

# NIH Public Access

Author Manuscript

*Comput Methods Programs Biomed.* Author manuscript; available in PMC 2008 February 1.

Published in final edited form as: *Comput Methods Programs Biomed*. 2007 February ; 85(2): 109–114.

# Use of OSWALD for analyzing longitudinal data with informative dropout

Amy E. Begley,  $MA^{(1)}$ , Gong Tang,  $PhD^{(2)}$ , Sati Mazumdar,  $PhD^{(2)}$ , Patricia R. Houck, MSH <sup>(1)</sup>, John Scott,  $AM^{(1),(2)}$ , Benoit H. Mulsant,  $MD^{(1)}$ , and Charles F. Reynolds III,  $MD^{(1)}$  (1)Western Psychiatric Institute and Clinic, Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15213

(2) Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, 15213

# Abstract

OSWALD (Object-oriented Software for the Analysis of Longitudinal Data) is flexible and powerful software written for S-PLUS for the analysis of longitudinal data with dropout for which there is little other software available in the public domain. The implementation of OSWALD is described through analysis of a psychiatric clinical trial that compares antidepressant effects in an elderly depressed sample and a simulation study. In the simulation study, three different dropout mechanisms: completely random dropout (CRD), random dropout (RD) and informative dropout (ID), are considered and the results from using OSWALD are compared across mechanisms. The parameter estimates for ID-simulated data show less bias with OSWALD under the ID missing data assumption than under the CRD or RD assumptions. Under an ID mechanism, OSWALD does not provide standard error estimates the usage of OSWALD for analyzing longitudinal data with dropouts and how to draw appropriate conclusions based on the analyzing longitudinal data with dropouts regarding the dropout mechanism.

#### Keywords

OSWALD; non-ignorable missing; longitudinal data analyses; simulation

# 1. Introduction

Longitudinal studies follow the same group of individuals over a period of time. The response variable for each subject in the study is observed on several occasions and the repeated responses from each individual are correlated. Dropouts are common in longitudinal studies and care must be taken when analyzing the data. For example, in a psychiatric study a depression score might be followed over time. If patients whose depression were not relieved even with some treatment tended to dropout, we would observe a benefit of the treatment over time which would be reflected by a decline of the average depression score among patients who remained with the study. Three dropout processes are generally considered in the

Corresponding Author: Amy E. Begley, WPIC (E1109), 3811 O'Hara Street, Pittsburgh, PA, 15213, telephone: 412-246-6442, fax: 412-246-5300, email: begleyae@upmc.edu

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

This paper is from the Advanced Center for Interventions and Services Research for Late Life Mood Disorders; and the John A. Hartford Center of Excellence for Geriatric Psychiatry.

literature: completely random dropout (CRD), random dropout (RD), and informative dropout (ID)[1,2]. In CRD, dropout is independent of both observed and missing data. For example, an individual may drop out of a study because of relocation, change of insurance policy, or car accident. In RD, dropout depends on observed data but not on missing data. RD might occur in a longitudinal psychiatric study if patients who did not observe improvement in the past recorded depression scores thought that the treatment was not helpful and tended to discontinue participation. It would also occur when patients who observed greatly improved depression scores over time thought it would be not necessary to stay in the trial any more. In ID, dropout may depend on both observed and unobserved data. For example, if individuals with deteriorating physical health, often not quantified and recorded in practice, and related to the depression level, tended to drop out.

For longitudinal data with dropouts, a statistical model is usually considered for the dropout process in addition to the model for the repeatedly measured outcomes, or the complete data model, where maximum likelihood method can be used to perform inference [1,2,3,4,5]. When the dropout process is RD and the parameters for the dropout process are distinct from the parameters for the complete data model, the dropout process is called ignorable for the maximum likelihood method and statistical analysis can be based solely on the observed values as if there were no missing values [1,2]. The widely used statistical software SAS Proc Mixed [6] assumes this process; however, estimates of standard errors for fixed effect parameters are usually underestimated and this leads to smaller p-values when the dropout process is not CRD [7]. When the dropout process is ID, a model for the dropout process needs to be specified, usually as a function of the covariates, observed responses, unobserved responses, or even some random effects from the complete data model [2,3,4,5]. Logistic or probit regression is often applied on an ID process. The likelihood function often involves intractable integration and numeric approximations are required.

