Random Regression Models for Multicenter Clinical Trials Data[1]

Donald Hedeker, Ph.D., Robert D. Gibbons, Ph.D., and John M. Davis, M.D.[2]

Abstract

A random-effects regression model is proposed for the analysis of data arising from multicenter clinical trials. Advantages of the random regression model (RRM) in this context include that it allows for varying numbers of subjects within the different centers, it can assess the influence of variables measured both at the level of the subject and at the level of the center on the subject's clinical outcome, and it controls for and estimates the amount of intracenter variation that is present in the data. An example utilizing data collected in the National Institute of Mental Health schizophrenia collaborative study, where subjects were clustered within nine centers, illustrates the usefulness of the statistical model. Other applications and extensions of RRM within a psychiatric framework are discussed.

Introduction

In a series of papers (Gibbons et al. 1988; Hedeker et al. 1989), we have illustrated some of the advantages of the random regression model (RRM) in analyzing longitudinal psychiatric data: RRM allows for the presence of missing data, time-varying or invariant covariates, and subjects measured on different occasions. In a similar way, RRM has many advantages in the analysis of data arising from multiple centers, that is, clustered or hierarchical data. RRM permits a varying number of subjects within each cluster as well as the inclusion of covariates measured at both the level of the individual and the cluster.

Clustered data arise when subjects are nested or clustered within some larger unit (e.g., center, ward, clinic, rater) and the focus of the analysis is to examine relationships at the individual level while controlling for the variability at the level of the cluster. Since data from subjects within a cluster are likely to be correlated, an analysis which ignores this association would be misleading. For example, in a multicenter clinical trial our intent should be to evaluate the drug vs. placebo difference at the patient level while simultaneously accounting for the differences between the centers.

Furthermore, in some cases it may be of interest to assess the influence of center-level variables (those that vary only with center, e.g., center size and location) in addition to subject-level variables (those that vary for each subject within each center, e.g., treatment, sex, and age) on the subject's outcome. Analysis by RRM can assess the effect of variables at either the level of the center or the subject on an outcome variable measured at the subject level. In addition to providing tests of significance for the terms in the model, RRM can also estimate the unique effect for each center as well as the overall variability of these center effects. These additional statistics can be useful in determining the intracenter correlation, as well as in identifying centers that are having an undue influence on the data.

Model Description

Consider the following regression model for assessing the effects of the center-level variable x and the subject-level variable z on the measurement y from subject i within center j:

$$y_{ij} = \mu_\alpha + \alpha_j + \beta x_j + \gamma z_{ij} + \epsilon_{ij} \qquad (1)$$

$j = 0, 1, \ldots, N$ centers

$i = 1, 2, \ldots, n_j$ subjects in center j

where

y_{ij} is the measurement for subject i within center j

μ_α is the overall population grand mean

α_j is the effect due to center j

[1] This work was supported in part by a grant from the Office of Naval Research.

[2] University of Illinois at Chicago and Illinois State Psychiatric Institute, Chicago, IL

Reprint requests: Dr. Donald Hedeker, Biometric Laboratory, Illinois State Psychiatric Institute, #531W, 1601 West Taylor Street, Chicago, IL 60612.

β is the regression coefficient for x_j
x_j is a center-level explanatory variable
γ is the regression coefficient for z_{ij}
z_{ij} is a subject-level explanatory variable
ϵ_{ij} is an independent residual distributed normally with mean 0 and variance σ^2.
We also assume that the distribution of the individual center effects (α_j) is normal $N(\mu_\alpha, \sigma^2_\alpha)$ in the population.

In terms of the influence of the center, this model represents the measurement of y as a function of center at both the individual (α_j) and population (μ_α) levels. In addition, this model can assess the residual variance σ^2 and the variance due to centers σ^2_α, from which the intracenter correlation can be calculated as $\sigma^2_\alpha/(\sigma^2_\alpha + \sigma^2)$. The intracenter correlation is the proportion of variance in the data that is attributable to the center.

