Data Analysis in Practice-Based Research

Stephen Zyzanski, PhD

Department of Family Medicine

Case Western Reserve University

School of Medicine

October 2008

Multilevel Data

- **Statistical analyses that fail to recognize the hierarchical structure of the data, or the dependence among observations within the same clinician, yield inflated Type I errors in testing the effects of interventions.**

Multilevel Data

Inflation of the Type I error rate implies that interventions effects are more likely to be claimed than actually exist.

Unless ICC is accounted for in the analysis, the Type I error rate will be inflated, often substantially.

Multilevel Data

When ICC>0, this violates the assumption of independence. Usual analysis methods are not appropriate for group-randomized trials.

Application of usual methods of analysis will result in a standard error that is too small and a p-value that overstates the significance of the results

Table 1.1 The inflation of the alpha level of 0.05 in the presence of intra-class correlation (Barcikowski, 1981, p. 270)

	ρ		
N_i	0.01	0.05	0.20
10	0.06	0.11	0.28
25	0.08	0.19	0.46
50	0.11	0.30	0.59
100	0.17	0.43	0.70

The values in the body of the table are the observed alpha levels; N_i denotes the number of observations within a group

Traditional Response to Nesting

- **Ignore nesting or groups**
- **Conduct analysis with aggregated data**
  - **Use clinician as the unit of analysis**
- **Spread group data across lower units**
  - **Patients of a given clinician get the same value for clinician level variables**

Analysis of Aggregated Data

- **Analyses of aggregated data at higher levels of a hierarchy can produce different results from analyses at the individual level.**
- **Sample size will become very small and statistical power is substantially reduced**
- **Aggregation bias (meaning changed after aggregation)**

Miscalculation of Standard Errors

- **Nested data violate assumptions about independence of observations**
- **Exaggerated degrees of freedom for group data (e.g., clinicians) when spread across lower units (patients)**
- **Increased likelihood of Type I error due to unrealistically small confidence intervals**

Reduction in Standard Error

Basic formula for standard error of a mean is:

$$\text{Standard Error} = \frac{\text{Standard Deviation}}{\text{Sq. Rt. Sample Size}}$$

If data are for 100 clinicians spread across 1000 patients, the standard error for clinician variables will be too small (roughly 1/3 its actual size in this example)

Example of Two-Group Analysis

The primary aim of many trials is to compare two groups of patients with respect to their mean values on a quantitative outcome variable

Example of Two-Group Analysis

Testing mean differences for statistical significance, in group trials, requires the computation of standard errors that take into account randomization by groups.

Analysis example

Assume we have 32 clinicians, 16 randomized to Intervention and 16 to Control conditions

Intervention is a weight loss program and the outcome is BMI at 2 years.

Mean (I) = 25.62; Mean (C) = 25.98

Sample (I) = 1929; Sample (C) = 2205 (4134)

Standard t-test

$$t = \frac{M_1 - M_2}{\text{Sq. Rt. } (\text{Var } (1/N_1 + 1/N_2))}$$

$$= \frac{25.62 - 25.98}{0.152} = \frac{0.36}{0.152} = -2.37 \quad (p = 0.02) \quad (df = 4132)$$

P=0.02 is too small when ICC>0

Adjusted two-sample t-test

$$t = \frac{M_1 - M_2}{\text{Sq. Rt. } (\text{Var } (C_1/N_1 + C_2/N_2))}$$

ICC = 0.02; C_1=VIF/Grp1 = $(1 + (N_1-1)p)$

$$= \frac{25.62 - 25.98}{0.28} = \frac{0.36}{0.28} = -1.27 \quad (p = 0.21)$$

(df = 30)

Post Hoc Correction for Analyses that Ignore the Group Effect.

The VIF can be used to correct the inflation in the test statistic generated by the observation-level analysis.

Test statistics such as F-and chi-square tests are corrected by dividing the test by the VIF. Test statistics such as t or z-tests are corrected by dividing the test by the square root of the VIF.

