MIMEO  James J. Grady
SERIES  STRUCTURAL COVARIANCE
#2102T  MATRICES FOR INCOMPLETE, I
LONGITUDINAL DATA

12-19-93  J. J. Lee
STRUCTURAL COVARIANCE MATRICES FOR INCOMPLETE LONGITUDINAL DATA

by

James J. Grady

Department of Biostatistics, University of North Carolina at Chapel Hill, NC.

Institute of Statistics Mimeo Series No. 2102T

July 1992
STRUCTURAL COVARIANCE MATRICES FOR
INCOMPLETE LONGITUDINAL DATA

by

James J. Grady

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the Doctor of Public Health in the Department of Biostatistics.

Chapel Hill

1992

Approved by:

Advisor

Reader

Reader

Reader
© 1992

James J. Grady

ALL RIGHTS RESERVED
ABSTRACT

JAMES JONATHAN GRADY. Structured Covariance Matrices for Incomplete Longitudinal Data. (Under the Direction of Ronald W. Helms.)

Longitudinal studies offer effective ways to study individual change over time, but they typically present problems such as missing data and data that are irregularly timed, which complicate data analysis. Missing data are handled poorly with standard multivariate analysis methods which use fully parametrized covariance matrices and which result in a proliferation of estimated parameters and essentially require case-wise deletion. Although the theory and methods for the mixed model analysis of incomplete longitudinal data exist, these methods have not been readily available to practicing statisticians, due, in part, to limited exposure and to obscure software. The objective of this paper is to show how these new analysis methods can be used to model the covariance structure with current software, without the need for case-wise deletion of incomplete data.

Standard covariance structure models include the mixed model with random effects, the compound symmetry model, and the AR(1) model. Other models include a generalized version of the AR(1) model, a compound symmetric model plus an autoregressive error and a mixed model with random effects plus an autoregressive error. These different covariance structures are applied to longitudinal data and evaluated for goodness of fit. In addition to likelihood ratio tests, graphical methods are presented as a measure of goodness of fit for the covariance structure. The results are parsimonious models for both the expected value and covariance part of the model. Treatment of baseline values as covariates or additional time points is also investigated.

In conclusion, a brief analysis paradigm is presented and discussed as a guide for fitting the various covariance models and their fixed effects. Focused experience with these methods highlights the process of fitting these models and choosing among them the best fitting model to the data.
ACKNOWLEDGEMENTS

I would like to thank Dr. Ronald Helms for his continued guidance throughout this endeavor. His enthusiastic and refreshing approach to research kept me both interested and challenged from start to finish. Committee members Dr. Gerardo Heiss, Dr. Shirkant Bandiwalla, Dr. Paul Stewart and Dr. Gary Koch are acknowledged for their constructive comments.

My parents, Barbara and Jim Grady, deserve special thanks for their emotional support at every step along this sometimes rocky path. My academic achievements are a tribute to their appreciation and respect for higher education. My sisters Cheryl and Karen were supportive as well.

I would also like to express gratitude to my friends from West Haven, especially Mike McAllister, for their encouragement and support during my graduate studies.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xv</td>
</tr>
<tr>
<td>I. INTRODUCTION AND LITERATURE REVIEW</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Notation</td>
<td>4</td>
</tr>
<tr>
<td>1.2. Introduction to Models for Longitudinal Data</td>
<td>4</td>
</tr>
<tr>
<td>1.2.2 The General Linear Model</td>
<td>5</td>
</tr>
<tr>
<td>1.2.3 The General Linear Mixed Model</td>
<td>6</td>
</tr>
<tr>
<td>1.3 Historical Review of the Literature</td>
<td>7</td>
</tr>
<tr>
<td>1.3.1 Regular or Irregular and Consistent Data - Complete Data Case</td>
<td>7</td>
</tr>
<tr>
<td>1.3.2 Regular or Irregular and Consistent Data - Incomplete Data Case</td>
<td>8</td>
</tr>
<tr>
<td>1.3.3 Inconsistent and Incomplete Data</td>
<td>11</td>
</tr>
<tr>
<td>1.3.4 Models With Special Covariance Structures</td>
<td>17</td>
</tr>
<tr>
<td>1.4 General Computing Algorithms</td>
<td>17</td>
</tr>
<tr>
<td>1.4.1 The E-M Algorithm</td>
<td>19</td>
</tr>
<tr>
<td>1.4.2 Gradient Methods</td>
<td>20</td>
</tr>
<tr>
<td>1.5 Statement of the Problem and Outline</td>
<td>20</td>
</tr>
<tr>
<td>II. DESCRIPTION OF THE MODELS</td>
<td>23</td>
</tr>
<tr>
<td>2.1 The General Linear Model</td>
<td>23</td>
</tr>
<tr>
<td>2.1.1 Model Terminology and Equation</td>
<td>23</td>
</tr>
<tr>
<td>2.1.2 Introduction to Structures for Covariance Matrices</td>
<td>23</td>
</tr>
<tr>
<td>2.1.3 Introduction to Computing Aspects</td>
<td>26</td>
</tr>
</tbody>
</table>
2.2 The Mixed Model ......................................................... 26
  2.2.1 Model Equation and Terminology .......................... 26
  2.2.2 Description of the Covariance Structure .................. 27
  2.2.3 Computing Aspects for the Mixed Model ................. 27

2.3 Models with Stationary Variance and Exponentially
    Decreasing Correlation ........................................... 30
  2.3.1 Model Terminology and Equation ......................... 30
  2.3.2 Description of the Covariance Structure ................. 30
  2.3.3 Computing Aspects for these Models ................... 31

2.4 Mixed Model with Compound Symmetry Plus an
    Autoregressive Error ........................................... 37
  2.4.1 Model Equation and Terminology ........................ 37
  2.4.2 Description of the Covariance Structure ................. 38
  2.4.3 Computing aspects for these Models .................. 39

2.5 Mixed Model with Random Effects Plus an Autoregressive Error .... 44
  2.5.1 Model Equation and Terminology ......................... 44
  2.5.2 Description of the Covariance Structure ................. 44
  2.5.3 Computing aspects for these Models .................. 44

III. ANALYSIS OF TOTAL CHOLESTEROL DATA
  3.1 Total Cholesterol Measures .................................. 49
  3.2 Assessment of the Unmodeled Data .......................... 50
    3.2.1 Introduction ............................................. 50
    3.2.1 Trends in the Correlations and Covariances ........... 51

      Trends in the Correlations and Covariances
      for these data ........................................... 54
3.3 Modeling the Expected Value Part of the Model with Unstructured Covariance .................................................. 55

3.3.1 Initial Models .................................................................. 55

3.3.2 Mean Regression Plots ...................................................... 57

3.3.3 Checking Trends in the Covariances and Correlations .............................. 59

3.3.4 Homogeneity of the Variances of Repeated Outcome Measures ................. 60

   Assessment of Homogeneity for these Data .................................. 62

3.3.5 Conclusions ..................................................................... 64

3.4 Modeling the Covariance Matrix Using the Mixed Model, Compound Symmetry, and AR(1) Covariance Structures .......... 65

3.4.1 Introduction ..................................................................... 65

3.4.2 Evaluation of the Covariance Structure for the Cell Mean Model ................ 66

   Evaluation of Fixed Effects ...................................................... 68

   Conclusions ......................................................................... 68

3.4.3 Evaluation of the Covariance Structure for the Cubic Model ...................... 68

   Evaluation of Fixed Effects ...................................................... 69

   Conclusions ......................................................................... 70

3.5 Model With Stationary Variance and Exponentially Decreasing Correlation .......... 71

3.5.1 Introduction ..................................................................... 71

3.5.2 Evaluation of the Covariance Structure ...................................... 72

   Model with $\nu$ set to 1 ......................................................... 72

   Model with $\nu$ set to 2 ......................................................... 72

   Model with $\nu$ estimated as $\tilde{\nu}$ ........................................ 72

   Conclusions ......................................................................... 73
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.6</td>
<td>Mixed Model with Compound Symmetry Plus an Autoregressive Error</td>
<td>74</td>
</tr>
<tr>
<td>3.6.1</td>
<td>Introduction</td>
<td>74</td>
</tr>
<tr>
<td>3.6.2</td>
<td>Evaluation of the Covariance Structure</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Model with $\nu$ set to 1</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Model with $\nu$ set to 2</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Model with $\nu$ set to 3</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Conclusions</td>
<td>76</td>
</tr>
<tr>
<td>3.7</td>
<td>General Mixed Model With an Autoregressive Error</td>
<td>77</td>
</tr>
<tr>
<td>3.7.1</td>
<td>Introduction</td>
<td>77</td>
</tr>
<tr>
<td>3.7.2</td>
<td>Evaluation of the Covariance Structure</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Model with $\nu$ set to 1</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Model with $\nu$ set to 2</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Model with $\nu$ estimated as $\tilde{\nu}$</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Conclusions</td>
<td>79</td>
</tr>
<tr>
<td>3.8</td>
<td>Choosing the Best Model</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Some Overall Comments</td>
<td>82</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>INCLUSION OF BASELINE VALUES</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>The Lipid Research Clinics Coronary Primary Prevention Trial</td>
<td>83</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Description of the Study</td>
<td>83</td>
</tr>
<tr>
<td>4.2</td>
<td>Incorporating Baseline Data into the Analysis</td>
<td>87</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Choice of Analysis Variables</td>
<td>87</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Choice of Analysis of the Repeated Measures</td>
<td>89</td>
</tr>
<tr>
<td>4.3</td>
<td>Treating Baseline Values as Additional Time Points</td>
<td>93</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Introduction - Unstructured $\hat{\Sigma}$</td>
<td>93</td>
</tr>
</tbody>
</table>
4.3.2 Using Individual Baseline Measures as Additional
Time Points ................................................................. 94

A Note on Unequal Spacing and BMDP-5V ................. 96
Evaluation of the Models .............................................. 96
Graphical Assessment of the Covariance Structures .... 98

4.3.3 Using Average Baseline Measures as Additional
Time Points ................................................................. 99
Evaluation of the Models .............................................. 101
Graphical Assessment of the Covariance Structures ... 103
To Average or Not to Average ..................................... 103

4.4 Treating Baseline Values as Covariates ..................... 104
4.4.1 Introduction ....................................................... 104
4.4.2 Choosing a Covariate .......................................... 106
Preliminary Plots ..................................................... 107

4.4.3 Models with Cholesterol Measure at Visit 1 Treated
as a Covariate with Other Baseline Measures Treated
as Additional Time Points ........................................ 108
Evaluation of the Models ............................................. 108
Graphical Assessment of the Covariance Structures ... 109

4.4.4 Models with Mean Baseline Measure from Visits
1 and 2 Treated as a Covariate, Mean Baseline
Measure from Visits 3 – 5 is an Additional Time Point ..... 110
Evaluation of the Models ............................................. 110
Graphical Assessment of the Covariance Structures ... 111

