Comparison of generalized estimating equations (GEE), mixed effects models (MEM) and repeated measures ANOVA in analysis of menorrhagia data

Parisa Naseri [1], Hamid Alavi Majd [1,] *, Nourossadat Kariman [2], Atefeh Sourtiji [3]

[1]Department of Biostatistics, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
[2]Department of Midwifery and Reproductive Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
[3]Department of Midwifery, Babol Islamic Azad University, Babol, Iran.

*Corresponding Author: email address: alavimajd@gmail.com (H. Alavi Majd)

ABSTRACT

Menorrhagia is one of the most common gynecological problem and leading causes of poor quality of life and iron deficiency anemia in women of reproductive age. Research in gynecological field relies heavily on repeated measure designs. Repeated measure studies are helpful in understanding how factors of interest change over time. Our goal is to apply statistical methods which are appropriate for analyzing repeated measure data such as gynecological data. Three statistical methods were performed by data collection from 100 patients with menorrhagia. One-hundred patients were randomly assigned to two groups, i.e. intervention group (Urtica Dioica and mefenamic acid) and control group (placebo and mefenamic acid) with an equal size of 50. In this study, generalized estimating equations (GEE) and mixed effects models (MEM) were used for analyzing menorrhagia data to determine the effect of hydroalcoholic extract of Urtica Dioica on Menorrhagia. Finally, these methods are compared to the conventional repeated measures ANOVA (RM-ANOVA).

Based on the results, the three methods are found to be similar in terms of statistical estimation, the amount of bleeding before and after treatment between and within groups was compared. Results showed the average amount of bleeding was reduced significantly (P<0/001). The average menorrhagia score in the third month (second cycles after intervention) were 91.38(71.432) and 149.40(127.823) in Urtica Dioica and control groups, respectively. The difference between the two groups was statistically significant (p =0.036). Because their advantages, GEE and MEM should be strongly considered for the analysis of repeated measure data. In particular, GEE should be utilized to explore overall average effects. When in addition to overall average effects, subject-specific effects are of primary interest, MEM should be utilized. With respect to these methods, it seems the extract of Urtica Dioica can be effective in reducing the amount of menstrual bleeding in women of reproductive age with Menorrhagia.

Keywords: Repeated Measure-ANOVA; Generalized Estimating Equations; Mixed Effects Models; Repeated measure study; Menorrhagia

INTRODUCTION

Menorrhagia is an abnormally heavy and prolonged menstrual period at irregular intervals and is an important cause of ill health in women. A normal menstrual cycle is 21–35 days in duration, with bleeding lasting an average of 5 days and total blood flow between 25 and 80 mL. A blood loss of greater than 80 mL or lasting longer than 7 days constitutes menorrhagia[1]. It is one of the most common gynecological problem[2] and leading causes of poor quality of life and iron deficiency anemia in women of reproductive age[3]. More than 30 percent of women of reproductive age and 60 percent of the total female population experience menorrhagia [2] and 44.7 % of Iranian girl experience it, too[4]. Often the bleeding is severe enough to adversely affect the patient's social, physical, and emotional well-being[5]. Menorrhagia accounts for about 25% of all gynecological surgeries[6]. It described as excessive uterine bleeding, affects approximately 10 million American women annually, most of whom are in their 40's and 50's and accounts for 20% of visits to gynecologists[7]. It reduces quality of life and caused Discomfort and anxiety [2, 8-10].

In recent year, medical treatment, surgical and pharmaceutical plants have been used for the treatment of menorrhagia. It seems, the extract of Urtica Dioica can be effective in reducing the amount and duration of menstrual bleeding in women of reproductive age with menorrhagia.

In this study, we consider a two-group comparative (i.e. treatment vs. control) Repeated measure or longitudinal study (LS) in clinical example.

Here we use a two-group design but statistical inferences can be generalized for different number of groups.