Besides the maximum likelihood method, inverse-probability weighted estimating equations (IPWEE) weighs each individual with complete records by the inverse of the estimated inclusion probability as if such an individual represents him- or herself and also a few other individuals who had similar characteristics but dropped out earlier by chance [8,9]. However, IPWEE is not easy to implement in practice, especially when dropouts occurred at several occasions. Other statistical methods include conditional likelihood method and pseudo-likelihood method that focus on the inference of certain components of the distribution of the repeated-measured outcomes and are more ad-hoc. [10,11]

OSWALD (Object-oriented Software for the Analysis of Longitudinal Data) is a flexible and powerful object-oriented software package developed by the Statistics Group at the University of Lancaster for use with S-PLUS for the analysis of longitudinal data with dropout. [12,13] In addition to CRD and RD mechanisms, OSWALD can model longitudinal data with ID when the reason for dropout depends on the underlying unobserved value of the outcome measured and draws inference under the likelihood framework.

Currently there is no software in the public domain other than OSWALD that can analyze longitudinal data with potentially informative dropouts. Here we illustrate how to implement OSWALD in practice via analysis of a dataset from a psychiatric clinical trial and a simulation study. The dataset was first analyzed under the ID assumption and the parameter estimates were then used to generate CRD, RD and ID datasets in order to illustrate the utility of OSWALD. We applied OSWALD and compared the results assuming the three different dropout mechanisms for each dataset. We also applied OSWALD to the original clinical trial under all three dropout mechanisms and compared the results. Since OSWALD does not provide the standard errors for the parameter estimates under informative dropout assumption, a bootstrap procedure was implemented to derive them.

Our aims are to emphasize how various assumptions about the dropout process affect the results of analysis and to provide an illustration of how OSWALD is able to analyze longitudinal data under CRD, RD and ID assumptions about the dropout process. We also demonstrate an addition to the original program to derive the standard errors under the ID assumption in order to make inference on parameters for the complete data model.

#### 2. Method

In OSWALD, a linear mixed-effect model is implemented for the repeatedly measured response and a logistic regression for the dropout process. Suppose a response variable for each subject is measured at *K* timepoints:  $y_1, y_2, ..., y_K$ . For longitudinal studies in which the response variable is a continuous measurement, a popular analytic approach is a linear mixed-effect model that allows for the individual variability to be modeled by including random intercept and a serial correlation term:

$$\mathbf{Y}_{ij} = \boldsymbol{\mu}_{ij} + \mathbf{U}_i + \mathbf{W}_i (\mathbf{t}_{ij}) + \mathbf{Z}_{ij}$$

where  $Y_{ij}$  represents the jth response from the ith subject measured at time  $t_{ij}$ ;  $\mu_{ij}$  is the expected value of  $Y_{ij}$  where the relationship between covariates, such as treatment, time, and the repeatedly measured outcome variable is modeled;  $U_i \sim N(0, v^2)$  is the random intercept shared by the responses from subject i;  $W_i(t_{ij}) \sim N(0, \sigma^2)$  is the realization from the stationary Gaussian process that reflects the serial correlation among the responses from subject i; and  $Z_{ij} \sim N(0, \tau^2)$  is the measurement error.

OSWALD is flexible enough to model the variance components in such a linear mixed model. These variance components are 1) a random intercept common for repeated-measures from the same subject, with variance  $v^2$ ; 2) serially correlated component with variance  $\sigma^2$  and autocorrelation function  $\rho(\cdot)$  i, and 3) a measurement error with variance  $\tau^2$ .