4.5 Conclusions ............................................................ 112
4.5.1 When Baseline Values are Treated as Additional
Time Points ................................................................. 112
4.5.2 When Baseline Values are Treated as Covariates .... 114
4.5.3 Comparing the Covariate Method to the Additional
Time Point Method ....................................................... 116
Comparisons of Treatment Differences ....................... 116
V SUMMARY AND SUGGESTIONS FOR FUTURE DEVELOPMENTS

5.1 An Analysis Paradigm ................................................................. 118
  5.1.1 Analysis Paradigm ............................................................... 118

5.2 Future Developments .................................................................. 122
  5.2.1 Some Notes on BMDP-5V ....................................................... 125

REFERENCES .................................................................................... 227
# LIST OF TABLES

| Table 3.1 | Pattern of Missing Data for Total Cholesterol | 127 |
| Table 3.2 | Simple Covariance and Correlation Matrices, and Group Means | 128 |
| Table 3.3 | Estimated Covariance and Correlation Matrices From the Cell Mean Model | 129 |
| Table 3.4 | Estimated Covariance and Correlation Matrices From the Cell Mean Model, Equal Variances and Unstructured Covariances | 130 |
| Table 3.5 | Estimated Covariance Matrices Using the Cell Mean Model | 131 |
| Table 3.6 | Estimated Correlation Matrices of Using the Cell Mean Model | 132 |
| Table 3.7 | Model Parameters for the Cell Mean Model | 133 |
| Table 3.8 | Estimated Covariance Matrices Using the Cubic Model | 134 |
| Table 3.9 | Estimated Correlation Matrices Using the Cubic Model | 135 |
| Table 3.10 | Estimated Covariance and Correlation Matrices Using the Cubic Model with Unstructured Covariance | 136 |
| Table 3.11 | Model Parameters for the Cubic Model | 137 |
| Table 3.12 | Estimated Covariance and Correlation Matrices Model With Stationary Variance and Exponentially Decreasing Correlation, $\nu = 1$ | 138 |
| Table 3.13 | Estimated Covariance and Correlation Matrices Model With Stationary Variance and Exponentially Decreasing Correlation, $\nu = 2$ | 139 |
| Table 3.14 | Estimated Covariance and Correlation Matrices, Model With Stationary Variance and Exponentially Decreasing Correlation, $\nu = \hat{\nu}$ | 140 |
| Table 3.15 | Covariance Parameter Estimates and Standard Errors for Models with Stationary Variance and Exponential Decreasing Correlation | 141 |
| Table 3.16 | Estimated Covariance and Correlation Matrices, Models With Compound Symmetry Plus an Autoregressive Error, $\nu = 1$ | 142 |
| Table 3.17 | Estimated Covariance and Correlation Matrices, Models With Compound Symmetry With an Autoregressive Error, $\nu = 2$ | 143 |
Table 3.18  Estimated Covariance and Correlation Matrices  
Models With Compound Symmetry With an  
Autoregressive Error, $\nu = 3$ ......................................................... 144

Table 3.19  Covariance Parameter Estimates for Models with  
Compound Symmetry Plus an Autoregressive Error —  
Cell Mean Model ................................................................. 145

Table 3.20  Covariance Parameter Estimates for Models with  
Compound Symmetry Plus an Autoregressive Error —  
Cubic Model ................................................................. 146

Table 3.21  Estimated Covariance and Correlation Matrices,  
Mixed Models With Random Effects Plus an  
Autoregressive Error, $\nu = 1$ ......................................................... 147

Table 3.22  Estimated Covariance and Correlation Matrices,  
Mixed Models With Random Effects Plus an  
Autoregressive Error, $\nu = 2$ ......................................................... 148

Table 3.23  Estimated Covariance and Correlation Matrices,  
Mixed Models With Random Effects Plus an  
Autoregressive Error, $\nu = 3$ ......................................................... 149

Table 3.24  Covariance Parameter Estimates for the Cell Mean Model,  
Mixed Model With Random Effects Plus an  
Autoregressive Error ................................................................. 150

Table 3.25  Covariance Parameter Estimates for the Cubic Model  
Mixed Model With Random Effects Plus an  
Autoregressive Error ................................................................. 151

Table 3.26  Covariance Parameter Estimates for Models with  
Compound Symmetry Plus an Autoregressive Error —  
Cell Mean Model ................................................................. 152

Table 3.27  Likelihood Ratio Tests, Cell Mean Model ......................... 153

Table 3.28  Likelihood Ratio Tests, Cubic Model ................................. 154

Table 4.1  Time Sequence of Protocol Visits, Diet, and Treatment  
Randomization for First Year ......................................................... 155

Table 4.2  Model with 2 Averaged Baseline Values Treated as  
Additional Time Points, Unstructured Covariance Matrix  
with a Cell Mean Model ......................................................... 156

Table 4.3  Model with 5 Baseline Values Treated as Additional  
Time Points, Unstructured Covariance Matrix  
with a Cell Mean Model ......................................................... 157

Table 4.4  Likelihood Ratio Tests, Model with 5 Baseline Values  
Treated as Additional Time Points, Cell Mean Model  
Plus Cubic Contrasts ......................................................... 158
<table>
<thead>
<tr>
<th>Table 4.5</th>
<th>Estimates for Mixed Model With Compound Symmetry Plus AR(1) Error, Model with 5 Baseline Values Treated as Additional Time Points, Cell Mean Model Plus Cubic Contrasts ..................................................... 159</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 4.6</td>
<td>Estimates for Models with Stationary Variance and Exponentially Decreasing Correlation, Model with 5 Baseline Values Treated as Additional Time Points, Cell Mean Model Plus Cubic Contrasts ..................................................... 160</td>
</tr>
<tr>
<td>Table 4.7</td>
<td>Likelihood Ratio Tests, Model with 2 Average Baseline Values Treated as Additional Time Points, Cell Mean Model Plus Cubic Contrasts ..................................................... 161</td>
</tr>
<tr>
<td>Table 4.8</td>
<td>Estimates for Mixed Model With Compound Symmetry Plus AR(1) Error, Model with 2 Averaged Baseline Values Treated as Additional Time Points, Cell Mean Model Plus Cubic Contrasts ..................................................... 162</td>
</tr>
<tr>
<td>Table 4.9</td>
<td>Estimates for Models with Stationary Variance and Exponentially Decreasing Correlation, Model with 2 Average Baseline Values Treated as Additional Time Points, Cell Mean Model Plus Cubic Contrasts ..................................................... 163</td>
</tr>
<tr>
<td>Table 4.10</td>
<td>Model with 2 Averaged Baseline Values Treated as Additional Time Points, Comparison of $\hat{\Sigma}$ for Diggle's Model ..................................................... 164</td>
</tr>
<tr>
<td>Table 4.11</td>
<td>Model with 2 Averaged Baseline Values Treated as Additional Time Points, $\hat{\Sigma}$ for the General Mixed Model ..................................................... 165</td>
</tr>
<tr>
<td>Table 4.12</td>
<td>Likelihood Ratio Tests, Model with Cholesterol Measure at Visit 1 Treated as a Covariate, Baseline Measures from Visits 2–5 are Additional Time Points, Cell Mean Model Plus Cubic Contrasts ..................................................... 166</td>
</tr>
<tr>
<td>Table 4.13</td>
<td>Estimates for Models with Stationary Variance and Exponentially Decreasing Correlation, Model with Cholesterol Measure at Visit 1 Treated as a Covariate, Baseline Measures from Visits 2–5 are Additional Time Points, Cell Mean Model Plus Cubic Contrasts ..................................................... 167</td>
</tr>
<tr>
<td>Table 4.14</td>
<td>Covariance Matrices, Model with Cholesterol Measure at Visit 1 Treated as a Covariate, Baseline Measures from Visits 2–5 are Additional Time Points, Cell Mean Model Plus Cubic Contrasts ..................................................... 168</td>
</tr>
<tr>
<td>Table 4.15</td>
<td>Likelihood Ratio Tests, Model with Mean Baseline Measure from Visits 1 &amp; 2 Treated as a Covariate, Mean Baseline Measure from Visits 3–5 is an Additional Time Point, Cell Mean Model Plus Cubic Contrasts ..................................................... 169</td>
</tr>
<tr>
<td>Table 4.16</td>
<td>Likelihood Ratio Tests, Model with Mean Baseline Measure from Visits 1 &amp; 2 Treated as a Covariate, Mean Baseline Measure from Visits 3–5 is an Additional Time Point, Cell Mean Model Plus Cubic Contrasts ..................................................... 170</td>
</tr>
</tbody>
</table>
Table 4.17 Covariance Matrices, Model with Mean Baseline Measure from Visits 1 & 2 Treated as a Covariate, Mean Baseline Measure from Visits 3-5 is an Additional Time Point, Cell Mean Model Plus Cubic Contrasts

171
LIST OF FIGURES

Figure 3.1 Simple Group Means ................................................. 172
Figure 3.2 Correlation Plot, Simple Correlation Matrix .................. 173
Figure 3.3 Covariance Plot, Simple Covariance Matrix ................... 174
Figure 3.4 Model-Based Estimates of Group Means, Cell Mean Model,
Unstructured Covariance .................................................. 175
Figure 3.5 Model-Based Estimates of Group Means, Cubic Model,
Unstructured Covariance .................................................. 176
Figure 3.6 Model-Based Estimates of Group Means, Quadratic Model,
Unstructured Covariance .................................................. 177
Figure 3.7 Model-Based Estimates of Group Means, Linear Model,
Unstructured Covariance .................................................. 178
Figure 3.8 Correlation Plot, Cubic Model with Unstructured Covariance..... 179
Figure 3.9 Covariance Plot, Cubic Model with Unstructured Covariance ...... 180
Figure 3.10 Correlation Plot, Cell Mean Model, Unstructured Covariance ...... 181
Figure 3.11 Covariance Plot, Cell Mean Model, Unstructured Covariance ...... 182
Figure 3.12 Covariance Plot, Ordinary Least Squares Method .................. 183
Figure 3.13 Covariance Plot, Cell Mean Model, Mixed Model Covariance ...... 184
Figure 3.14 Correlation Plot, Cell Mean Model, Mixed Model Covariance ...... 185
Figure 3.15 Covariance Plot, Cell Mean Model with AR(1) Covariance ............ 186
Figure 3.16 Correlation Plot, Cell Mean Model with AR(1) Covariance ............ 187
Figure 3.17 Covariance Plot, Cubic Model with Mixed Model Covariance ...... 188
Figure 3.18 Correlation Plot, Cubic Model with Mixed Model Covariance ...... 189
Figure 3.19 Covariance Plot, Cubic Model with AR(1) Covariance ............... 190
Figure 3.20 Correlation Plot, Cubic Model with AR(1) Covariance ............... 191
Figure 3.21 Model-Based Estimates of Group Means, Cubic Model,
Mixed Model Covariance .................................................. 192
Figure 3.22 Model-Based Estimates of Group Means, Cubic Model,
Mixed Model Plus AR Error Covariance .................................... 193
Figure 4.10 Model-Based Means, Murray’s Model Covariance, 5 Baseline Values are Added Time Points ................................................. 211

