender, i.e. -0.324 (0.5434).

Across all four stages of cancer, it was found that MCMC method estimates parameter with less variation as compared to other imputation techniques.

Table 4 summarises the pain score values imputed by various techniques at visit two and three respectively. The mean (SD) calculated through Predictive Mean Matching Method found to be precise as compared to other techniques for visit two and three.

The Mean (SD) of observed data was 3.638 (3.175) whereas the imputed Mean (SD) values were 3.356 (2.6603), 3.502 (2.6100), 3.406 (2.4334), 3.474 (2.6285) and 3.264 (2.6336) respectively for the imputation methods Regression, Predictive Mean Matching, Propensity Score, EM algorithm and MCMC methods for pain score values at visit three (detailed in Table 4). The Mean (SD) of observed data was 3.528 (3.1112) whereas the imputed Mean (SD) were 3.231 (2.8715), 3.253 (2.8691), 3.278 (2.7935), 3.268 (2.8725) and 3.227 (2.8952) respectively, for the imputation methods Regression, Predictive Mean Matching, Propensity Score, EM algorithm and MCMC methods for pain score values at visit two (detailed in Table 4). Figure 1 describes mean pain score obtained by various imputation techniques for visit two and visit three.

The Propensity Score Method ensures that imputed values are plausible and might be more appropriate than the Regression Method if the normality assumption is violated [20].

CONCLUSION

It should be emphasised that sophisticated statistical analysis is not a substitute for a good clinical plan in order to mitigate subjects dropping out of a study. It is important to continue following subjects even after they have dropped out of a clinical research. In addition, understanding both the disease and the therapy being studied can be helpful in selecting an appropriate statistical method [9]. The choice of a particular method for handling missing data depends on whether one is considering a more pragmatic or a more explanatory perspective [10]. In this study, we have compared the parameter estimates obtained from

TABLE 3. Relationship between change in pain score at visit three and other baseline characteristics (Regression Analysis): a comparison between various imputation techniques.

		COMPLETE CASE METHOD (N=174)		REGRESSION METHOD (N=326)		PROPENSITY SCORE METHOD (N=326)		PREDICTIVE MEAN MATCHING METHOD (N=326)		EM ALGORITHM METHOD (N=326)		MCMC METHOD (N=326)	
Characteristics	**Category**	**Estimate (Standard Error)**	**P-value**	**Estimate (Standard Error)**	**P-value**	**Estimate (Standard Error)**	**P-value**	**Estimate (Standard Error)**	**P-value**	**Estimate (Standard Error)**	**P-value**	**Estimate (Standard Error)**	**P- value**
Intercept		-0.383 (2.5577)	0.8812	-0.472 (2.3761)	0.8428	0.534 (2.2529)	0.8128	-0.825 (2.1419)	0.7	-0.571 (2.1682)	0.79	-0.171 (2.2382)	0.9391
Age		0.025791 (0.0216)	0.2336	0.025 (0.02143)	0.2541	0.0102 (0.0179)	0.5697	0.034 (0.0181)	0.0606	0.027 (0.0245)	0.28	0.023 (0.0215)	0.2752
Diagnosis	Head & Neck	-0.857 (2.1950)	0.6968	-1.299 (1.9956)	0.5156	-1.644 (2.0049)	0.4123	-1.011 (1.9482)	0.6038	-1.210 (2.0348)	0.55	-1.455 (1.9916)	0.4654
	Female genito-urinary	-0.679 (2.2970)	0.7677	-0.612 (2.0236)	0.7627	-0.614 (2.0356)	0.763	-0.428 (2.1183)	0.84	-0.445 (2.1504)	0.84	-0.888 (2.0005)	0.6571
	Male genito-urinary	-0.235 (2.3093)	0.919	-0.889 (2.1527)	0.68	-0.715 (2.1537)	0.74	-0.530 (2.0781)	0.7989	-0.679 (2.2531)	0.76	-1.063 (1.9708)	0.5897
	GIT	0.071 (2.2418)	0.9749	-0.335 (2.1076)	0.874	-0.209 (2.0433)	0.9186	-0.097 (1.8797)	0.9585	-0.232 (2.0656)	0.91	-0.547 (1.9500)	0.779
	Sarcoma	-0.351 (2.6652)	0.8952	-0.657 (2.3801)	0.7827	-1.090 (2.3539)	0.6433	-0.334 (2.3786)	0.8884	-0.439 (2.5918)	0.87	-0.673 (2.2127)	0.7607
	Breast	-0.831 (2.4429)	0.7341	-0.977 (2.1447)	0.6492	-0.843 (2.0563)	0.6816	-0.744 (2.2027)	0.7359	-0.453 (2.4248)	0.85	-1.049 (2.0567)	0.6102
	Haematological	1.877 (2.5532)	0.4633	0.677 (2.3049)	0.7691	0.359 (2.3602)	0.879	0.793 (2.1677)	0.7145	0.864 (2.2616)	0.7	0.458 (2.2213)	0.8366
	Lung	-0.280 (2.1957)	0.8984	-0.405 (2.0813)	0.846	-0.582 (2.0636)	0.7779	-0.231 (1.9155)	0.9038	-0.238 (2.0132)	0.91	-0.657 (1.9442)	0.7352
	Unknown	1.540 (2.3744)	0.5175	0.397 (2.2807)	0.8623	0.085 (2.1538)	0.9685	0.395 (2.0518)	0.8474	0.705 (2.1204)	0.74	0.177 (2.1440)	0.9341
	Melanoma	0	.	0	.	0	.	0	.	0	.	0	.
Sex	Male	-0.811 (0.6939)	0.2442	-0.191 (0.6368)	0.7652	-0.120 (0.5860)	0.8373	-0.624 (0.5656)	0.2728	-0.335 (0.6625)	0.62	-0.324 (0.5434)	0.5508
	Female	0	.	0	.	0	.	0	.	0	.	0	.
Stage	1	1.253 (1.8454)	0.4981	1.632 (1.8746)	0.384	1.855 (1.9968)	0.3527	1.235 (1.8443)	0.5029	1.559 (1.8579)	0.4	1.588 (1.8281)	0.3849
	2	0.551 (0.8564)	0.5209	0.669 (0.8779)	0.4491	0.458 (0.9198)	0.6194	0.442 (0.8260)	0.5937	0.726 (0.9223)	0.44	0.836 (0.7479)	0.2642
	3	0.512 (0.5126)	0.3189	0.789 (0.5222)	0.1395	0.770 (0.4911)	0.1207	0.574 (0.4515)	0.2074	0.707 (0.4804)	0.15	0.563 (0.4289)	0.1915
	4	0	.	0	.	0	.	0	.	0	.	0	.