LS design, involving consecutive measurements on the same individual, has become popular for examining trends in outcomes over time. Compared to a cross-sectional study, which assesses the outcome at a single point in time, repeated measure study can provide information about changes in both individual and average group outcomes over time. Several factors need to be considered in the statistical analysis of repeated measures data. The measurements obtained from an individual at a particular point in time are correlated with the measurements obtained at prior time point(s) and this dependence needs to be taken into account in the analysis. Often, these data contain missing values due to loss to follow up or other reasons, such as the participant being unavailable to have the outcome measured at some time point(s). These missing data result in information from different individuals available at different time points, that is, an unbalanced design and also existence of multiple covariates. If these complexities are properly addressed, a LS has the advantage of being able to answer important and clinically questions with higher precision than a study with simpler design[11].

All of these factors have resulted in the growing popularity of repeated measure design and have become important reasons for the development of new statistical methods during the past few decades. The relatively new and more advanced methods of Generalized Estimating Equations (GEE) [12] and Mixed Effects Models (MEM) [13] have started to replace the traditional methods of Repeated Measures ANalysis Of VAriance (RM-ANOVA) [14] and t-test, as the older methods are not flexible enough to accommodate all of features of repeated measure designs. Repeated measure studies are particularly important in the field of gynecological research.

Therefore, in this manuscript, we apply RM-ANOVA, GEE, and MEM in simple terms for analyzing data from a clinical study with continuous outcome and compare the results from each method.

The research questions we focus on are: Is there an improvement in the menorrhagia over time in the group that received Urtica Dioica?

We selected the menorrhagia as our continuous outcome measure, which was derived from the Pictorial Blood Loss Assessment Chart (PBLAC) [15].

A randomized triple blinded clinical trial was carried out on 100 women affected by menorrhagia, selected by convenience sampling, which had inclusion criteria. Data collection tools were data form, weight, meter and PLBAC chart. Measurements from 100 women were taken at three month so that, first month is before intervention, second and third month are first and second cycles after intervention, respectively.

MATERIALS AND METHODS

In this paper, we consider a two-group comparative (i.e. Urtica Dioica vs. control) repeated measure design in clinical example. Although here we use a two-group design, statistical inferences can be generalized for repeated measure design involving a different number of groups.

Assessing the change in outcome measure over time

Figures show that means of the menorrhagia scores at each time point across all individuals.

For repeated outcome measures where the individual measurements are correlated, the visualization of means and confidence intervals at different time points does not provide an adequate representation of the change in the

outcome over time within individuals. The within-individual variation in repeated outcome measures can be explored by joining the measurements of the same individual over time.
Figure 1, commonly known as a spaghetti plot, shows individual trajectories along with superimposed means of the menorrhagia scores at each time point for the two groups. For the purpose of illustration, in Figure 1, we have included trajectories of all participants from the study. The change in menorrhagia scores over time seems to differ substantially between the two groups. For Urtica Dioica group, variability across time points seems to decrease relative to control group, this shows Urtica Dioica is effective treatment for menorrhagia.

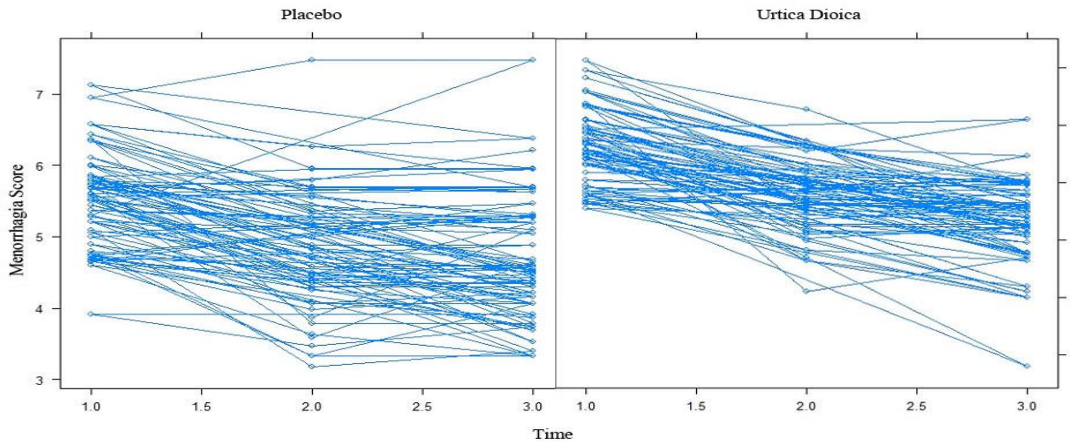

Figure 1. Spaghetti plot showing individual trajectories of the menorrhagia scores for two groups Urtica Dioica and Placebo