Figure 4.11 Covariance Plot: 2 Mean Baseline Values are Added Time Points, Unstructured Covariance ......................................................... 212

Figure 4.12 Covariance Plot: 2 Mean Baseline Values are Added Time Points, Mixed Model Covariance ........................................................... 213

Figure 4.13 Covariance Plot: 2 Mean Baseline Values are Added Time Points, Diggle’s Model ............................................................................. 214

Figure 4.14 Covariance Plot: 2 Mean Baseline Values are Added Time Points, Murray’s Model ................................................................. 215

Figure 4.15 Model-Based Estimates of Group Means, Covariate is CHL at Visit 1 ............................................................................................... 216

Figure 4.16 Covariance Plot: Cell Mean Model, Covariate is CHL at Visit 1 ........................................................................................................ 217

Figure 4.17 Model-Based Estimates of Group Means, Covariate is Mean CHL of Visits 1 and 2 ................................................................. 218

Figure 4.18 Covariance Plot: Cell Mean Model, Covariate is Mean CHL of Visits 1 and 2 ............................................................................. 219

Figure 4.19 Covariance Plot: Unstructured Covariance, Covariate is CHL at Visit 1 ....................................................................................... 220

Figure 4.20 Covariance Plot: General Mixed Model, Covariate is CHL at Visit 1 ............................................................................................. 221

Figure 4.21 Covariance Plot: Murray’s Model, Covariate is CHL at Visit 1 ................................................................................................. 222

Figure 4.22 Covariance Plot: Unstructured Covariance, Covariate is Mean CHL of Visits 1 and 2 ................................................................. 223

Figure 4.23 Covariance Plot: General Mixed Model, Covariate is Mean CHL of Visits 1 and 2 ................................................................. 224

Figure 4.24 Covariance Plot: Murray’s Model, Covariate is Mean CHL of Visits 1 and 2 ................................................................. 225

Figure 4.25 Model-Based Estimates of Group Means Models: Covariate V1 vs 5 Added Values ................................................................. 226

Figure 4.26 Model-Based Estimates of Group Means Models: Mean Covariate vs 2 Added Means ................................................................. 227
CHAPTER 1
INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

Longitudinal studies are among the principle research designs used in medical and social research because an individual's changes over time can be observed and the effects of these changes can be studied. However, along with the known advantages of longitudinal studies we find that the typical problems associated with them, such as missing data and data that are irregularly or inconsistently timed, cause problems when analyzing the data. Missing data are handled poorly with standard multivariate analysis methods which use fully parametrized covariance matrices since they require case-wise deletion: all data are discarded for each subject with missing data. In most longitudinal studies there are both missing data and data collected on occasions other than those specified by the study protocol, making the data both incomplete and irregularly timed. When different individuals have different schedules resulting in highly unbalanced data, the methods mentioned earlier, involving general multivariate models with fully parametrized covariance matrices, are not amenable to analysis. This is mainly due to the proliferation of covariance parameters, many of which will be poorly estimated as the covariance matrix becomes large, as discussed by Ware (1985).

Although the theory and methods for the analysis of incomplete longitudinal data exist, these methods have not been readily available to
practicing statisticians, due, in part, to limited exposure and obscure software. As a result, these new methodologies for analyzing longitudinal data have not found their way into the repertoire of many statisticians. The objective of this paper is to show how these new analysis 'tools' can be used to model longitudinal data, in particular the covariance structure, without the need for case-wise deletion of incomplete data. The results are parsimonious models for both the expected value and covariance part of the model. Focused experience with these tools will be presented to highlight not just the results of a data analysis, but also the process of fitting these models and choosing among the various covariance structures available to the analyst. Demonstration analysis will also provide a useful introduction for those unfamiliar with these incomplete longitudinal data methodologies.

Due to the nature of longitudinal studies, which can last for months or even years and require data to be collected from subjects on multiple occasions, irregularly and inconsistently timed data are common problems. Data are regularly timed if the time interval between measurement occasions is the same throughout the study (e.g. monthly occasions). Irregularly timed data have time intervals that differ. An example is a study with occasions after the first, second, third, and sixth months, after one year and after five years. Data are consistently timed if all subjects are evaluated on the same schedule. Inconsistently timed data, for example, result from a study of pulmonary function in which data are collected at the onset of each episode of respiratory illness, leading to subjects observed for different sets of times. Inconsistently timed data can also arise from individuals who simply cannot keep scheduled appointments and upon rescheduling have data collected at odd time points, creating different schedules for different subjects. In this particular example most of the subjects might have regularly timed data while a subset would not. Longitudinal study data can be
considered to fall into four categories: regular and consistent, regular and inconsistent, irregular and consistent, and irregular and inconsistent.

The data are said to be balanced if there are no missing data in a consistently timed longitudinal study. Data from inconsistently timed studies are said to be unbalanced. A study with missing data values results in an incomplete data structure. The term complete would then describe any study with no missing data. Incomplete and unbalanced data both lead to similar analysis difficulties in longitudinal data studies.

Missing data, presumably from a random process (Rubin, 1976), can result from many causes over the course of a typical longitudinal study. Case-wise deletion is sometimes used to cope with missing data. While this may allow for the use of standard multivariate analysis methods, such deletion can result in a large number of subjects being dropped from the analysis -- after the expense of having them in the study. Case-wise deletion, therefore, is not a reasonable option when there is a large proportion of missing data. Ware (1985) reviewed the available methods for analyzing longitudinal studies when data are not discarded.

For these studies the choice of the longitudinal metameter is usually elapsed time but may be some other parameter such as height or weight. In a study of growing swine the variable of interest, the ratio of subcutaneous fat to overall weight, is considered to be related to overall weight. Weight at time of measure is an appropriate metameter since it is expected to be an important covariate. In a pulmonary function study of children, the metameter could be height at time of measurement if pulmonary function is thought to depend on body size. If the covariates were thought to depend on age, then the age at time of measure might be the proper metameter. Ware (1985) notes that the choice of metameter can influence the amount of imbalance in a data set, noting that in the above example, using height at time of measure as the metameter can lead to
greater imbalance than age at time of measure.

When the data set is highly unbalanced or incomplete, Ware (1985), Jennrich and Schluchter (1986), Chi (1989), and others suggest the use of parsimonious covariance structures. Estimating the covariance structure as an arbitrary positive definite symmetric matrix for individual subjects can be computationally burdensome and can lead to inefficient estimates of location parameters, compared to the use of simpler covariance structures. In some cases the intercorrelation of the errors of the serial measurements make the use of autoregressive models an attractive choice for a structured covariance matrix which can be estimated using fewer parameters than the unstructured covariance matrix. In other cases the mixed model with random effects or a compound symmetry model offer parsimonious alternatives to a more highly parametrized matrix. The mixed model and autoregressive models have also been used in combination by Diggle (1988) to model the covariance structure.

1.2 Notation

1.2.1 Introduction to Models for Longitudinal Data

The usual maximum likelihood approach assumes that the complete data vector $\mathbf{Y}$ can be modeled as multivariate normal. An important feature is that each element of $\mathbf{Y}$ is a measure of the same phenomenon at different time points and possibly under different conditions. Laird (1988) states that standard multivariate analyses are not particularly useful for analysis of longitudinal data vectors and for this reason many popular models for longitudinal data are based on some form of the linear growth model. The linear growth model assumes that the mean of $Y_{it}$, the $i$-th measure on the $i$-th person, is a linear function\(^1\) of time.

\(^1\)This linear function is not necessarily a straight-line function.
(or age) and covariates and design variables.

1.2.2 The General Linear Model

The following notation will be used to describe the general linear model for fixed effects. The model equation for the $i$-th subject is

$$ \mathbf{y}_i = \mathbf{x}_i \beta + \mathbf{e}_i $$

(1.2.1)

where $\mathbf{y}_i$ is an $n_i \times 1$ vector of observations on the $i$-th subject, $i = 1, \ldots, K$,

$\beta$ is a $p \times 1$ vector of unknown constant population parameters,

$\mathbf{x}_i$ is an $n_i \times p$ known constant design matrix, corresponding to the fixed effects $\beta$, and

$\mathbf{e}_i$ is an $n_i \times 1$ vector of random error terms.

The following assumptions are made:

$\mathbf{e}_i \sim \text{NID}(0, \Sigma_i)$, so that $\mathbf{e}_i$ is independent of $\mathbf{e}_j$ for all $i \neq j$,

$\Sigma_i$ is a positive-definite symmetric matrix whose elements are twice differentiable functions of a finite set number of constant, unknown parameters, $\xi_1, \ldots, \xi_m$, i.e. $\Sigma_i = \Sigma_i(\xi) = \xi \in \mathcal{T}$.

$\mathcal{T}$ is the set of parameter values that make $\mathcal{T}_i$ positive-definite.

The parameters in $\beta$ are functionally independent of those in $\xi$. It follows that $E[\mathbf{y}_i] = \mathbf{x}_i \beta$ and $\text{Var}[\mathbf{y}_i] = \Sigma_i$. As the name implies, this is a rather general model of which ANOVA and ordinary regression models are special cases. Also of note is that the general linear mixed model below may be considered a special case of this model in which $\Sigma_i = \mathbf{Z}_i \Delta \mathbf{Z}_i' + \sigma^2 \mathbf{I}_i$.

This model can be easily generalized to allow $\Sigma_i$ to depend upon selected covariates or design variables. $\Sigma_i$ can also be a structured function of a set of unknown parameters. Regardless of the assumed structure for $\Sigma_i$, the ML
estimate of $\hat{\beta}$ is given by

$$
\hat{\beta} = \left( \sum_{i=1}^{k} X_i' \hat{\Sigma}_i^{-1} X_i \right)^{-1} \left( \sum_{i=1}^{k} X_i' \hat{\Sigma}_i^{-1} Y_i \right),
$$

where $\hat{\Sigma}_i$ is the ML estimate of $\Sigma_i$. The amount of complexity of the computation for the complete data case is directly related to the assumed structure of $\Sigma_i$ and $X_i$ (Laird, 1988). An advantage of this model is that it is quite general, but even with complete data, closed form solutions for the ML estimates usually do not exist. Closed form ML estimates for $\beta$, $\Sigma_i$ and var($\hat{\beta}$) exist only when $X_i$ and/or $\Sigma_i$ take special forms (Lange and Laird, 1986; Sztawrowski, 1983; Andrade and Helms, 1984). When closed form solutions do not exist iterative methods such as the EM algorithm or general scoring algorithms must be used to obtain parameter estimates.