different multiple imputation techniques. The Propensity Score Method has appeared to be the most appropriate method for the pain score data. The multiple imputation techniques also have few advantages; since, the imputed values are drawn from a distribution, they inherently contain some variation by introducing an additional form of error in the parameter estimates across the imputation. It replaces each missing item with two or more acceptable values, representing a distribution of possibilities. Multiple imputations have the same optimal properties as ML, and it removes some of its limitations [21]. Multiple imputations can be used with any kind of data and model with conventional software. When the data is MAR, multiple imputations can lead to consistent, asymptotically efficient, and asymptotically normal estimates.

Multiple imputations have limitations. It is a bit challenging to successfully use it. It produces different estimates (hopefully, only slightly different) every time you use it, which can lead to situations where different researchers get different numbers from the same data using the same method [20, 14].

Acknowledgement

We wish to thank to Dr. Biji, MCC for providing the data used in this paper. The authors would like to express sincere gratitude to two anonymous referees and support of the native English speaker for their cautious reading and constructive suggestions which led to improvement over the earlier versions of the manuscript.

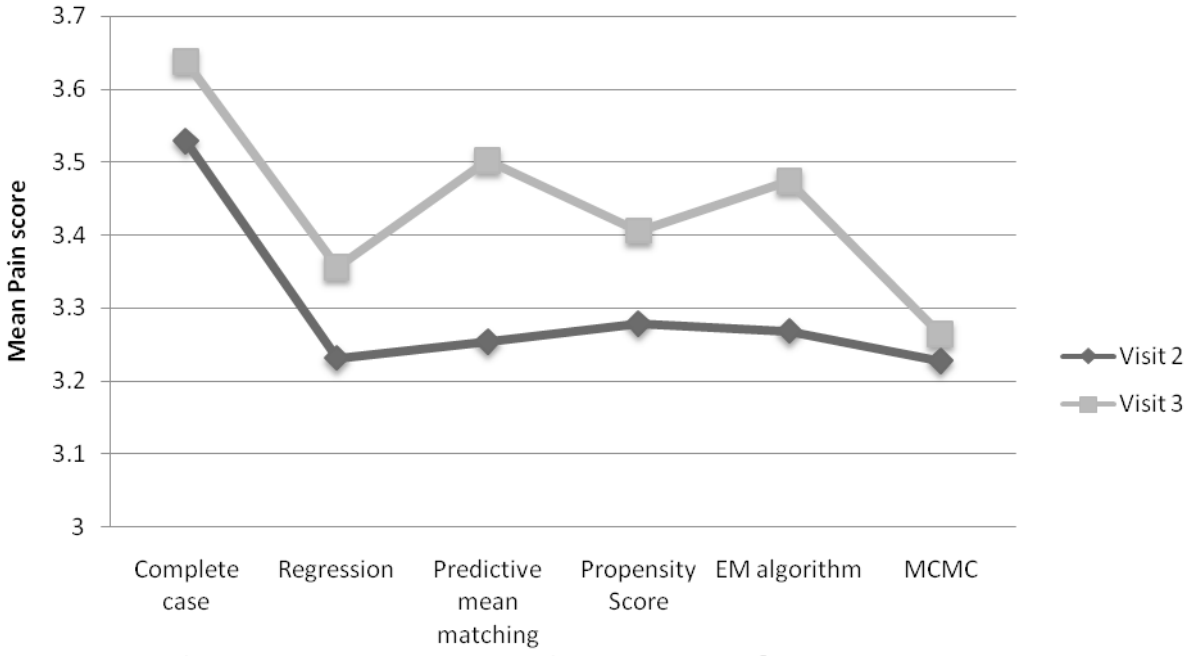
Conflict of Interest

Presentation in AROICON-2015, Abstract: 218.

TABLE 4. Summary of Imputed values through different techniques

VARIABLE	STATISTICS	COMPLETE CASE METHOD	REGRESSION	PREDICTIVE MEAN MATCHING	PROPENSITY SCORE	EM ALGORITHM	MCMC
Pain score at visit 3	N	174	326	326	326	326	326
	Mean	3.638	3.356	3.502	3.406	3.474	3.264
	SD	3.175	2.6603	2.6100	2.4334	2.6285	2.6336
Pain score at visit 2	N	244	326	326	326	326	326
	Mean	3.528	3.231	3.253	3.278	3.268	3.227
	SD	3.1112	2.8715	2.8691	2.7935	2.8725	2.8952

FIGURE 1. Average Pain score values



References

1. Morey, et al. Cancer and Its Psychosomatic Ill Effects - A Vision In Ayurveda. Sangyaharan Shodh 2015, 18(2).