Conceiving correlations of the outcome measure between time points

Scatter plots [16] of the outcome measure assessed at one time point versus the outcome measured at another time point (e.g. menorrhagia score at first month vs menorrhagia score at second month) allow exploration of the within-individual correlation between the different times of follow up. Within-individual correlations of the menorrhagia scores among the three month of assessment for the two groups are shown in Figure 2. There seems to be strong correlations between the assessments, and these correlations do not systematically increase or decrease with the time lag between the assessments.

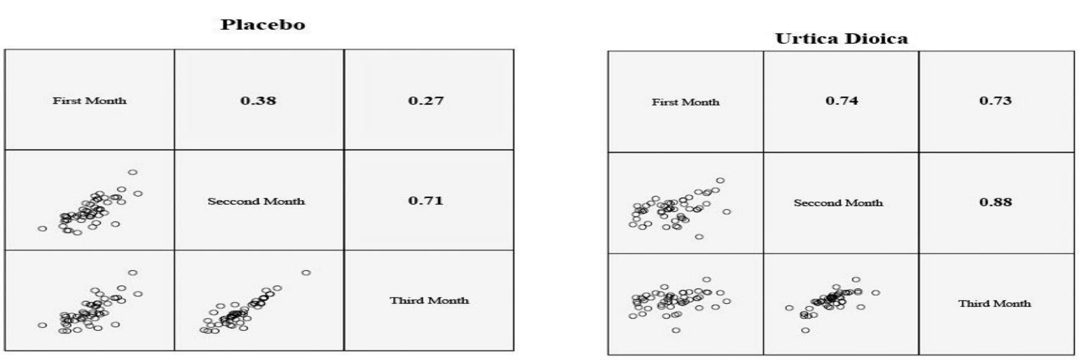

Figure 2. Scatter plots showing within-individual correlations of the menorrhagia scores between time points for the two groups.

Statistical Analysis of the Data

Repeated Measures ANOVA

RM-ANOVA relates the study outcome variable to a set of covariates (e.g. treatment group, time) and compares the mean outcome at multiple time points or between groups. Although RM-ANOVA (one of the earliest proposed methods for analyzing correlated responses) has gained widespread popularity, it has several unattractive features. However, the weaknesses of this method for analyzing repeated measure data are well documented in the literature [17, 18]. First, RM ANOVA requires the outcome variable to be quantitative (i.e. a continuous variable) and normally distributed. It also requires the covariates to be discrete (i.e. categorical variables). Second, RM-ANOVA requires that the outcome have constant variance across time points as well as constant correlation between any two time points (i.e. assumption of sphericity). The assumption of constant correlation of repeated measures is often unrealistic in medical research as repeated measures often become less correlated with increasing time from treatment. This kind of violation of the sphericity assumption may cause inflated Type I error[19]. Third, RM-ANOVA can only handle repeated measure studies in which all subjects have the same numbers of repeated measurements. Specifically, RM-ANOVA excludes those subjects who have missing observations at one or more time points. Inclusion of only those subjects who have complete data for all variables has unfavorable subsequences. The group of subjects with complete data may not represent a random sample from the target population, thus causing biased results.