Without loss of generality, we assume that  $y_1$  is observed for all individuals. At each of the subsequent time points, individuals are subject to dropout for various reasons such as relocation and treatment failure. We can write  $\{y_1, y_2, ..., y_K\} = \{y_{obs}, y_{mis}\}$ , where  $y_{obs}$  is the observed part and  $y_{mis}$  is the missing part after dropout. At each time point, OSWALD uses a logistic regression to model the dropout:

$$\text{logit}\Big[P_k\Big(y_1, y_2, \dots, y_{k-1}, y_k\Big)\Big] = \theta_0 + \theta_1 y_k + \sum_{j=2}^{q} \theta_j y_{k+1-j}$$

where  $P_k$ = conditional probability of a subject dropping out at time  $t_k$  given the previously recorded history of  $y_1, y_2, ..., y_{k-1}$  and underlying unobserved value of  $y_k; \theta = (\theta_0, \theta_1, ..., \theta_q)$  is a vector of (q+1) parameters relating the dropout process to the unobserved ( $y_k$ ) and previously observed responses [3].

When  $\theta_j = 0$  for all  $1 \le j \le q$ , the dropout mechanism is CRD; when  $\theta_1 = 0$  and  $\theta_j$  may be nonzero for  $2 \le j \le q$ , the mechanism is RD; when  $\theta_1 \ne 0$ , the mechanism is ID since the dropout probability depends on the underlying value of  $y_k$ .

With the dropout process modeled, OSWALD draws inference based on the likelihood function [3]. In S-Plus, the primary OSWALD function to analyze longitudinal data with dropout under various assumptions (CRD, RD or ID) is *pcmid* that takes the general form

pcmid(formula, vparms, drop.parms=NULL, drop.cov.parms, dropmodel),

where:

• *formula* refers to the model that predicts outcome with covariates such as treatment group and time

```
Informative.Dropout <-
```

function(NTPAXIL.balanced)

{ NTPAXIL.ldf <-as.ldframe(NTPAXIL.balanced)</pre>

```
results <-pcmid(formula = y ~ as.factor(group) * (Time),
vparms = c(10, 7, 0.4),
```

data=NTPAXIL.ldf, drop.parms = c(-2, -2),

dropout.est = T, correxp=1)

- *vparms* are the initial estimates of the variance parameters ( $v^2$ ,  $\tau^2$ ,  $\phi$ )
- *drop.parms* are the initial values for the dropout parameters. The first value is the regressor on the unobserved dropout value: if this value is non-zero, this is an ID model; if it is zero, this is a RD model. Further values are regressors on previous observed values.
- *drop.cov.parms* are the initial estimates for the covariate parameters in the dropout model.
- dropmodel is a model (like formula) describing the covariate part of the dropout model. If not supplied but drop.cov.parms is supplied, the constant model ~1 is assumed [4].

The ID model does not provide standard error estimates. We implemented a bootstrap method to obtain these. This was done by using the internal S-PLUS function "*bootstrap*" on a wrapper function containing a call to a OSWALD's *pcmid*, with dropout parameters corresponding to ID [13].

The standard *pcmid* output gives a summary of the model, the maximized log-likelihood, estimates and the corresponding standard errors of the parameters for the mean structure, estimates of the variance components and dropout process parameters. The bootstrap output gives the standard errors of the parameter estimates for the mean structure when an ID process is assumed in the application of *pcmid*.

A randomized clinical trial compared the antidepressant effects of two medications [14]. There were 116 subjects randomized to either nortriptyline (N=54) or paroxetine (N=62) and followed for 12 weeks. The aim of the study was to compare treatment and side effects across time.

Depression symptoms were measured weekly with the Hamilton 17-item Depression Scale (HRS-D). Higher HRS-D scores indicate more severe depressive symptoms [15]. Figure 1 shows the bi-weekly HRS-D profiles for the two treatment groups. There was approximately 52% dropout by week 12. The sample sizes for each group are shown on the graph.

We applied OSWALD to the original randomized clinical trial data. Under the ID assumption, OSWALD does not provide standard error estimates, so we implemented a bootstrap method. The function and commands to perform the bootstrap to get estimates of the standard error for the ID analyses were:

```
return(as.vector(results$coefficients$mean))
```

}

```
boot.result <-bootstrap(NTPAXIL.balanced,
Informative.Dropout, B=500, block.size=500)
```

In the simulation study, datasets were generated under each of the 3 types of dropout mechanism using a function the authors wrote for S-PLUS. In the simulated data the parameters for intercept and time came from the analysis of the example described earlier. However, we added a significant group effect and removed the interaction term for ease of illustration. The measurement error variance and the random variance terms were also increased in the simulated data. Only the bi-weekly HRSD scores were simulated.