2. Nath DC, Bhattacharjee A. Pattern Mixture Modeling. An Application in Anti-Diabetes Drug Therapy on Serum Creatinine on Type 2 Diabetes Patients. Asian Journal of Mathematics and Statistics 2012; 5(3):71-81.
3. Little RJA. Modeling the drop-out mechanism in repeated measures studies. J Am Stat Assoc 1995 (90):1112-21.
4. Rubin DB. Inference and missing data. Biometrika 1976;(63):581-92.
5. Laird NM. Missing data in longitudinal studies. Stat Med 1988;(7):305-15.
6. Diggle P, Kenward MG. Informative dropout in longitudinal data analysis (with discussion). Appl Stat 1994;(43):49-94.
7. Little RJA, Rubin DB. Statistical Analysis with Missing Data. New York: John Wiley and Sons, 1987.
8. Little RJA. A test of missing completely at random for multivariate data with missing values. J Am Stat Assoc 1988;(83):1198-202.
9. William R Myers. Handling missing data in clinical trials: an overview. Drug Information Journa 2000;(34):525–33.
10. Little RJA, Schenker N. Handbook of Statistical Methodology Missing Data. New York: Plenum Press, 1995.
11. Heyting A, Tolboom J, Essers J. Statistical handling of drop-outs in longitudinal clinical trials. Stat Med 1992;(11):2043-61.
12. Allison PD. Missing Data. Thousand Oaks, CA: Sage Publications, 2001
13. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons, Inc. 1987
14. Yang C Yuan. Multiple Imputation for Missing Data: Concepts and New Development. SAS Institute Inc.
15. Lavori PW, Dawson R, Shera D. A Multiple Imputation Strategy for Clinical Trials with Truncation of Patient Data. Statistics in Medicine 1995;(14):1913–25.
16. Heitjan F, Little RJ. A Multiple Imputation for the Fatal Accident Reporting System. Applied Statistics 1991;(40):13-29.
17. Schenker N, Taylor JMG. Partially Parametric Techniques for Multiple Imputation. Computational Statistics and Data Analysis 1996; 22:425–46.
18. Schafer JL. Analysis of Incomplete Multivariate Data, New York: Chapman and Hall, 1997.
19. Dempster AP, Laird NM, Rubin DB. Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society, Ser. B. 1977;(39):1-38.
20. Nakai M, Weiming K. Review of Methods for Handling Missing Data in Longitudinal Data Analysis. Int Journal of Math Analysis 2011;(5)1:1-13.
21. Horton NJ, Lipsitz SR. Multiple Imputation in Practice: Comparison of Software Packages for Regression Models with Missing Variables. The American Statistician 2001;(55):244–54.

*