Generalized estimating equations (GEE)

The GEE method focuses on average changes in response over time and the impact of covariates on these changes. The method models the mean response as a linear function of covariates of interest via a transformation or link function. To accommodate various types of outcomes that are not necessarily normally distributed, different link functions are employed for modeling the relationship between outcome and covariates. For example, an identity link function is used for a continuous outcome, a logit link function for a binary outcome and a log link function for count data[20]. These transformations can be considered repeated measures analogs of linear regression, logistic regression, and Poisson regression, respectively. GEE allows specification of the correlation structure from a wide variety of choices to account for variation in correlation between repeated measures. Popular choices, among others, include the compound symmetry (CS) correlation structure and the autoregressive (AR(1)) correlation structure. The CS correlation structure assumes a common correlation for any pair of responses at different time points, while the AR correlation structure assumes that measurements closer in time have a higher correlation than those that are further apart. GEE also has attractive and robust property in parameter estimation. Unlike RM-ANOVA, GEE does not require the outcome variable to have a particular distribution. This feature can greatly benefit studies in which data are skewed or the distribution of data is difficult to verify due to a small sample size.

Mixed effects models (MEM)

MEM describes how the response of the individual participant changes over time. It considers between-individual heterogeneity by adding random effects to a subset of covariates of interest. These added random effects allow covariate coefficients to vary randomly from one individual to another, thereby providing an individual response trajectory over time. The most common MEM in repeated measure studies are those with random effects attached to baseline values or time dependent variables reflecting heterogeneity among individual responses at baseline or variation between individual trajectories over time. In addition, like GEE, MEM allows specification of the correlation structure between repeated measurements from similar choices such as the CS and AR(1).

Clinical data analysis

Study [21] was conducted to determine the effect of hydroalcoholic extract of Urtica Dioica on menorrhagia among students of Babol Azad University 2012-13.

A randomized triple blinded clinical trial was carried out on 100 women affected by menorrhagia, selected by convenience sampling, which had inclusion criteria. Data collection tools were data form, weight, meter and PLBAC chart. All samples after a control cycle of primary bleeding and dividing with random allocation to trial or control group were subjected to mefenamic acid treatment (500 mg every 8 hours) and Urtica Dioica 5 cap. Per day, from first to end of bleeding, up to 7 days, for two consecutive cycles, for trial group and for the control group, mefenamic acid and placebo as the same way of trial group, was prescribed.

Menorrhagia was measured at first, second and third month postoperatively. Menorrhagia will be used as the continuous outcome.

This data provides an illustration of repeated measure data analysis for menorrhagia.

For illustrative purposes time, treatment group and the interaction between time and treatment were included in the models. All statistical analyses were performed in SAS software version 9.2(SAS Institute, Cary, NC), R 2.14.0 and SPSS (version 21; SPSS, Chicago, IL, USA).

"This research has been approved by Shahid Beheshti University of Medical Sciences's Medical Ethics Committee (Referred Code: 201111254529N3)"

RESULTS

The demographic and fertility characteristics of the study population are shown in Table 1. The average of age, BMI, Menarche age and Number of childbirth were not significant between control and Urtica Dioica group, so respect to the factors mentioned above, the two groups were matched.

Table 1. Demographic and fertility characteristics of the study population

Variable	Control group(n=50) Mean(SD)	Urtica Dioica group(n=50) Mean(SD)	P-value
Age	26.71(7.25)	25.24 (6.75)	0.324
BMI	23.36 (3.30)	22.36 (2.94)	0.228
Menarche age	13.33 (1.39)	13.53 (1.40)	0.546
Number of childbirth	0.56 (0.81)	0.29 (0.58)	0.092

The average of menorrhagia score in the third month (second cycles after intervention) was 91.38(71.432) and 149.40(127.823) in Urtica Dioica and control groups, respectively. The difference between the two groups was statistically significant (p =0.036), results are shown in table2.

Table 2 .Menorrhagia score of study population before intervention (first month) and after intervention (second and third month)

Menorrhagia score	Control group(n=50) Mean (SD)	UrticaDioica group(n=50) Mean (SD)	P-value
First month	303.98 (213.839)	315.16 (191.498)	0.412
Second month	149.51 (115.381)	113.07 (62. 798)	0.182
Third month	149.40 (127.823)	91.38 (71. 432)	0.036

The obtained results showed that the amount of bleeding reduced significantly in both groups after treatment, the reduction of bleeding amount in the Urtica Dioica group was significantly higher than that in the control group that figure 3 shows this, too.