We used the OSWALD *pcmid* function mentioned earlier to run the analyses. For CRD, RD and ID we used the *vparms* set at 2, 1, and 0.32 respectively. The *drop.parms* was 0 for all CRD runs. For RD, the *drop.parms* was c(0, .25) and for ID it was c(.95, .25). The RD numbers were chosen to generate approximately 50% dropout by week 12 as seen in the clinical trial data.

### 3. Results

While applying OSWALD to the original randomized clinical trial data, the initial estimates for the variance components in the model were estimated from the empirical variogram (see Figure 2). The parameter estimates for complete data model, based on the CRD, RD and ID assumptions on the dropout process, are shown in Table 1. For the analysis under ID assumption on the dropout, 200 bootstrap samples were drawn in order to calculate the standard errors. There is a significant time effect for all models but no significant group or group by time interaction effects. The parameter estimates and standard errors for CRD and RD are similar, and ID differs slightly. Based on these observations, assumptions on the dropout process do not seem to have much impact on the estimates of parameters for the repeated-measurement process in this dataset.

The simulation results are presented in Tables 2A, 2B and 2C where data were simulated under CRD, RD and ID dropout processes, respectively. Under each dropout process setting, 500 datasets were simulated and OSWALD was applied on these datasets with CRD, RD or ID assumptions. The parameter estimates, empirical standard deviations and coverages of the 95% confidence intervals are presented. The empirical standard deviations in the parentheses were calculated based on the corresponding parameter estimates on 500 simulated datasets. In table 2C, the datasets were simulated under an ID dropout process. In order to calculate the coverage of the 95% CI for the OSWALD analysis under an ID assumption, 50 instead of 200 bootstrap samples were drawn in order to reduce computation time. The simulation results suggest that if data were generated from an ID process but analyzed under a CRD or RD assumption, the estimates would have larger bias and the 95% confidence intervals could have worse coverage than those from analysis by OSWALD under an ID assumption.

## 4. Discussion

Although missing data is a well recognized problem in practice and there are theoretical techniques to approach it, there are few readily available software programs. OSWALD supplies a powerful tool for the analysis of longitudinal data with dropout under a few prevalent

and commonly assumed mechanisms: CRD, RD and ID. The simulation study demonstrates the effectiveness of OSWALD in this respect. When data were simulated under an ID mechanism, analyses with OSWALD under an ID assumption have less bias and better coverage of the 95% CI than analyses with OSWALD under a CRD or RD assumption on the dropout process. Though in general, analysis under CRD should have the same parameter estimates and smaller standard errors compared with analysis under RD [7], we did not see much difference in standard errors between CRD and RD. We suspect that this is due to our selected simulation parameters that came from estimates on a psychiatric clinical trial and

Although OSWALD is freely available, it has a few drawbacks. Initial estimates are often difficult to identify and great care needs to be taken that the models converge to the true maximum likelihood, although this is a problem common to most high dimensional optimization problems. Another drawback is that OSWALD is no longer being maintained or updated by its developers to handle newer S-PLUS versions. New user-friendly programs are needed to handle the problem of analyzing the longitudinal data with informative dropout. In the meantime, OSWALD is a reasonable and effective option.

OSWALD works with S-PLUS for both Windows and Unix Platforms (S-PLUS versions 2.2-2000 for Windows and 3.4-4.4 for Unix). It is written using source code in the S [4] and C programming languages. OSWALD for Windows or Unix can be downloaded from http://www.maths.lancs.ac.uk/Software/OSWALD. OSWALD version 3.1, 3.2 and 3.4 for S-PLUS for Windows and OSWALD version 4.4 for Unix are available from this site.