As an example of RRM with both center-level and subject-level covariates, consider the following illustration. A Hamilton Rating Scale for Depression (HAM-D; Hamilton 1960) score is measured for each subject at baseline and at the end of 4 weeks on subjects assigned either to placebo or drug therapy (tx = 0 for placebo and tx = 1 for drug). Furthermore, for the N centers the number of subjects in the center (size) is thought to influence a subject's improvement. Then, for a given center j the above model for the change in HAM-D scores between baseline and the end of the study (Δ HAM-D) would be represented in matrix form by the following:

$$\begin{bmatrix} \Delta\text{HAM-D}_{j1} \\ \Delta\text{HAM-D}_{j2} \\ \cdots \\ \Delta\text{HAM-D}_{jn_j} \end{bmatrix} = \mu_\alpha \begin{bmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{bmatrix} + \alpha_j \begin{bmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{bmatrix} + \begin{bmatrix} \text{size}_j & tx_{j1} \\ \text{size}_j & tx_{j2} \\ \cdots & \cdots \\ \text{size}_j & tx_{jn_j} \end{bmatrix} \begin{bmatrix} \beta \\ \gamma \end{bmatrix} + \begin{bmatrix} \varepsilon_{j1} \\ \varepsilon_{j2} \\ \cdots \\ \varepsilon_{jn_j} \end{bmatrix}$$

Notice, the number of subjects within a given center n_j is not assumed to be equal between centers. Also, although this model only contains two covariates (size and tx), more explanatory variables can be included in the model in the same manner as in a multiple regression model. In fact, the above model differs from the "usual" multiple regression model only because the effect of center (α_j) is regarded as a random, and not a fixed, effect.

From a statistical perspective, the randomness of the center effect implies interest in characterizing the population of centers from which the current sample of centers was drawn. By positing a distribution for the center effects in the population (i.e., $N\{\mu_\alpha, \sigma^2_\alpha\}$), the estimation of the fixed regression coefficients (β and γ), and the structural (σ^2) and population (μ_α, σ^2_α) parameters can be accomplished through the method of maximum likelihood, whereas the estimation of the individual center effects (α_j) can be accomplished using empirical Bayesian methods. Details on the estimation can be found in Hedeker (1989) or Bock (1989).

Example

The following data were collected as part of the National Institute of Mental Health schizophrenia collaborative study on treatment-related changes in overall severity using the Inpatient Multidimensional Psychiatric Scale (IMPS; Lorr & Klett 1966). Item 79, "Severity of Illness," was scored in the following way:

1 = normal, not at all ill
2 = borderline mentally ill
3 = mildly ill
4 = moderately ill
5 = markedly ill
6 = severely ill
7 = among the most extremely ill

The sample sizes for the 434 subjects within the nine centers broken down by treatment are given in Table 1. The number of subjects within each center varies greatly, a feature of the data which can be easily accommodated by RRM, since the assumption of equal cell sizes is unnecessary.

Previous analysis of these data (Gibbons et al. 1988) revealed similar effects for the three drug

TABLE 1. Sample Sizes by Center and Treatment Group.

Center	Placebo	Chlorpromazine	Fluphenazine	Thioridazine	Total
1	13	9	8	8	38
2	20	22	25	23	90
3	13	8	7	10	38
4	15	18	18	17	68
5	13	15	13	13	54
6	7	9	8	7	31
7	10	10	11	10	41
8	10	12	15	10	47
9	6	7	7	7	27

groups so they were combined in the present analysis. Subjects were rated on the IMPS79 at baseline and weekly up to 6 weeks. For each subject, an IMPS79 change score was calculated as the difference between the final and baseline IMPS79 values. Thus a negative IMPS79 change score would indicate an improvement from baseline. The change score mean for the 107 placebo subjects is −0.693 (SD = 1.450) and −2.244 (SD = 1.603) for the 327 subjects receiving antipsychotics.

The model specified in the equation (1) was first fit to these data. This model had no center-level covariates and only included a dummy-coded term for the drug group (0 = placebo and 1 = antipsychotic), a subject-level covariate. The results for this model are given in Table 2. The highly significant drug effect indicates the increased improvement in the IMPS79 change scores that is observed for the group receiving antipsychotics relative to the placebo group. Although the estimate of the center variance (σ^2_α) is not statistically significant, the intracenter correlation, which is the proportion of total variance that is attributable to the center, equals 0.026 (or in percentage terms, 2.6%). This value for the intracenter correlation, though small, is not negligible and is in the expected range, since as Jacobs and colleagues (1989) note, estimates of this correlation are typically 5 to 12 percent for data from spouse pairs and 0.05 to 0.85 percent for data clustered by counties.

TABLE 2. Random Regression Model Results—Simple Model.

Parameter	Estimate	SE	p
Grand mean μ_α	−0.665	0.172	.0001
Center variance $\gamma\sigma^2_\alpha$	0.062	0.054	.24
Error variance σ^2	2.379	0.163	.0001
Drug effect γ	−1.561	0.172	.0001

Adding terms to the model for the sex, age, and IMPS79 baseline score of the subject (subject-level covariates) and the sample size of each center (a center-level covariate), both as main effects and as interactions with drug, yields the results presented in Table 3. The significance of all of the interactions indicates that the drug effect does vary by the sex, age, and baseline severity of the subject and by the sample size of the center.