1.2.3 The General Linear Mixed Model

The following notation will be used to describe mixed models which accommodate both fixed and random effects. Note that the error terms ($\varepsilon_i$) within a vector of observations need not be independent. Consider the model:

$$
Y_i = X_i \beta + Z_i \delta_i + \varepsilon_i
$$

(1.2.2)

where $Y_i$ is an $n_i \times 1$ vector of observations on the $i$-th subject, $i = 1, ..., K$,

$\beta$ is a $p \times 1$ vector of unknown constant population parameters,

$X_i$ is an $n_i \times p$ known constant design matrix, corresponding to the fixed effects $\beta$,

$\delta_i$ is a $q \times 1$ vector of unknown, random individual parameters,

$Z_i$ is an $n_i \times q$ known constant design matrix, corresponding to the random effects $\delta_i$, and

$\varepsilon_i$ is an $n_i \times 1$ vector of random error terms.

The following assumptions are made:

$\delta_i \sim \text{NID}(0, \Delta)$, and is independent of $\varepsilon_i \sim \text{NID}(0, \sigma^2 Y_i)$,
so that \( \text{Var} \begin{bmatrix} \mathcal{Y}_i \\ \mathcal{E}_i \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{Q} \\ \mathbf{Q} & \sigma^2 \mathbf{Y}_i \end{bmatrix} \)

where \( \mathbf{A} \) is a \( q \times q \) positive-definite, symmetric covariance matrix of the random effects \( (\mathcal{Y}_i) \) which depends upon a set of unknown parameters (e.g., \( \mathbf{I} \)), \( \mathbf{Y}_i \) is a known constant positive-definite symmetric matrix, and \( \sigma^2 \) is the scaler within-subject variance parameter. Also, \( \text{E}[\mathbf{Y}_i] = \mathbf{X}_i \mathbf{B} \) and \( \text{Var}[\mathbf{Y}_i] = \Sigma_i = \mathbf{Z}_i \mathbf{A} \mathbf{Z}_i^T + \sigma^2 \mathbf{Y}_i \) is the \( n_i \times n_i \) positive definite symmetric covariance matrix of \( \mathbf{Y}_i \), and \( \mathbf{Y}_i \) is independent of \( \mathbf{Y}_j \) for \( i \neq j \).

1.3 Historical Review of the Literature

This section will be a brief synopsis of the literature on approaches to analyzing data from longitudinal studies. It will be historical in the sense that it will attempt to chronicle the methodology starting with the "early" (1960's) approaches and ending with recent developments. The notions of regular, consistent, and complete data, as described in the Introduction, will be helpful for this section.

1.3.1 Regular or Irregular and Consistent Data - Complete Data Case

The fundamental design for longitudinal studies is the complete design where all subjects are measured on the same occasions and there is the assumption that measures taken from the same subject are correlated. One model for this design is a special case of the general lineal model called the general linear multivariate model: GLMM(\( \mathbf{Y}; \mathbf{X}_i \mathbf{B}, \Sigma_i \)) with one dependent variable for each occasion (or time point). Each subject has the same expected value and covariance structure, namely \( \text{E}[\mathbf{Y}_i] = \mathbf{X}_i \mathbf{B} \) and \( \text{V}[\mathbf{Y}_i] = \Sigma_i = \Sigma, i=1,\ldots,K. \) In this model \( \Sigma \) is unstructured, meaning that it can be any positive definite symmetric
matrix. For the GLMM with complete data, the m.l.e. of \( \hat{\beta} \) is 
\[ \hat{\beta} = (X'X)^{-1}X'Y \]
and the 'usual' unbiased estimator of \( \Sigma \) is
\[ \hat{\Sigma} = (Y - X\hat{\beta})(Y - X\hat{\beta})/[N - \text{rank}(X)]. \]
(See for example Rao 1965; Johnson and Wichern, 1988).

For balanced and complete "growth curve" data with no missing data, Potthoff and Roy (1964) and Grizzle and Allen (1969) examined the "growth curve model", a special case of the general multivariate analysis of variance model
\[ y_i = X_i \beta + \varepsilon_i, \quad i = 1, \ldots, N, \]
where \( y_i \) is the \( p \times 1 \) vector of observations on the \( i \)-th individual, \( X_i \) is the \( p \times q \) within individual design matrix, \( \beta \) is the \( q \times r \) matrix of unknown parameters, \( \varepsilon_i \) is the \( r \times 1 \) vector of non-time varying covariates for individual \( k \), and \( \varepsilon_i \) are \( p \times 1 \) errors assumed to be independently distributed as multivariate normal with mean 0 and unstructured \( p \times p \) covariance matrix \( \Sigma \). This model is a special case of the GLMM with a structured covariance matrix. An important feature of the model is the use of "error contrasts" as variance-reducing covariates as discussed by Timm (1970).

1.3.2 Regular or Irregular and Consistent Data - Incomplete Data Case

The studies considered in this section are ones that were planned to be complete and consistently timed but had incomplete data at the end of the study. This can be thought of as the imperfect version of the complete, consistently timed longitudinal study described in Section 1.3.1. There have been many proposals as to how to handle such data and below a few are mentioned.

One early paper on missing data techniques was by Wilkes (1932). His idea was to simply replace the missing data values with their means and to then analyze the data using methods for the complete data case.
Buck (1960) suggested using multiple regression estimates to replace missing values in a GLMM. For each dependent variable with missing data a regression equation was obtained using other dependent variables as regressors and using all available data. The available values of the other dependent variables from the $i$-th person are then used in the resulting regression equation to compute the “new” value for that person’s missing value.

Glasser (1964) was apparently the first to publish a description of pairwise deletion as a missing value technique. His idea was to use the mean and variance of the nonmissing observations as estimates for the missing data. The covariance of, say, variables $X$ and $Y$ is estimated using all of the observations with data for both $X$ and $Y$. Hence the pairwise deletion comes in estimating the covariance when an observation is missing either $X$ or $Y$ or both. One disadvantage of pairwise deletion is that $\hat{\Sigma}$ may have negative eigenvalues.

Pairwise deletion was compared to casewise deletion by Haitovsky (1968) in a Monte-Carlo study which concluded casewise deletion to be superior to pairwise deletion under some circumstances. Timm (1970) compared casewise deletion to the techniques of Buck (1960) and Wilks (1932) for estimating $\hat{\Sigma}$ concluding Buck’s method to be better than casewise deletion and Wilks’ method. Gleason and Staelin (1975) compared heuristic techniques in a large Monte-Carlo study for estimating the correlation matrix. They concluded that their own method, which involved using principle component scores, was better than methods of Wilks and Buck for estimating $\hat{\Sigma}$.

Hosking (1980) conducted a Monte-Carlo study to compare three missing data techniques for estimating $\hat{\Sigma}$ and $\hat{\beta}$ in the general linear multivariate model setting. The comparison between pairwise and casewise deletion for estimating $\hat{\Sigma}$ showed casewise deletion to be the worst method in the setting he studied.
Barton (1986) compared methods that do not require deletion of any observed data to the method of casewise deletion for test size and test power in the GLMM with incomplete data. He found Kleinbaum’s method (1970, 1973) using the Wald statistic and the Hotelling-Lawley trace criterion with adjustment for error degrees of freedom to perform poorly (inflated Type I error rates), while results using the EM algorithm were more promising. Rao’s F approximation for Wilk’s $\Lambda$ with adjustment for error degrees of freedom yielded slightly more conservative test power than casewise deletion.

A major development in missing data techniques was the use of maximum likelihood estimation based on the assumption of normality. Orchard and Woodbury (1972) proposed the “missing information principle” to deal with incomplete data based on the GLMM with normality. Their basic procedure had three steps. The first was to estimate the expected values of the missing data conditional on the observed data and the initial estimates of the unknown parameters. The second step involved producing maximum likelihood estimates of the parameters based on the now ‘complete data’ (the observed and estimated data). The third step was to re-estimate the missing data with the new parameters, repeating the procedure until the new estimates converge. For example, the first step in the GLMM setting would be the estimation of $\hat{\beta}$ and $\hat{\Sigma}$ using the set of individuals with complete data. The estimates of $\hat{\beta}$ and $\hat{\Sigma}$ would then be used to estimate the missing values. Each iteration increases the likelihood of the observed data given $\hat{\beta}$ and $\hat{\Sigma}$, until convergence of $\hat{\beta}$ and $\hat{\Sigma}$. The motivation for this approach is that the maximum likelihood equations for the ‘complete data’ conditional on the observed data are more easily obtained than the equations based on the observed data alone. Rubin (1974) factored the likelihood of the data into the product of two likelihoods, conditioning on the patterns of missing data to obtain parameter estimates from incomplete data.
The work of Orchard and Woodbury (1972) was a precursor of the more general E-M algorithm of Dempster, Laird and Rubin (1977), which is described in Section 1.4. The M-step of the the EM algorithm is similar to the Orchard-Woodbury step of estimating missing data; the difference is that in the E-step one estimates a set of sufficient statistics (for all unknown parameters). In the M-step the estimates of the sufficient statistics are used as "data" to compute ML estimates of unknown parameters. The advantage of the EM algorithm is that it can be applied in much wider circumstances, i.e. to models other than GLMM and to distributions other than multivariate normal.

Little and Scluchter (1985) developed maximum likelihood procedures for analyzing mixed continuous and categorical data with missing data. The model handles missing data in both the continuous and the categorical variables. The key model assumption is that the conditional distribution of the continuous response variable, given the categorical variable, is a multivariate normal distribution with certain mean and unstructured covariance matrix. Maximum likelihood estimation with incomplete data is achieved through the application of the E-M algorithm. Laird (1991) also discusses likelihood-based methods for longitudinal data analysis with categorical responses.

1.3.3 Inconsistent and Incomplete Data

In many longitudinal studies there are both missing data and data collected on occasions other than those specified by the study protocol, making the data both inconsistently timed and incomplete. When different individuals have different schedules resulting in highly unbalanced data, the methods mentioned earlier, involving general multivariate models with unrestricted covariance structures, are not amenable to analysis. This is mainly due to the proliferation of variance parameters, many of which will be poorly estimated as $\Sigma$ becomes large.
A class of purposefully incomplete longitudinal studies, called "linked cross-sectional studies", have been described by Rao and Rao (1966) and by Woolson et al. (1978). In these two studies, in which growth rates of young boys were obtained, the cost and length of time required for a longitudinal study were reduced dramatically by observing "cross-sectional" cohorts for two and four years, respectively.

Early uses of maximum likelihood (ML) as a technique for estimating variance components were reviewed by Harville (1977), involving a general mixed model which allowed for fixed and random effects. One criticism of the ML method is that variance component estimators for ordinary ANOVA methods obtained by solving the ML equations do not generally coincide with those obtained from ANOVA methods, even in the case of balanced data. The ML estimates tend to be biased downward (Patterson and Thompson 1971). Another criticism of the ML approach is that it does not account for the loss in degrees of freedom that results from estimating fixed effects, namely $\beta$.