#### Acknowledgements

 $Supported in part by P30 \,MH52247, P30 \,MH71944, P30 \,MH30915, R25 \,MH60473, AG020677, T32 \,MH073451 \,and \,GlaxoSmithKline.$ 

#### References

- [1]. Little, RJA.; Rubin, DB. Statistical Analysis with Missing Data. 2nd Ed.. Wiley; New York: 2002.
- [2]. Little RJA. Modeling the drop-out mechanism in repeated-measures studies. Journal of the American Statistical Association 1995;90:1112–1121.
- [3]. Diggle P, Kenward MG. Informative drop-out in longitudinal data analysis. Applied Statistics 1994;43:49–93.
- [4]. Wu MC, Carroll RJ. Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. Biometrics 1987;44:175–188.
- [5]. Diggle, PJ. Dealing with missing values in longitudinal studies. In: Everitt, BS.; Dunn, G., editors. Recent Advances in the Statistical Analysis of Medical Data. Arnold; London: 1998. p. 203-228.
- [6]. SAS/STAT Software Version 8. SAS Institue; Cary, NC.: 2000.

imposed large variation in the simulated datasets.

- [7]. Kenward MG, Molenberghs G. Likelihood based frequentist inference when data are missing at random. Statistical Science 1998;13:236–247.
- [8]. Robins JM, Rotnitzky A, Zhao LP. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. Journal of American Statistical Association 1995;90:106– 121.
- [9]. Rotnitzky A, Robins JM, Scharfstein DO. Semiparametric regression for repeated outcomes with nonignorable nonresponse. Journal of the American Statistical Association 1998;93:1321–1339.
- [10]. Wu MC, Beiley KR. Estimation and comparison of changes in the presence of informative right censoring: conditional linear model. Biometrics 1989;45:939–955. [PubMed: 2486189]
- [11]. Troxel AB, Lipsitz SR, Harrington DP. Marginal models for the analysis of longitudinal measurements with nonignorable non-monotone missing data. Biometrika 1998;94:1096–1120.

- [12]. Smith, DM.; Robertson, B.; Diggle, J. Department of Mathematics and Statistics, University of Lancaster, LA1 4YF; United Kingdom: 1996. OSWALD: Object-oriented software for the analysis of longitudinal data in S. Technical Report MA 96/192
- [13]. S-PLUS 2000 User's Guide. Mathsoft; Seattle, Washington: 1999.
- [14]. Mulsant BH, Pollock BG, Nebes R, Miller MD, Sweet RA, Stack J, Houck PR, Bensasi S, Mazumdar S, Reynolds CF III. A twelve-week, double-blind, randomized comparison of nortriptyline and paroxetine in older depressed inpatients and outpatients. American Journal of Geriatric Psychiatry 2001;9:406–414. [PubMed: 11739067]
- [15]. Hamilton M. A rating scale for depression. Journal of Neurology, Neurosurgery & Psychiatry 1960;23:56–62.



#### Figure 1.

Mean Hamilton 17-item Depression Scale Profiles by Treatment Group. Mean bi-weekly HRS-D scores for patients randomized to either notriptyline or paroxetine.



Estimated process variance: 20.518

Variogram:

LAG	1	2	3	4	5	6
VGRAM	7.957	9.432	9.320	13.743	13.146	10.468

#### Figure 2.

Initial Estimates of variances based on the variogram. Variogram of the HRS-D residuals. If the variogram levels off very quickly then  $\phi$  is large; if it does not level off until very large lags then  $\phi$  is small.

	Model: C	RD	Model: F	Ð	Model: I	D
Itercept	19.721	(0.676)	19.721	(0.676)	18.595	(0.739)
roup	-0.003	(0.924)	-0.004	(0.923)	-0.003	(0.855)
ime	-1.368	(0.101)	-1.368	(0.101)	-1.471	(0.139)
roup*Week	0.098	(0.136)	0.097	(0.136)	0.140	(0.129)

**NIH-PA Author Manuscript** 

1 June 1

**NIH-PA** Author Manuscript

Author Manuscript	ets, if available) from the Simulation.
NIH-PA Author Manuscript	Table 2           parentheses) and coverage of 95% CI (in the brack
<b>NIH-PA</b> Author Manuscript	Parameter estimates, standard errors (in the