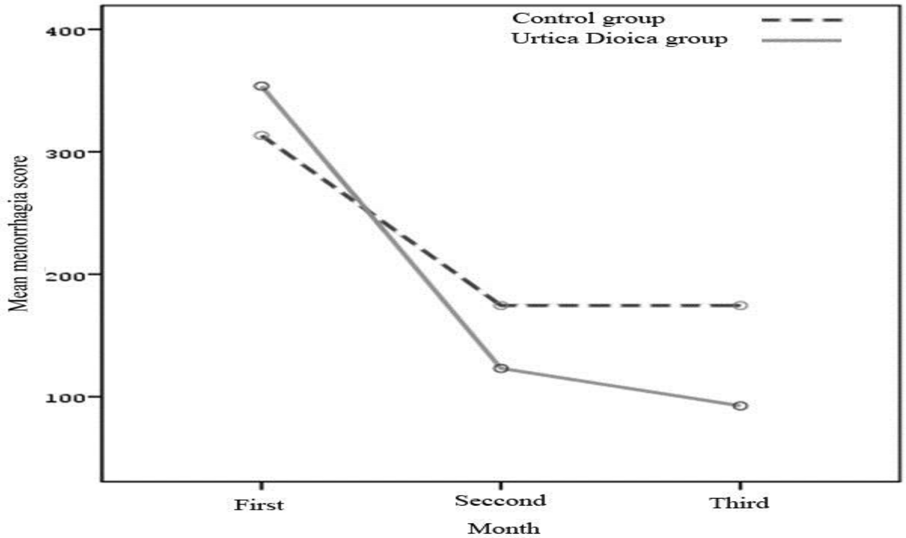

Figure 3. Mean of menorrhagia scores for two groups Urtica Dioica and Placebo at first month (before intervention), second month (first cycle after intervention) and third month (second cycle after intervention)

Table 3 reports the findings from all three methods in treatment effect, time effect and interaction effect of treatment and time. All three methods showed similar result, time, treatment and interaction term were significant.

Table 3. Estimated effects of treatment, time, and treatment × time on the menorrhagia

RM-ANOVA			
Effect	**Treatment**	**Time**	**Treatment× Time**
P-value	<0.001	<0.001	<0.001
MEM			
Effect	**Treatment**	**Time**	**Treatment× Time**
Parameter Estimate	0.4	-0.41	-0.26
Standard error	0.15	0.05	0.08
P-value	0.0098	<0.0001	0.0011
GEE			
Effect	**Treatment**	**Time**	**Treatment× Time**
Parameter	0.39	-0.41	-0.26
Standard error	0.16	0.04	0.07
P-value	0.0178	<0.0001	0.0003

Significancy the interaction between time and treatment showed that Urtica Dioica can be effective in reducing the amount of menstrual bleeding in women of reproductive age with menorrhagia.

GEE and MEM produced similar parameter estimates and both had population-averaged interpretations. For example, average increases of 0.13 (0.39-0.26) mL and 0.14 (0.4-0.26) mL menorrhagia in the treatment group compared to the control group, respectively, were identified using GEE and MEM.

While GEE and MEM estimates the magnitude of these effects and tests their significances, RM-ANOVA only assesses significance. The reported MEM has random effects attached to baseline Menorrhagia and time in order to consider the between-patient variations. The assumption of sphericity in RMANOVA is violated ($p<0.001$), showing non-constant variances and correlations over time.

The AR(1) correlation structure is specified in GEE and MEM to provide a more pragmatic modeling strategy.

DISCUSSION

Through the application of the three methods, we were able to show that GEE and MEM are more flexible than RM-ANOVA for investigating continuous outcomes and modeling a variety of correlation patterns between repeated measures. There are other important differences among the three approaches.