These "deficiencies" may be ameliorated through the use of restricted maximum likelihood estimation\(^2\) (REML), a technique explored by Russel and Bradley (1958), W. A. Thompson (1962) and generalized by Patterson and R. Thompson (1971 and 1974). In a simple random sample of $x_i \sim \text{i.i.d.} \ N(\mu, \sigma^2)$ the REML estimate of $\sigma^2$ is $\sum_i (x_i - \bar{x})^2 / (n-1)$ whereas the ML estimate of $\sigma^2$ is $\sum_i (x_i - \hat{x})^2 / (n-1)$.

\(^2\) The REML estimate of $\hat{\beta}$, where $L(\hat{\beta}; \tilde{y})$ is the likelihood function, is based on a full-rank set of error contrasts, $U'y$, rather than on the full data vector $y$. An error contrast is a linear combination of the data such that $E[U'y] = 0$, and any set is appropriate since different sets will only differ by a linear transformation, which does not depend on $\hat{\beta}$. In many situations it is necessary to use iterative techniques to determine the REML estimates, just as in ML estimation.
\[ \sum_i (x_i - \bar{x})^2 / n. \] Using the REML approach gives the same variance estimates as ordinary ANOVA methods for balanced ANOVA models. In such cases the REML estimates are unbiased. In many cases neither ML nor REML is unbiased when iterative techniques are required.

Searle (1988), while tending to agree with the above, was inconclusive about which method, ML or REML, he preferred for mixed models with unbalanced data. He noted that ML has the advantage of simultaneously providing estimators of both fixed effects and the variance components, with the fixed effects estimators having all the attendant properties of ML. The REML method, meanwhile, provides no such estimator for the fixed effects although intuitive alternates exist. However, REML has the advantage of providing variance component estimators that are unaffected by the fixed effects.

Laird and Ware (1982) examined the use of the mixed model, or “two stage random-effects model”, for longitudinal data, extending the work by Harville (1977). As described above, these models contain two sets of parameters allowing modeling for both within- and between-individual effects. While these two-stage models allow for the definition and estimation of (random) individual characteristics, their major limitation when compared to the general multivariate model is the special form assumed for the covariance structure. Laird and Ware considered both maximum likelihood and empirical Bayes estimation with the EM algorithm for parameter estimation.

Gibbons et al. (1988) compared traditional methods used to analyze longitudinal psychiatric data, including fixed-effects ANOVA, endpoint analysis (analysis of change scores) and growth curve models, to the use of mixed models. They concluded that the “random regression” (or mixed model) approach to the analysis of psychiatric data provided improvements and flexibility over the
traditional approaches. The random regression models work well with psychiatric patients due to the considerable amount of heterogeneity in the responses of individuals to treatment. In light of this heterogeneity, estimates of treatment related effects can be characterized with better precision using a mixed model. Helms and Edwards (1991) used mixed-models to analyze cross-over data.

Rosner and Muñoz (1988) apply autoregressive modeling to the analysis of longitudinal data with unequally spaced data. Their approach involves using linear interpolation to “fill in” missing data to create equally spaced data. For example, if a person has data at visits one and three but missed visit two, a model requiring equidistant occasions could not accommodate this person. Through linear interpolation, visit two would be given an interpolated value and this person would remain in the analysis. Estimation and hypothesis testing procedures for the model are done with weighted non-linear regression programs. A Monte-Carlo study showed this approach to have “no meaningful bias” and “considerable gains in efficiency” when compared to a linear regression model based only on observed equidistant pairs.

1.3.4 Models With Special Covariance Structures

A useful class of linear covariance structure models exist, characterized by \( \Sigma \) being linear in the parameters, that is, linear models in which

\[
\Sigma_i = \sum_{i=1}^{m} \tau_i G_i
\]

for \( n \times n \) symmetric matrices \( G_1, \ldots, G_m \) whose elements are known. A simple example of a linear covariance structure is the nested variance component model

\[
Y_{ij} = \mu + \alpha_i + \epsilon_{ij},
\]

for the \( j \)-th measure, \( j = 1, 2, \ldots, n_2 \) at the \( i \)-th level, \( i = 1, 2, \ldots, n_1 \). The model assumptions are

\[
\alpha_i \sim \text{NID}(0, \sigma^2_i),
\]
\[ \epsilon_{ij} \sim \text{NID}(0, \sigma_{\epsilon}^2), \] and that \( \alpha_i \) and \( \epsilon_{ij} \) independent.

If \( Y_{ij} = (y_{11}, y_{12}, \ldots, y_{1n_{2}}, \ldots, y_{n_{1}1}, y_{n_{1}2}, \ldots, y_{n_{1}n_{2}})' \) then
\[ \Sigma_i = \sigma_{\epsilon}^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \sigma_{\delta}^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \]
\[ = \tau_1 \mathcal{G}_1 + \tau_2 \mathcal{G}_2. \]

Other more complicated variance component models also follow this linear structure.

Fairclough and Helms (1984) and Andrade and Helms (1984) developed maximum likelihood estimation results for the general linear mixed model with linear covariance structures under the assumption of normality for unbalanced and incomplete data. They obtained \( \hat{\beta} \) and \( \hat{\xi} \) along with their asymptotic variances and covariances. Andrade and Helms (1984) also obtained estimates of both \( \hat{\beta} \) and \( \hat{\xi} \) with either (or both) subject to linear restrictions. They obtained the asymptotic central and noncentral distributions of likelihood ratio tests of hypotheses about linear restrictions on \( \beta \) and/or \( \xi \).

A model with a nonlinear covariance structure which allows for incomplete data in the case of regularly, consistently timed, possibly incomplete data was proposed by Rochon and Helms (1985 and 1989). They considered a general linear model in which the elements of \( \epsilon_i \) observe a given, stationary ARMA\((p,q)\) time series process where \( p \) and \( q \) are fixed known constants.

Using the ARMA notation the model for the error terms for the \( i \)-th individual may be written as:
\[ e_{i,t} - \phi_1 e_{i,t-1} - \cdots - \phi_p e_{i,t-p} = e_{i,t} - \theta_1 e_{i,t-1} - \cdots - \theta_q e_{i,t-q} \]
where \( \phi_p, \theta_q \neq 0 \) and \( t = 1, \ldots, n_i \). The error term at time \( t \) for the \( i \)-th individual, \( e_{i,t} \), is modeled as an explicit function of the \( p \) previous error terms for that individual (hence the autoregression) plus a weighted average of \( q \) 'random shocks'. These random shocks, \( \epsilon_{i,t} \), are referred to as white noise in the time
series literature and are independent and identically distributed where \( \xi_i = (\epsilon_{i,1}, \ldots, \epsilon_{i,n_i})' \) has mean 0 and variance \( \sigma^2 I_{n_i} \).

Rochon and Helms (1985) derived the likelihood equations and resulting maximum likelihood equations for these models and showed that under various regularity assumptions the maximum likelihood estimators are weakly consistent, asymptotically normal and asymptotically efficient. Also, \( \hat{\beta} \), the vector of estimated regression parameters, was shown to be asymptotically independent of the covariance parameter estimates.

In his review of the literature on the existing methods of analyzing repeated measures data, Louis (1988) briefly mentions a general linear model in which the covariance between the \( k \)-th and \( l \)-th observations on the \( i \)-th individual is \( \text{cov}(Y_{ik}, Y_{il}) = \sigma^2 \rho^{|t_{ik} - t_{il}|}, 0 \leq \rho < 1 \). When the data are equally spaced (regularly timed) this model reduces to a general linear model with AR(1) covariance structure.

Jones (1986) extended regression analysis with stationary errors to the case when observations are not regularly timed for one observation vector \( Y = X\beta + \xi \). The errors are modeled as a continuous-time autoregression, CAR(\( p \)), with observational error observed at arbitrary times. If the data are equally spaced this model reduces to a model with AR(\( p \)) covariance.

Diggle (1988) proposed a similar model to that of Louis (1988) which has parameters for variation between units in addition to parameters for modeling the serial correlation within units. Diggle's model is actually a reparametrization of the model proposed by Louis, with additional terms for variation between experimental units. Murray and Helms (1990) developed maximum likelihood and restricted maximum likelihood estimators for this model along with their asymptotic distributions. Two extensions of the model were also described.
Izenman and Williams (1989) explored the use of linear spectral models for an exploratory analysis of longitudinal data on growth in rats. The models are called ‘spectral’ models because they are based on the spectral representation of the response covariance matrix that is part of factor analysis. Spectral models are suggested as alternatives to either mechanistic or linear statistical models for the study of both random and nonrandom trends. The random part of these models is a factor analysis model where the metameters are factor loadings, and the coefficients of the random trends are common factor scores. Certain features of spectral models which make them particularly well suited for exploratory studies of longitudinal data are noted.

1.4 General Computing Algorithms

Seldom is the case where solutions to the maximum likelihood equations exist in closed form; iterative methods are usually required. There are several different methods available, and for certain applications, one method may be better than another.

1.4.1 The E-M Algorithm

Dempster, Laird, and Rubin (1977) presented a general approach to iterative computational methods for maximum likelihood estimates when the data are incomplete called the E-M algorithm. The basic idea of E-M is to view the actual data \( \mathbf{Y} \), as incomplete data from a larger system. An E-M algorithm has two steps in each iteration: an E-step and an M-step. The M-step performs the maximum likelihood estimation which would be carried out if the data were actually complete, but since the data are incomplete the maximization during the M-step is carried out on a “proxy” for the likelihood produced in the E-step. The
E-step finds the posterior expectation of the sufficient statistics for the complete data given the observed data \( \mathcal{Y} \) and given the current estimate of the parameters. The M-step then uses the expected sufficient statistics to produce the next iteration of the parameters. For the case of a general multivariate model analysis with missing data, the E-step is estimation of the missing data values given the current estimates of \( \mathcal{B} \) and \( \mathcal{S} \). The M-step is then ML estimation of \( \mathcal{B} \) and \( \mathcal{S} \) given the "complete" data of observed and estimated values.

The E-M algorithm is often favored because the ML equations based on the complete-data likelihood are often easy to obtain and in some cases the solutions to the M-step have closed form solutions. The literature, however, shows the effectiveness of the E-M algorithm to depend on the particular application. Fairclough and Helms (1984) found the E-M algorithm to require undesirably high numbers of iterations. Barton (1986), found the E-M algorithm approach to work well for his data.