Post Hoc Correction

Correction = t/VIF; where t=2.37, and

VIF=1+(M-1)p = 1+(129-1)(.02) = 3.56

Sq. Rt. of 3.56 = 1.89

Correction: 2.37/1.89 = 1.25 (computed 1.27)

Example of Adjusting for Clustering from the DOPC Study

<u>Outcome:</u> % time physicians spent chatting with adult pts.

<u>Hypothesis:</u> No pt. gender difference in time spent chatting

Mean percent time spent with:
Male Patients: 8.2%; (N = 1203)
Female patients: 7.2%; (N = 2181)

t = 3.30, p = 0.001
The intra-class correlation for chatting was: 0.15

The VIF for males was: 2.75 and 3.70 for females
After adjusting for clustering: t = 1.89, p = 0.08

Multilevel Models

This example illustrates a method for adjusting individual level analyses for clustering based on a simple extension of the standard two-sample t-test.

We now move to a more comprehensive, but computationally more extensive, approach called Multilevel Modeling

What is Multilevel Modeling?

A general framework for investigating nested data with complex error structures

Multilevel models incorporate higher level (clinician) predictors into the analysis

Multilevel models provide a methodology for connecting the levels together, i.e., to analyze variables from different levels simultaneously, while adjusting for the various dependencies.

Multilevel Models

Combining variables from different levels into a single statistical model is a more complicated problem than estimating and correcting for design effects.

Multilevel Models

- **Multilevel models are also known as: random-effects models, mixed-effects models, variance-components models, contextual models, or hierarchical linear models**

Multilevel Models

Use of information across multiple units of analysis to improve estimation of effects.

Statistically partitioning variance and covariance components across levels

Tests for cross-level effects (moderator)

A Multilevel Approach

Specifies a patient-level model within clinicians. Level 1 model

Treats regression coefficients as random variables at the clinician level

Models the mean effect and variance in effects as a function of a clinician-level model