Model: ID

Model: RD

Model: CRD

A. CRD simulated data<sup>*a*</sup>

Eived Effects							
Intercept Group	19.753 2.969 1.260	(0.465) [94.0] (0.602) [92.2]	19.754 2.969	(0.465) [95.0] (0.602) [92.8]	19.812 2.974 1.240	(0.471) (0.609) (0.651)	
ume Variance Estimates	-1.209	(0.048) [90.4]	600.1-	(0.048) [94.0]	-1.349	(100.0)	
Nu Square	1.946	(1.927)	1.758	(1.836)	1.740	(1.901)	
I au Square Sigma Square	27.940	(1.90) (2.748)	28.224	(1.096) (2.735)	28.482	(2.751) (2.751)	
Phi Dronout Parameters	0.325	(0.0/3)	0.317	(0.055)	0.319	(0.058)	
Intercept	I	1	-2.635	(0.198)	-2.820	(0.207)	
YD - 1		1 1	0.019	- (0.012)	0.067	(0.048) (0.041)	
B. RD simulated data <sup>d</sup>							
	W	odel: CRD	M	lodel: RD	Mod	lel: ID	
Fixed Effects					102.01	0 400	
Intercept Group	2.984	[0.601] $[94.0]$	2.981	[0.67] [0.600] [03.8] [0.6600] [03.8]	3.010	(0.498) (0.619)	
Lime Variance Estimates	-1.305	[0.26] (860.0)	-1.308	[4.68] (100.0)	-1.300	(0.060)	
Nu Square	2.100	(2.399)	1.879	(2.157)	1.985	(2.408)	
1 au Square Sigma Square	1.094 27.842	(3.124)	1.008 28.150	(2.172) (3.031)	1.485 28.446	(2.188) (3.214)	
Phi Dronout Parameters	0.323	(0.087)	0.316	(0.067)	0.318	(0.071)	
Intercept	ł	1	-11.149	(0.997)	-11.590	(1.912)	
YD - 1	: :	: :	0.456	(0.043)	0.019	(0.103) (0.075)	
· ·			2				
C. ID simulated data <sup>u</sup>	Mod	el: CRD	Mode	i: RD	Model	Ð	
Fixed Effects Intercent	078.81	(0 410) [40 6]	18 877	0.110) [41.8]	10 680	533) [94 0]	
Group	2.713	(0.487) [92.8] (0.487) [92.8]	2.704	(0.490) [91.0] (0.450) [91.0]	2.965 (0		
Variance Estimates	0001-	[7:76] (ocn:n)	700.1-	[0:16] (/c0:0)	n) 700.1-		
Nu Square Tan Sunare	0.562 22.914	(1.065)	0.559	(1.060)	0.720 (1	.546)	
Sigma Square Phi	6.394 1.835	(4.994) (10.155)	12.286 2.108	(11.029) (10.393)	7.747 (6 2.109 (1	589) 1.313)	
Dropout Parameters		~					
Intercept VD	: :	1 1	-5.433 (	(0.591)	-18.705 (7 0.6801 (0	.075)	
YD - 1	1	-	0.250	(0.033)	0.317 (0	.121)	

Comput Methods Programs Biomed. Author manuscript; available in PMC 2008 February 1.

Page 11

NIH-PA Author Manuscript

C. ID simulated data"	Model: CRD	Model: RD	Model: ID
Cumulative % Missing: 0, 10.4, 2	22.4, 30.2, 39.7, 46.6, 51.7		
Cumulative % Missing: 0, 32.3, 4	11.4, 45.2, 46.9, 47.7, 48.0		
Cumulative % Missing: 0, 56.0, 6	53.1, 66.1, 67.4, 68.1, 68.5		
<sup>a</sup> Simulated Data Parameters: Inte	srcept = 19.72, Group = 3, Time =-1.37, Nu Square = 2, Tau	a Square = 1, Sigma Square=28.82, Phi=.32	
<sup>a</sup> Simulated Data Parameters: Inte	srcept = 19.72, Group = 3, Time =-1.37, Nu Square = 2, Tau	a Square = 1, Sigma Square=28.82, Phi=.32, b0 (Interce	pt) = -11, b1 (YD) = 0.45, b2 (YD - 1) = 0

a Simulated Data Parameters: Intercept = 19.72, Group = 3, Time =-1.37, Nu Square = 2, Tau Square = 1, Sigma Square=28.82, Phi=.32 b0 (Intercept) = -25.0, b1 (YD) = 0.95, b2 (YD - 1) = 0.25