Laird and Ware (1982) derived an E-M algorithm for computing ML and REML estimates in the general linear mixed model for \( \mathcal{Y}_i = \mathcal{X}_i \eta_i \) and unstructured \( \mathcal{D} \). Little and Schluchter (1985) obtained estimates for \( \mathcal{Q} \) in a model which contained both categorical and continuous response variables using the E-M approach. Jenrich and Schluchter (1986) describe an algorithm which combines scoring with generalized E-M for incomplete data models and large unstructured \( \mathcal{D} \). Their algorithm makes use of a separate "scoring step" which is used to estimate \( \mathcal{S} \) after \( \mathcal{Q} \) has been updated. Laird, Lange and Stram (1987) considered the use of the EM algorithm for both maximum likelihood and restricted maximum likelihood estimation in a general repeated measures setting using a multivariate normal data model with linear mean and mixed model covariance structure.
Dempster, Rubin and Tsutakawa (1981) describe and illustrate computational methods required for estimation of covariance components models. The estimation of fixed and random effects when the variance and covariances are known is presented in Bayesian terms. Estimation in three applications is done with the E-M algorithm, with computational details for special models presented.

1.4.2 Gradient Methods

Two of the best known gradient methods are the steepest ascent algorithm and the Newton-Raphson algorithm. Harville (1977), in reviewing ML estimation and some iterative methods, reported that although supported by convergence theorems, the rate of convergence of the steepest ascent method was often too slow.

The Newton-Raphson procedure utilizes second-order partial derivatives of the log-likelihood function and when supplied with good initial values (near the maximum), often converges in relatively few iterations. A similar procedure, the method of scoring, is identical to the Newton-Raphson method except that the expected values of the second-order partial derivatives are used in place of the actual derivatives. For a detailed description of gradient methods see Judge, Griffiths, Hill and Lee (1980).

Harville (1977) also discusses the use of the extended or modified Newton-Raphson procedure which addresses the problem of sensitivity of both methods to initial values. The modified N-R procedure attempts to improve the performance of the ordinary N-R procedure when supplied with poor initial estimates. Yet at the same time, it uses the good performance of the N-R procedure when it is started close to the maximum.
Improvements to the Newton-Raphson (NR), the method of scoring, the method of successive approximations (MSA) and the E-M algorithm are suggested by Callanan and Harville (1991) for the computation of restricted maximum likelihood estimates. This is achieved by reparametrizing the likelihood equations, i.e. to "linearize" the likelihood, before applying an iterative algorithm. The idea is to linearize the likelihood and obtain a new set of likelihood equations, which are equivalent to the original versions, but are "more nearly linear". It is hoped that these more nearly linear equations will converge more rapidly than the original ones. Based on numerical results this linearization seems to improve the MSA and NR algorithms and is better than the method of scoring and the EM algorithm.

1.5 Statement of the Problem and Outline

The data typically obtained from longitudinal studies contains missing data or misstimated data that require methods that can utilize a subject's available data. Analyses that utilize fully parametrized covariance structures, resulting in a proliferation of poorly estimated parameters, or methods that use case-wise deletion may not be particularly useful for these studies.

As discussed in Section 1.1, the use of structured covariance models is helpful in studies where the data are highly unbalanced or incomplete. Estimates of covariance parameters from unstructured covariance matrices may be difficult to obtain due to sparse data; they are also typically less efficient, i.e. have larger standard errors, than those obtained from structured covariances, where the number of parameters will be much smaller. Methods for analyses of longitudinal data with missing time points are available but have not been easy to use. In the past many of the models required writing one's own program or using a non-standardized program from some other investigator. These maximum likelihood
methods are easier to access with the 1990 version of BMDP-5V (and future releases from SAS) which offers direct access to some covariance structures and allows for others to be programmed by supplying FORTRAN subroutines.

The aim of this dissertation is to demonstrate the use of various structured covariance models through the analysis of longitudinal data. General suggestions about strategies for fitting these models are made and implemented during the analysis. Several popular covariance structures will be applied to the data and allowed to "compete" to see which ones fit the data best as assessed through likelihood ratio tests, evaluation of covariance parameters, and graphical methods. During this process we will note whether specific conditions present in the data lead to the favorable application of some covariance structures over others. For example, by noting conditions in the data such as the behavior of correlations between a subject's responses over time it may be possible to determine which covariance structures are most likely to fit the data before trying numerous different structures. At the end of the dissertation, a brief analysis paradigm will be presented and discussed as a possible guide for fitting the various covariance models and their fixed effects.

In Chapter 2 the structured covariance models to be used in the dissertation are described in detail along with the specific computing requirements for parameter estimation. The third chapter introduces the Lipids Research Clinic (LRC) data on repeated measures of total cholesterol and treats the data as a group of subjects being followed over time, without reference to baseline data or any specific clinical outcome. The structural covariances of Chapter 2 are applied to the data and comparisons are made. Special attention is also given to the fixed effects part of the model. Chapter 4 reanalyzes the same data, only this time baseline data are included and the data are analyzed in the spirit of a clinical trial. The issues of how to treat baseline values, as covariates or addition time
points, is the central theme of the chapter. In Chapter 5 an analysis paradigm is presented for using the models from this dissertation and some suggestions about further developments for standard software are made.
CHAPTER II
DESCRIPTION OF THE MODELS

2.1 The General Linear Model

2.1.1 Model Terminology and Equation

The general linear model equation (1.2.1) is the basic expression used to describe $Y_i$ as a linear function of covariates and design variables. The model equation for the $i$-th subject is

$$Y_i = X_i \beta + \varepsilon_i$$

where $Y_i$ is an $n_i \times 1$ vector of observations on the $i$-th subject, $i = 1, \ldots, K$, $\beta$ is a $p \times 1$ vector of unknown constant population parameters, $X_i$ is an $n_i \times p$ known constant design matrix, corresponding to the fixed effects $\beta$, and $\varepsilon_i$ is an $n_i \times 1$ vector of random error terms.

Note that $E[Y_i] = X_i \beta$ and $\text{Var}[Y_i] = \Sigma_i$. Also $\text{Var}[Y] = \text{Var}[Y_1 \parallel Y_1 \parallel \ldots \parallel Y_K] = \text{Diag} [\Sigma_i]$. See Section 1.2.2 for details of the general linear model.

2.1.2 Introduction to Structures for Covariance Matrices

Covariance matrices for the models which are used to describe repeated measures data come in different forms with different mathematical properties and different underlying assumptions. A covariance matrix, $\Sigma_i$, can be structured or
unstructured, and it can be linear or non-linear in its parameters.

The covariance matrix can be unstructured, in which case the dispersion matrix has \( t(t+1)/2 \) unique parameters, where \( t \) is the number of repeated outcome measures. Another term for this is 'fully parametrized', meaning that all possible parameters are estimated. For most models, especially when the number of repeated outcome measures is large or there are mistimed data, the number of covariance parameters for an unstructured covariance will become inflated. For example, if \( Y_i \) has 10 elements, \( \Sigma_i \) has 55 covariance parameters to be estimated. When the data are sparse due to missing or mis-timed data, this can lead to poor estimates of the covariance parameters and large standard errors. A favorable aspect of unstructured covariances is that no assumptions need to be made about them; they can take any form.

Structured covariances can be generally thought of as ones in which a trend is observed in the dispersion matrix and the covariances are modeled according to that trend. The "structure" of the covariance is imposed through a model, complete with a model equation and accompanying model assumptions. The information contained in the covariance matrix is modeled as a function of covariance parameters. This introduces the main drawback to modeled covariance structures: they are dependent upon a reduced number of parameters to accurately convey the information of the covariance matrix. The advantage, if the model assumptions hold, is a covariance described by a small number of parameters, say \( t' \) (where \( t' \ll t(t+1)/2 \)), with relatively small standard errors.

The mixed model with compound symmetry is an example of a structured covariance in which the variances are modeled by one parameter and covariances are modeled by a second, regardless of the number of repeated measures in the data. The AR(1) model also uses two parameters, \( \sigma^2 \) and \( \rho \), to model decreasing correlations of measures as time between measures increases. The assumptions of
a covariance model are often implicit through their model equations and must be met if the covariance model is to provide a good fit to the unstructured covariance matrix. The assumptions for some models can be informally assessed by inspection of the unstructured within-subject covariance matrix.\footnote{The methods to obtain the within-subject covariance and correlation matrices are described in Chapter 3.} For example, in the mixed model with compound symmetry, which has two covariance parameters, the variances of the unstructured covariance matrix must all appear equal to one value and the off-diagonal covariances must be approximately equal to another value. For the two parameter AR(1) model, the variances must again be equal and the covariances must decrease over time so that the trend can be modeled. Hence the term structured covariance comes from the nature of the structure imposed on the covariance matrix through a particular model. If the fitted covariance structure models the unstructured covariance matrix successfully, it is said to have ‘good fit’, analogous to ordinary regression models when referring to the expected value model. Recall that unstructured matrices can take any form or pattern; their forms are completely dependent on the data.

Covariance matrices can be linear or non-linear in the parameters that comprise them. Linear covariance structures were described in section 1.3.4 as models characterized by $\Sigma_i$ being linear in the parameters, that is, linear models in which

$$\Sigma_i = \sum_{i=1}^{m} \tau_i \mathcal{G}_i$$

for $n \times n$ symmetric matrices $\mathcal{G}_1, \ldots, \mathcal{G}_m$ whose elements are known. Mixed models have linear covariance structures, with compound symmetry being an important example.

Models with non-linear covariance structures cannot be described this way and are comprised of parameters which are non-linear in the parameters.
example is Murray's generalized model with curvature parameter \( \nu \), which will be described in section 2.3.

2.1.3 Introduction to Computing Aspects

The parameter estimates for these models are obtained using BMDP-5V (1990). Some of the models can be specified easily using built-in covariance structures. These do not require starting values, usually converge without problems, and do so quickly. The built-in structures include the AR(1) model, mixed model, compound symmetric model and the unstructured covariance.

For specialty models such as those described in 2.3, 2.4, and 2.5 for which there are not built-in structures, FORTRAN subroutines must be used to define the elements of \( \Sigma_i \) and the derivatives with respect to the parameters which define \( \Sigma_i \). Initial parameter estimates must be supplied for models utilizing FORTRAN subroutines. These models are less likely to converge than the built-in models, mainly due to their complexity and their dependence on user-supplied starting values. Details of computing aspects will be given for each model described in this chapter and also in later chapters and appendices.

2.2 The Mixed Model

2.2.1 Model Equation and Terminology

Recall the mixed model equation 1.2.2 from Section 1.2.3:

\[
Y_i = X_i \beta + Z_i a_i + \varepsilon_i
\]

where \( Y_i \) is an \( n_i \times 1 \) vector of observations on the \( i \)-th subject, \( i = 1, \ldots, k \),

\( \beta \) is a \( p \times 1 \) vector of unknown constant population parameters,

\( X_i \) is a \( n_i \times p \) known constant design matrix, corresponding to the fixed effects, \( \beta \),
\( \mathbf{d}_i \) is a \( q \times 1 \) vector of unknown, random individual parameters,

\( \mathbf{Z}_i \) is a \( n_i \times q \) known constant design matrix, corresponding to the random effects \( \mathbf{d}_i \), and

\( \varepsilon_i \) is an \( n_i \times 1 \) vector of random error terms.