Since the coefficient of the sex-by-drug term is positive, it indicates a less pronounced effect of the antipsychotic medication, relative to placebo, for the males (who are coded = 1 on the dummy-coded sex variable, whereas females are coded = 0). The cell means of the IMPS79 change scores support this conclusion: for the females, a mean of −2.47 (SD = 1.57) was observed for the 170 subjects receiving antipsychotics and −0.52 (SD = 1.48) for the 58 placebo subjects. For the males, the means were −2.00 (SD = 1.61) for the 157 subjects treated with antipsychotics and −0.91 (SD = 1.40) for the 49 placebo subjects.

TABLE 3. Random Regression Model Results—Full Model.

Parameter	Estimate	SE	p
Grand mean μ_α	0.695	1.077	.52
Center variance σ^2_α	0.028	0.032	.39
Error variance σ^2	1.941	0.133	.0001
Main Effects			
Drug γ_1	2.479	1.231	.04
Sex γ_2	−0.395	0.273	.15
Age γ_3	0.012	0.021	.56
Baseline IMPS79 γ_4	−0.371	0.167	.03
Center size β	0.008	0.007	.25
Interactions			
Sex by drug γ_5	0.652	0.315	.04
Age by drug γ_6	−0.051	0.024	.03
Baseline IMPS79 by drug γ_7	−0.383	0.188	.04
Center size by drug γ_8	−0.015	0.008	.04

NOTE: IMPS79 = Inpatient Multidimensional Psychiatric Scale, Item 79.

Conversely, since the coefficients for both the age-by-drug and baseline IMPS79-by-drug terms are negative, more pronounced relationships are indicated between higher levels of these covariates and improvement of IMPS79 scores (negative change scores) for the subjects treated with the antipsychotic drugs, relative to placebo. The observed correlations between these covariates and the change in IMPS79 scores separated for the drug- and placebo-treated subjects illustrate this interpretation. The correlation between the IMPS79 change and baseline scores was -0.42 ($p < .001$) for the 327 subjects treated with antipsychotics and -0.20 ($p < .04$) for the 107 placebo subjects. Likewise, the correlation between the IMPS79 change score and age was -0.16 ($p < .003$) in subjects receiving antipsychotics and 0.08 ($p < .4$) in the subjects given placebo. Increasing age and baseline IMPS79 scores were more strongly associated with improvement in the IMPS79 change scores for the antipsychotic drug group, relative to the placebo group.

Finally, the significant negative coefficient for the center size-by-drug interaction suggests that the antipsychotic drug effect was more pronounced in the larger centers. The cell means given in Table 4, and ordered by center size, show that this effect is primarily a result of the three largest centers having the greatest drug improvement, relative to placebo.

TABLE 4. Change in IMPS79 Score Means Broken Down by Center and Treatment Type.

Center		Antipsychotic			Placebo			Difference in means
ID	Size	*n*	*Mean*	*SD*	*n*	*Mean*	*SD*	
2	90	70	−2.68	1.96	20	−0.75	1.31	−1.93
4	68	53	−2.37	1.46	15	−0.33	1.53	−2.04
5	54	41	−1.81	1.17	13	0.18	0.94	−1.99
8	47	37	−1.78	1.22	10	−0.45	1.23	−1.33
7	41	31	−2.31	1.51	10	−1.15	1.38	−1.16
3	38	25	−2.69	1.73	13	−1.31	1.84	−1.38
1	38	25	−2.12	1.60	13	−1.15	1.69	−0.97
6	31	24	−1.98	1.78	7	−0.64	1.25	−1.34
9	27	21	−1.95	1.39	6	−0.72	1.38	−1.23

Discussion

In previous work (Gibbons et al. 1988; Hedeker et al. 1989), we have highlighted some of the advantages of the random regression approach to the analysis of longitudinal psychiatric data. In a similar manner, RRM is useful in the analysis of multicenter clinical trials data. For multicenter data, RRM allows the testing of relationships at the level of the subject, while controlling for the variability due to the center. Furthermore, RRM can represent a subject-level dependent variable in terms of both subject-level and center-level covariates, as well as the interactions from these two types of covariates. In the present analysis of the change in severity scores of schizophrenic patients, it was shown that the proportion of variance attributable to the center was about 2.6 percent, a value which is consistent with what has been observed with other types of clustered data (Jacobs et al. 1989). Also, the center-level variable of center size was seen to interact significantly with drug treatment group, as did the subject-level variables sex, age, and baseline severity.