Besides the noted differences, the three methods were used to answer different research questions. With data collected from every individual at every time point, we wanted to make statistical inferences regarding the change in mean response over time (a population averaged inference) or the individual trajectory over time (a subject-specific inference). Both RM-ANOVA and GEE measured population-averaged effects of covariates of interest. MEM by contrast, could identify subject-specific effects of covariates on the changes in the response over time. Therefore, MEM would be helpful where an intervention is likely to affect some individuals differently than others as compared to RM-ANOVA and GEE which do not take individual response into account in their interpretations. MEM could allow for a more various analysis of individuals of this sub-population such as predicting individual risk of complications.

Yan and et.al in their study compared GEE, MEM and RM-ANOVA through study conducted in the Department of Anesthesiology at the Hospital for Special Surgery to assess limb preconditioning on reduction pain. They provided an illustration of longitudinal data analysis for C-reactive protein (CRP) as a continuous outcome[11].

Cleophas and et.al reviewed Methods for analyzing cardiovascular studies with repeated measures. They compared MEM with RM-ANOVA [22].

Livera and et.al using data collected from a trial evaluating an intervention for managing asthma and chronic obstructive pulmonary disease, demonstrated ways of statistically analyzing repeated measures data. They applied MEM and RM-ANOVA [16].

The purpose of this paper was to present methods to consider the special nature of repeated measures data in assess the effect of Urtica Dioica in reducing the amount of bleeding in women of reproductive age with heavy menstrual bleeding. This drug, due to its short course of prescription, fewer side effects and not having hormonal effects, seems to be an effective drug for treating heavy menstrual bleeding. It is a well known ethnomedicinal plant and in most parts of the world is used to treat various diseases, it includes different chemical products, such as caffeic malic acid, Polysaccharides, Lectin, Agglutinin, Scopoletin, Serotonin and D, C and K vitamins [23, 24].

Farahmand and et.al in their study showed that both vitamin E and Ibuprofen reduces bleeding amount in primary dysmenorrheal and does not cause any digestive disorder or fatigue [25].

Rahnama and et.al showed that ginger may be an effective and safe therapy for relieving pain in women with primary dysmenorrhoea if administered at the onset or during the 3 days prior to menstruation[26].

Safari and et.al indicated both vitamin E and mefenamic acid could reduce the amount of bleeding similarly, nevertheless complications of vitamin E consumption, such as digestive disorders like, aspyrosis, stomach-ache, nausea and fatigue were significantly less compared to mefenamic acid [27].

Most of results from other studies were comparable to our results. Obtained results from this study, showed that the extract of Urtica Dioica can be effective in reducing the amount of menstrual bleeding but difference between two groups at second month weren't statistically significant. Maybe, one reason for the lack of significance between the two groups was short follow-up period and consumption Urtica Dioica over a short period of time (only two month), so for Achieving more accurate results, study with relatively long follow-up period is suggested.

ACKNOWLEDGMENTS

This work was supported by the Biostatistics department of Shahid Beheshti University of Medical Sciences. We would like to express our gratitude to all the students of Babol Azad University who participated in the study.

"The authors declare no conflict of interest"

REFERENCES

1.Berk J, Novak (2012). Berek &Novak gynecology. Vol. 1. Fifteenth edition. Lippincott Williams & Wilkins. Philadelphia. Pp: 404-409.

2.Ghazizadeh S, Bakhtiari F, Rahmanpour H, Davari-Tanha F, Ramezanzadeh F. A randomized clinical trial to compare levonorgestrel-releasing intrauterine system (Mirena) vs trans-cervical endometrial resection for treatment of menorrhagia. International journal of women's health. 2011;3:207.

3.Moukhah S, Mazari Z, Goshtasbi A, Moaed Mohseni S. The effect of tranexamic acid on the quality of life and blood loss of women with menorrhagia: A clinical trial. Arak Medical University Journal. 2012;15(3):75-84.

4.Farahmand M AZ, Rasekh A. Effect of Vitamine E and Iboprofen on Intensity of Heamorrage in primary dysmenorrea. Kosar Medical Journal. 2010;14: 49-52.