Also see section 1.2 on Notation in Chapter 1.

2.2.2 Description of the Covariance Structure

One of the advantages of the mixed model is the ability to supply a design matrix for the individual covariates (or variables) which are represented as columns of \( \mathbf{Z}_i \). The covariance for person \( i \) is

\[
\Sigma_i = \mathbf{Z}_i \Delta \mathbf{Z}_i' + \sigma^2 \mathbf{Y},
\]

(2.2.1)

where \( \mathbf{Y} \) is a known constant matrix and often \( \mathbf{Y} = \mathbf{I} \). This covariance model has \( q(q + 1)/2 \) parameters in \( \Delta \) plus the parameter \( \sigma^2 \). The columns of \( \mathbf{Z}_i \) can be the same for all subjects (when all subjects have equal valued covariates) or they can vary individually for each subject. For example, one column might represent fixed values such as the age at which height measures were made on children during a study: that column of \( \mathbf{Z}_i = [8, 10, 12, 14]' \) for all \( i \). That is, this column of \( \mathbf{Z}_i \) is the same for all subjects if they were all measured at the same ages during the study.

When a model has individual effects that vary for each individual over time such as the average number of cigarettes smoked per day during a study period, the columns of \( \mathbf{Z}_i \) will vary across subjects. For example, the different columns of \( \mathbf{Z}_i \) might be: \( \mathbf{Z}_1 = [40, 38, 35, 20]' \), \( \mathbf{Z}_2 = [0, 0, 0, 0]' \), etc.

2.2.3 Computing Aspects for the Mixed Model

Estimates for the mixed model can be obtained from BMDP-5V using one
of the built-in covariance structure options. The elements of $\Sigma$ are obtained by
defining the individual random effects in $\mathcal{Z}_i$. Both types of covariates, those that
vary for individuals and those that do not, can be fit.

Consider a longitudinal study with five time points, say $t = 1$ through
$t = 5$. Then the complete data matrix $Y$ is $N \times 5$ and $\text{var}(Y) = \Sigma$ is block
diagonal with $\Sigma_i$ of dimension $5 \times 5$ on the diagonal. For the $i$-th person the
general framework for estimating the parameters of $\Sigma_i$ and their derivatives are
given below. The description of $\Sigma_i$ and its derivatives are for a simple mixed
model in which Time is the only covariate in the random effects parts of the
model. Note that in this example $\mathcal{Z}_i$ is constant across all subjects:

$$\text{Var}[Y_i] = \Sigma_i = \mathcal{Z}_i \Delta \mathcal{Z}_i' + \sigma^2 \mathcal{V}_i$$

$$= \mathcal{Z}_i \Delta \mathcal{Z}_i' + \sigma^2 \mathcal{I}_i$$, where $\mathcal{V}_i = \mathcal{I}_i$ for the simple case,

$$\mathcal{Z}_i = \begin{bmatrix}
1 & 1 \\
1 & 2 \\
1 & 3 \\
1 & 4 \\
1 & 5 \\
\end{bmatrix}$$, and $\Delta = \begin{bmatrix}
\delta_{11} & \delta_{12} \\
\delta_{21} & \delta_{22} \\
\end{bmatrix}$, $\delta_{12} = \delta_{21}$.

So $\Sigma_i =$

$$\begin{bmatrix}
1 & 1 \\
1 & 2 \\
1 & 3 \\
1 & 4 \\
1 & 5 \\
\end{bmatrix} \begin{bmatrix}
\delta_{11} & \delta_{12} \\
\delta_{21} & \delta_{22} \\
\end{bmatrix} \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 2 & 3 & 4 & 5 \\
\end{bmatrix} + \sigma^2 \mathcal{I}_5$$.
Multiplying and adding terms gives \( \Sigma_i = \)

\[
\begin{bmatrix}
\delta_{11} + 2\delta_{12} + \delta_{22} + \sigma_e^2 & \delta_{11} + 3\delta_{12} + 2\delta_{22} + \sigma_e^2 & \delta_{11} + 4\delta_{12} + 3\delta_{22} + \sigma_e^2 \\
\delta_{11} + 4\delta_{12} + 4\delta_{22} + \sigma_e^2 & \delta_{11} + 5\delta_{12} + 6\delta_{22} + \sigma_e^2 & \delta_{11} + 6\delta_{12} + 9\delta_{22} + \sigma_e^2 \\
\end{bmatrix}
\]

(sym)

\[
\begin{bmatrix}
\delta_{11} + 5\delta_{12} + 4\delta_{22} + \sigma_e^2 & \delta_{11} + 6\delta_{12} + 5\delta_{22} + \sigma_e^2 \\
\delta_{11} + 6\delta_{12} + 8\delta_{22} + \sigma_e^2 & \delta_{11} + 7\delta_{12} + 10\delta_{22} + \sigma_e^2 \\
\delta_{11} + 7\delta_{12} + 12\delta_{22} + \sigma_e^2 & \delta_{11} + 8\delta_{12} + 15\delta_{22} + \sigma_e^2 \\
\delta_{11} + 8\delta_{12} + 16\delta_{22} + \sigma_e^2 & \delta_{11} + 9\delta_{12} + 20\delta_{22} + \sigma_e^2 \\
\delta_{11} + 10\delta_{12} + 25\delta_{22} + \sigma_e^2 & \end{bmatrix}
\]

(Note: matrix is split after third column).

The derivatives are:

\[
\frac{\delta \Sigma_i}{\delta \delta_{11}} = I_5 , \quad \frac{\delta \Sigma_i}{\delta \sigma_e^2} = I_5 ,
\]

\[
\frac{\delta \Sigma_i}{\delta \delta_{12}} = \frac{\delta \Sigma_i}{\delta \delta_{21}} = \begin{bmatrix}
2 & 3 & 4 & 5 & 6 \\
4 & 5 & 6 & 7 \\
6 & 7 & 8 \\
8 & 9 \\
10
\end{bmatrix}, \text{ and}
\]

\[
\frac{\delta \Sigma_i}{\delta \delta_{22}} = \begin{bmatrix}
1 & 2 & 3 & 4 & 5 \\
4 & 6 & 8 & 10 \\
9 & 10 & 15 \\
16 & 20 \\
25
\end{bmatrix}
\]

29
2.3 Models with Stationary Variance and Exponentially Decreasing Correlation

2.3.1 Model Terminology and Equation

Louis investigated a statistical model with stationary variance and exponentially decreasing correlation. The model equation for the observations from the \(i\)-th subject is

\[
Y_i = X_i \beta + \xi_i + \varepsilon_i \tag{2.3.1}
\]

where \(Y_i\) is vector of responses for the \(i\)-th person, \(i = 1, ..., K\),
\(X_i\) is a full rank \(n_i \times p\) known, constant design matrix for fixed effects,
\(\beta_i\) is \(p \times 1\) vector of unknown fixed effect population parameters,
\(\varepsilon_i\) is an \(n_i \times 1\) vector of random errors, and
\(\xi_i\) is the \(n_i \times 1\) vector of observation times for the \(i\)-th subject.

Borrowing ideas from time series concepts, Louis discussed covariance structures from the AR(\(p\)) class of models and Louis and Shapiro (1986) developed SAS macros for fitting the AR(1) model,

\[
cov(y_{ij}, y_{ik}) = \sigma^2 \rho^{|t_{ij} - t_{ik}|}, \quad 0 \leq \rho < 1, \tag{2.3.2}
\]

where \(cov(y_{ij}, y_{ik})\) represents the \(j\)-th and \(k\)-th element of \(\Sigma_i\). For non-negative \(\rho\), the unit of time does not need to be integral spaced and the follow-up time intervals can vary. As do most other covariance structures, this model requires iterative methods to estimate the parameters.

2.3.2 Description of the Covariance Structure

The above model was generalized by Murray and Helms\(^2\) (1990) to include a curvature parameter which models the strength and pattern of the correlations over time. Murray's statistical model for the \(i\)-th person is equation (2.3.1) and

\(^2\) For brevity this model will be referred to as Murray's model hereinafter.
the covariance model is
\[ \text{cov}(y_{ij}, y_{ik}) = \sigma^2 \rho (|t_{ij} - t_{ik}|)^\nu, \quad 0 \leq \rho < 1, \nu \geq 0, \] (2.3.3)

where \( \nu \) is a "curvature parameter". It is clear from this model that equation 2.3.2 is a special case of the more general model when \( \nu = 1 \). The parameter \( \nu \) influences the shape of the correlation curve: values of \( \nu > 1 \) indicate that the correlation among observations fall off quickly as time between measures increases. Values of \( \nu < 1 \) indicate the opposite, that the correlations are more steady across time. As \( \nu \to 0 \), the correlations are near constant and the model approaches the familiar uniform correlation model (or compound symmetry). Values of \( \nu < 0 \), which make little sense, are not considered.

If the timing of the observations is one unit apart and \( \nu = 1 \), Murray's model reduces to the general linear model with AR(1) covariance structure. This can also be viewed as a special case of the model considered by Rochon and Helms (1985 & 1989). Once the timing becomes irregular, however, the model no longer fits into the general scheme of Rochon and Helms. The major advantage of Murray's model, when compared to the model of Rochon and Helms, is that irregularly timed data can be accommodated. When compared to Louis' model, Murray's model can accommodate correlations which decrease more or less swiftly than the special case when \( \nu = 1 \), making it more flexible.

### 2.3.3 Computing Aspects for these Models

Estimates of parameters for this model can be obtained from BMDP-5V using a FORTRAN subroutine which is linked to the program and called once for each iteration. The elements of \( \Sigma \) must be defined along with the derivatives with respect to each of the covariance parameters. There are two strategies for fitting this covariance model: (1) All three parameters, \( \nu, \sigma^2, \) and \( \rho \), are estimated simultaneously and (2) \( \nu \) is set to a constant and \( \sigma^2 \) and \( \rho \) are estimated for that
value of $\nu$, and $\nu$ is varied manually. The reason for two strategies is that the 
main model with $\nu$ estimated as a parameter often does not converge. However, 
the model with fixed $\nu$ and parameters $\sigma^2$ and $\rho$ rarely fails to converge. 