Correlates of Alcohol Consumption

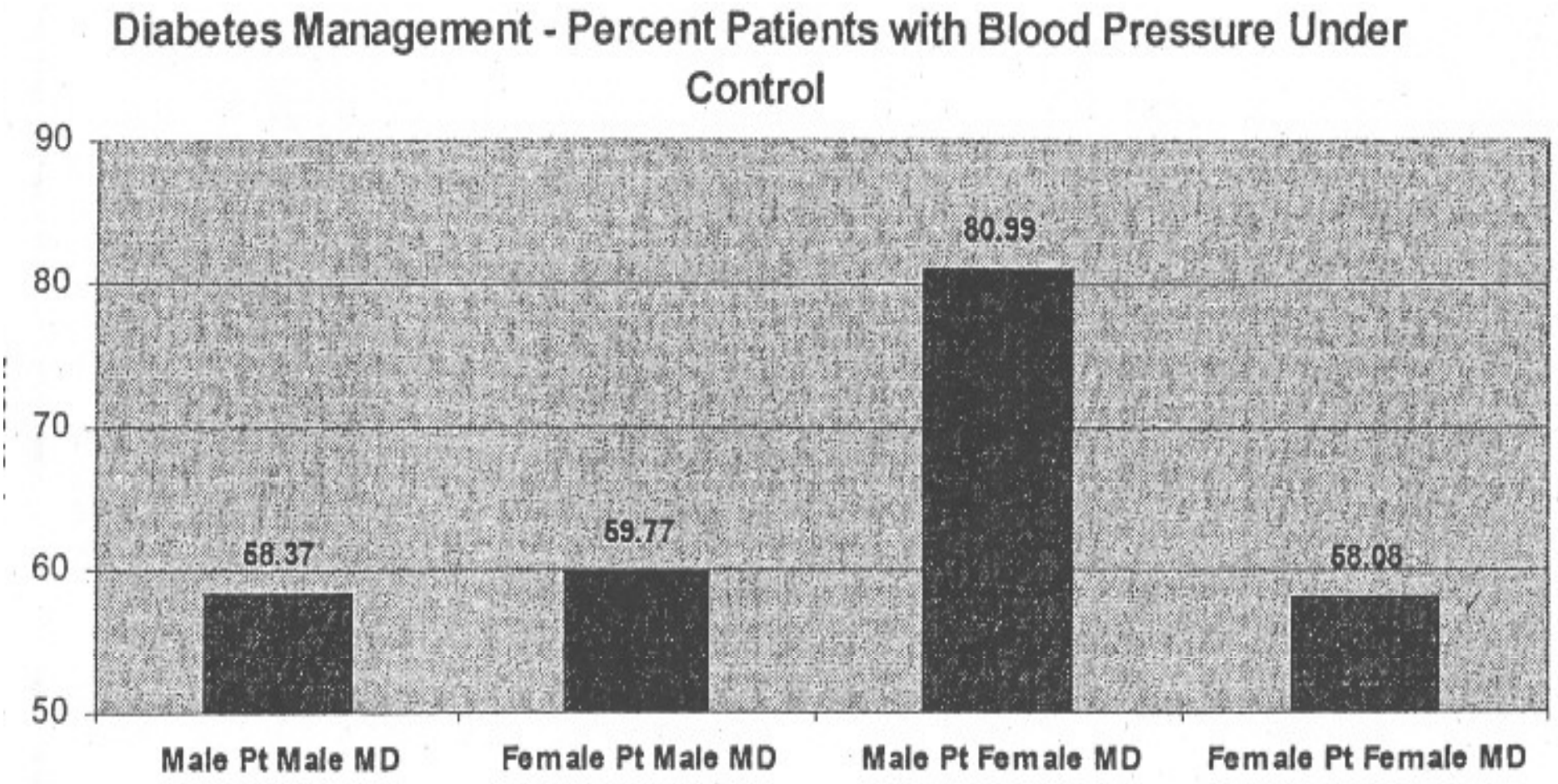
	β	S.E.	P value
Intercept	**2.06**	**0.46**	**<.001**
Individual Coefficients			
Distance to Outlet	**.0001**	**.035**	**.997**
Age	**-.008**	**.001**	**<.001**
Female	**-.678**	**.053**	**<.001**
Education	**.145**	**.034**	**.001**
Black	**-.527**	**.069**	**<.001**
Census Tract Coefficients			
Mean Distance to Outlets	**-.477**	**.194**	**.024**
Mean Age	**.014**	**.017**	**.435**
Percent Female	**.292**	**.957**	**.763**
Mean Education	**.345**	**.408**	**.410**
Percent Black	**-.407**	**.334**	**.238**
Percent Variance Explained			
Within Census Tracts	**8.9**		
Between Census Tracts	**80.3**		

ICC=11.5%

(Scribner, 2000)

Gender Differences in CV Risk Factors Management Using Multiple Levels With Interaction Analysis

Management	Patient gender	Physician gender	Patient & MD interaction
Weight management			
1. Obesity documented	**F>M p = 0.001, OR = 1.8**		
2. Physical activity advice		**F>M p = 0.032, OR = 2.21**	
Hypertension management			
1. Advice for diet/wt loss	**F>M p = 0.07, OR = 2.5**		
2. DM medication	**F<M p = 0.03, OR = 0.49**		
3. Aspirin Therapy	**F<M p = 0.0003, OR = 0.3**		
4. ACEI/ARB therapy			**P = 0.035**
5. BP <130/85			**P = 0.05**
6. Physical activity advice		**F>M p = 0.0002, OR = 6.55**	


Diabetes Management - Percent Patients with Blood Pressure Under Control
90
80
70
60
50
58.37
59.77
80.99
58.08
Male Pt Male MD
Female Pt Male MD
Male Pt Female MD
Female Pt Female MD

Software Packages

- **MBDP-V (www.ssicentral.com)**
- **VARCL (www.assess.com.VARCL)**
- **SAS Proc Mix (www.sas.com)**
- **MLwiN (www.ioe.ac.uk/mlwin)**
- **HLM (www.ssicentral.com)**

Take Home Messages

- **Clustered data inflate stand errors & p-values**
- **Standard statistical analyses are invalid**
- **Post hoc corrections for clustering**
- **Multilevel data require multilevel analyses**
- **MM designed to analyze variables from different levels simultaneously & cross-level interactions**
- **Computationally extensive, requiring expertise**
- **Parameters to be estimated increase rapidly**
- **Missing data at Level-2 more problematic**