5.Levy R, Prasad S, Rowland K. Consider this option for heavy menstrual bleeding. The Journal of family practice. 2011;60(7):410.

6.Chattopdhyay B, Nigam A, Goswami S, Chakravarty P. Clinical outcome of levonorgestrel intra-uterine system in idiopathic menorrhagia. Eur Rev Med Pharmacol Sci. 2011;15(7):764-8.

7.Livdans-Forret AB, Harvey PJ, Larkin-Thier SM. Menorrhagia: A synopsis of management focusing on herbal and nutritional supplements,
and chiropractic. The journal of the Canadian chiropractic association. 2007;51(4):235.

8.Marjoribanks J, Lethaby A, Farquhar C. Surgery versus medical therapy for heavy menstrual bleeding. The Cochrane Library. 2006.

9.Lethaby A, Irvine GA, Cameron IT. Cyclical progestogens for heavy menstrual bleeding. The Cochrane Library. 2008.

10.Rahimi M ZS, Babai Gh, Rabiy M. Effect of vitamin E in menorrhagia treatment in patients with dysfunctional uterine bleeding. Scientific medical Daneshvar Journal. 2009;82: 59-66.

11.Ma Y, Mazumdar M, Memtsoudis SG. Beyond Repeated measures ANOVA: advanced statistical methods for the analysis of longitudinal data in anesthesia research. Regional anesthesia and pain medicine. 2012;37(1):99.

12.Zeger SL, Liang K-Y. Longitudinal data analysis for discrete and continuous outcomes. Biometrics. 1986:121-30.

13.Demidenko E. Mixed models: theory and applications with R: John Wiley & Sons; 2013.

14.Girden ER. ANOVA: Repeated measures: Sage; 1992.

15.Higham JM, O'brien P, Shaw R. Assessment of menstrual blood loss using a pictorial chart. BJOG: An International Journal of Obstetrics & Gynaecology. 1990;97(8):734-9.

16.De Livera AM, Zaloumis S, Simpson JA. Models for the analysis of repeated continuous outcome measures in clinical trials. Respirology. 2014;19(2):155-61.

17.Vermeulen KM, Post WJ, Span MM, van der Bij W, Koëter GH, TenVergert EM. Incomplete quality of life data in lung transplant research: comparing cross sectional, repeated measures ANOVA, and multi-level analysis. Respiratory research. 2005;6(1):101.

18.Krueger C, Tian L. A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points. Biological research for nursing. 2004;6(2):151-7.

19.Huynh H, Feldt LS. Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. Journal of Educational and Behavioral Statistics. 1976;1(1):69-82.

20.McCullagh P, Nelder JA. Generalized linear models: CRC press; 1989.
21.Considering the effect of hydroalcoholic extract of Urtica Dioica on menorrhagia. 2012.
22.Cleophas T, Zwinderman A, van Ouwerkerk B. Methods for analysing cardiovascular studies with repeated measures. Netherlands Heart Journal. 2009;17(11):429-33.
23.Gruenwald J, Brendler T, Jaenicke C. PDR for Herbal Medicines. Montvale, NJ: Medical Economics Co. Inc; 1998.
24.Sourteji A KN, Mojab F, Alavi Majd H Comparing the effect of mefenamic acid and the hydroalcoholic extract of Urtica dioica on the volume of heavy menstrual bleeding in Babol Azad University students (2011-2012). Arak Medical University Journal (AMUJ). 2013; 16((73)):27-36.
25.Farahmand M, zahediasl m, abbaspour z, rasekh a. Effects of vitamin e and ibuprofen on amount, duration and interval of menstruation period bleeding in primary dysmenorrheal. 2009.
26.Rahnama P, Montazeri A, Huseini HF, Kianbakht S, Naseri M. Effect of Zingiber officinale R. rhizomes (ginger) on pain relief in primary dysmenorrhea: a placebo randomized trial. BMC complementary and alternative medicine. 2012;12(1):92.
27.Safari A, Shah RG, Damavandi A. Comparison of The effects of Vitamin E and Mefenamic Acid on the severity of primary dysmenorrhea. 2006.