The example given illustrates the use of RRM for the analysis of data from a multicenter study; in addition, this model can be useful in the analysis of other types of clustered data sets. For example, RRM can be useful in the analysis of data clustered by rater. In this case, each of N raters provide measures (perhaps a measure of psychiatric severity) on different subjects, with the raters typically measuring a different number of subjects as well. One might then be interested in assessing how much of the total variance in the measured variable is attributable to the raters, as well as the impact any rater-level covariates might have on the measures (perhaps the experience of the rater was thought to have an influence on the rating). The random regression model that has been presented would be useful for data of this type, since it provides an estimate of the variability attributable to the cluster (in this case, the rater) as well as examining the effect of cluster- or individual-level covariates on the dependent measure of the model (which is measured at the level of the individual). In a similar way, RRM can be useful in the analysis of data clustered within families. Since much of the data obtained in psychiatric research is clustered, RRM can be useful in providing researchers with a way of examining the influence of the cluster on the data.

From a statistical point of view, a distinguishing feature of the random regression approach is that the specific clusters (e.g.,centers or raters) used in the study, are considered to be a representative sample from a larger population of potential clusters, and so, the cluster is regarded as a "random-effect" in the model. Conversely, if interest is only in making inferences about the specific clusters of a data set (e.g., is there a difference between two particular centers), then the cluster would be regarded as a "fixed effect," and the usual (fixed-effects) linear regression model could be used in the data analysis. In other words, if interest is only in testing whether the mean of, say, cluster A differs from that of cluster B, a fixed-effects approach is recommended. If instead, there is interest in assessing the overall effect that any potential cluster may have on the data, and thus, in determining the degree of variability that the cluster accounts for in the data, then the random-effects approach is advised. Therefore the main advantage of RRM for clustered data is that it allows a statistical treatment of the cluster in a manner that is consistent with the way in which it is most often conceptualized; that is, the cluster was drawn from a population of potential clusters, and the 5 or 7 or 10 clusters used in a particular study are not the population itself.

Finally, it is not uncommon in psychiatric research that subjects, who are clustered within centers or raters, are also measured longitudinally. Here, repeated observations are nested within subjects who in turn are nested within the center or rater. Notice that the data presented in the example was of this type; however, to simplify the analysis, change scores were computed for each subject to remove the nesting of observations within subjects. A more thorough analysis would take into account all of the time-related observations of each subject while accounting for the differences between subjects as well as between clusters. This type of data structure would entail another level of the RRM and a solution for this type of model has been presented by Longford (1987). However, Longford's solution does not allow for autocorrelated errors from the repeated observations of each subject, and, as we have shown in an earlier article (Hedeker et al. 1989), autocorrelated errors are typical in longitudinal data. Therefore, an extension of the current RRM is underway in order to allow for subjects to be nested within clusters and also to allow for autocorrelated errors from the repeated observations of the subjects. We hope this research will provide a valuable statistical model for psychiatric researchers faced with clustered longitudinal data.

References

Bock, R.D. Measurement of human variation: A two-stage model. In: Bock, R.D., ed. *Multilevel Analysis of Educational Data.* New York: Academic Press, 1989. pp. 319–342.

Gibbons, R.D.; Waternaux, C.; Hedeker, D.; and Davis, J.M. Random regression models: A comprehensive approach to the analysis of longitudinal psychiatric data. *Psychopharmacol. Bull.* 24:438–443, 1988.

Hamilton, M. A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* 23:56–62, 1960.

Hedeker, D.; Gibbons, R.D.; Waternaux, C.; and Davis, J.M. Investigating drug plasma levels and clinical response using random regression models. *Psychopharmacol. Bull.* 25:227–231, 1989.

Hedeker, D. "Random Regression Models With Autocorrelated Errors." Unpublished Ph.D. dissertation, Department of Behavioral Sciences, University of Chicago, 1989.

Jacobs, D.R.; Jeffery, R.W.; and Hannan, P.J. Methodological issues in worksite health intervention research: II. Computation of variance in worksite data: Unit of analysis. In: Johnson, K.; LaRosa, J.H.; Scheirer, C.J.; and Wolle, J.M., eds. *Proceedings of the 1988 Methodological Issues in Worksite Research Conference.* Airlie, VA: United States Department of Health and Human Services, 1989. pp. 77–88.

Longford, N.T. A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika* 74:817–827, 1987.

Lorr, M., and Klett, C.J. *Inpatient Multidimensional Psychiatric Scale: Manual,* revised. Palo Alto, CA: Consulting Psychologists Press, 1966.