Consider the same longitudinal study with five regularly spaced time 
points, $t = 1$ through $t = 5$. Again, the complete data matrix $\mathbf{Y}$ is $N \times 5$ and 
$\text{var}(\mathbf{Y}) = \mathbf{\Sigma}$ is block diagonal with $\mathbf{\Sigma}_i$ of dimension $5 \times 5$ on the diagonal. For the 
i-th person the general framework for estimating the parameters of $\mathbf{\Sigma}_i$ and their 
derivatives when $\nu$ is a constant is given below. The covariance of the two 
measurements for 

person $i$ (omitting subscript $i$) is 

$$
cov(y_j, y_k) = \sigma^2 \rho^{\nu (|t_j - t_k|)} \nu, \quad 0 \leq \rho < 1, \nu \geq 0, 
$$

(2.3.4) 

where $cov(y_j, y_k)$ represents the $j$-th row and $k$-th column of $\mathbf{\Sigma}_i$ with the 
curvature parameter $\nu = c$, a constant. Then the $5 \times 5$ covariance matrix is 

$$
\mathbf{\Sigma}_i = 
\begin{bmatrix}
\sigma^2 \rho^{(|t_1 - t_1|)} & \sigma^2 \rho^{(|t_1 - t_2|)} & \sigma^2 \rho^{(|t_1 - t_3|)} & \sigma^2 \rho^{(|t_1 - t_4|)} & \sigma^2 \rho^{(|t_1 - t_5|)} \\
\sigma^2 \rho^{(|t_2 - t_1|)} & \sigma^2 \rho^{(|t_2 - t_2|)} & \sigma^2 \rho^{(|t_2 - t_3|)} & \sigma^2 \rho^{(|t_2 - t_4|)} & \sigma^2 \rho^{(|t_2 - t_5|)} \\
\sigma^2 \rho^{(|t_3 - t_1|)} & \sigma^2 \rho^{(|t_3 - t_2|)} & \sigma^2 \rho^{(|t_3 - t_3|)} & \sigma^2 \rho^{(|t_3 - t_4|)} & \sigma^2 \rho^{(|t_3 - t_5|)} \\
\sigma^2 \rho^{(|t_4 - t_1|)} & \sigma^2 \rho^{(|t_4 - t_2|)} & \sigma^2 \rho^{(|t_4 - t_3|)} & \sigma^2 \rho^{(|t_4 - t_4|)} & \sigma^2 \rho^{(|t_4 - t_5|)} \\
\text{(sym)} & \text{(sym)} & \text{(sym)} & \text{(sym)} & \text{(sym)} 
\end{bmatrix}
$$

(note: $\rho^{X^\nu} = \rho^{(X^\nu)}$ for these models)
\[
\begin{bmatrix}
\sigma^2 & \sigma^2 \rho & \sigma^2 \rho^2c & \sigma^2 \rho^3c & \sigma^2 \rho^4c \\
\sigma^2 & \sigma^2 \rho & \sigma^2 \rho^2c & \sigma^2 \rho^3c & \sigma^2 \rho^4c \\
\sigma^2 & \sigma^2 \rho & \sigma^2 \rho^2c & \sigma^2 \rho^3c & \sigma^2 \rho^4c \\
\sigma^2 & \sigma^2 \rho & \sigma^2 \rho^2c & \sigma^2 \rho^3c & \sigma^2 \rho^4c \\
\sigma^2 & \sigma^2 \rho & \sigma^2 \rho^2c & \sigma^2 \rho^3c & \sigma^2 \rho^4c \\
\end{bmatrix}
\]

(sym)

\[
\begin{bmatrix}
1 & \rho & \rho^2c & \rho^3c & \rho^4c \\
1 & \rho & \rho^2c & \rho^3c & \\
1 & \rho & \rho^2c & \\
1 & \rho & \\
1 & \\
\end{bmatrix}
\]

\[
= \sigma^2
\begin{bmatrix}
1 & \rho & \rho^2c & \rho^3c & \rho^4c \\
1 & \rho & \rho^2c & \rho^3c & \\
1 & \rho & \rho^2c & \\
1 & \rho & \\
1 & \\
\end{bmatrix}
\]

(sym)

The derivatives for this model with \( \nu = c \) are:

\[
\frac{\partial \Sigma_i}{\partial \sigma^2} = \begin{bmatrix}
1 & \rho & \rho^2c & \rho^3c & \rho^4c \\
1 & \rho & \rho^2c & \rho^3c & \\
1 & \rho & \rho^2c & \\
1 & \rho & \\
1 & \\
\end{bmatrix}, \text{ and}
\]

(sym)

\[
\begin{bmatrix}
0 & \sigma^2 & \sigma^2(2c)\rho(2c-1) & \sigma^2(3c)\rho(3c-1) & \sigma^2(4c)\rho(4c-1) \\
0 & \sigma^2 & \sigma^2(2c)\rho(2c-1) & \sigma^2(3c)\rho(3c-1) & \\
0 & \sigma^2 & \sigma^2(2c)\rho(2c-1) & \\
0 & \sigma^2 & \\
0 & \\
\end{bmatrix}
\]

(sym)

\[
\frac{\partial \Sigma_i}{\partial \rho} = \begin{bmatrix}
0 & \sigma^2 & \sigma^2(2c)\rho(2c-1) & \\
0 & \sigma^2 & \sigma^2(2c)\rho(2c-1) & \\
0 & \sigma^2 & \\
0 & \sigma^2 & \\
0 & \\
\end{bmatrix}
\]

(sym)

\[
\begin{bmatrix}
0 & \\
0 & \\
0 & \\
0 & \\
0 & \\
\end{bmatrix}
\]

33
When the curvature parameter $\nu = 1$, the model reduces to the AR(1) model. The covariance matrix and derivative are:

$$
\Sigma_i = \begin{bmatrix}
\sigma^2 & \sigma^2 \rho & \sigma^2 \rho^2 & \sigma^2 \rho^3 & \sigma^2 \rho^4 \\
\sigma^2 & \sigma^2 \rho & \sigma^2 \rho^2 & \sigma^2 \rho^3 \\
\sigma^2 & \sigma^2 \rho & \sigma^2 \rho^2 & \sigma^2 \rho^3 \\
\sigma^2 & \sigma^2 \rho & \sigma^2 \rho^2 & \sigma^2 \\
\sigma^2
\end{bmatrix},
$$

$$
\frac{\partial \Sigma_i}{\partial \sigma^2} = \begin{bmatrix}
1 & \rho & \rho^{2c} & \rho^{3c} & \rho^{4c} \\
1 & \rho & \rho^{2c} & \rho^{3c} \\
1 & \rho & \rho^{2c} \\
1 & \rho \\
1
\end{bmatrix}, \text{ and}
$$

$$
\frac{\partial \Sigma_i}{\partial \rho} = \begin{bmatrix}
0 & \sigma^2 & 2\sigma^2 \rho & 3\sigma^2 \rho^2 & 4\sigma^2 \rho^3 \\
0 & \sigma^2 & 2\sigma^2 \rho & 3\sigma^2 \rho^2 \\
0 & \sigma^2 & 2\sigma^2 \rho \\
0 & \sigma^2 \\
0
\end{bmatrix}.
$$
For models in which $\nu$ is estimated the derivative matrix $\frac{\partial \Sigma_i}{\partial \nu}$ must also be evaluated. If the data are regularly timed integers, as in this example, the resulting derivative matrix is:

\[
\frac{\partial \Sigma_i}{\partial \nu} = \begin{bmatrix}
0 & 0 & \sigma^2 \rho^{2\nu} \log \rho \cdot 2^\nu \cdot \log(2)
& \sigma^2 \rho^{3\nu} \log \rho \cdot 3^\nu \cdot \log(3)
& \sigma^2 \rho^{4\nu} \log \rho \cdot 4^\nu \cdot \log(4)
0 & 0 & \sigma^2 \rho^{2\nu} \log \rho \cdot 2^\nu \cdot \log(2)
& \sigma^2 \rho^{3\nu} \log \rho \cdot 3^\nu \cdot \log(3)
& \\
0 & 0 & \sigma^2 \rho^{2\nu} \log \rho \cdot 2^\nu \cdot \log(2)
& 0 & 0
\end{bmatrix}
\]

The diagonal elements and first off-diagonal elements are not functions of $\nu$ when the data are regularly timed with integer values 1, 2, 3, ..., 10. If the data are irregularly timed the first off-diagonal element will not be equal to zero because the element $\sigma^2 \rho^{2\nu} \log \rho \cdot 2^\nu \cdot \log(1)$ will not always contain the term $\log(1)$. For example, if the data collection has timing in years of [0.2, 0.4, 0.6, 1, 2, 3], the possible first off-diagonal elements will contain terms $\log(0.2)$, $\log(0.4)$ and $\log(1)$, in which case only some of the elements will equal zero. If the data are irregularly timed or are not consecutive integers (e.g. 6, 12, 18, 24), the first off-diagonal elements must be evaluated. If the data are regularly timed but are in a form such as [6, 12, 18, 24] months, then the derivatives must be evaluated for the first off-diagonal elements or the time values must be transformed to functions of 6-six month intervals such as 1, 2, 3, and 4.
The derivatives with respect to \( \sigma^2 \) and \( \rho \) remain the same with \( \nu \) replacing the constant:

\[
\frac{\delta \Sigma_i}{\delta \sigma^2} = \begin{bmatrix}
1 & \rho & \rho^{2\nu} & \rho^{3\nu} & \rho^{4\nu} \\
1 & \rho & \rho^{2\nu} & \rho^{3\nu} \\
1 & \rho & \rho^{2\nu} \\
1 & \rho \\
\text{(sym)} & 1
\end{bmatrix}, \text{ and}
\]

\[
\frac{\delta \Sigma_i}{\delta \rho} = \begin{bmatrix}
0 & \sigma^2 & \sigma^2(2\nu)\rho(2\nu-1) & \sigma^2(3\nu)\rho(3\nu-1) & \sigma^2(4\nu)\rho(4\nu-1) \\
0 & \sigma^2 & \sigma^2(2\nu)\rho(2\nu-1) & \sigma^2(3\nu)\rho(3\nu-1) \\
0 & \sigma^2 & \sigma^2(2\nu)\rho(2\nu-1) \\
0 & \sigma^2 \\
\text{(sym)} & 0
\end{bmatrix}.
\]
2.4 Mixed Model with Compound Symmetry and an Autoregressive Error

2.4.1 Model Equation and Terminology

Diggle (1988) describes his model as a linear model for repeated measures in which the correlation structure within each time sequence of measures includes parameters for measurement error, variation between subjects (experimental units) and serial correlation within units.

This approach assumes a sequence of measurements has been made on each experimental unit, that interest lies mainly in comparing the mean response profiles across treatments and at different times and that the mean response depends on treatment and time.

Desirable features for a general method of analysis for repeated measures data like those described include the following conditions mentioned in the article:

1. The model must be flexible enough to fit time trends within treatment groups and differences in trends between treatment groups.

2. Specification of the covariance structure within each time sequence should be flexible, but overspecification should be avoided. While over-parametrization of the covariance structure can lead to inefficient estimation and poor standard errors of the estimates of the mean responses, underparametrization can lead to invalid inferences about the mean response profile when the assumed covariance structure does not fit.

3. The method ought to be capable of handling irregularly timed data. This includes the ability to keep experimental units which are prematurely terminated in the analysis.

4. The analysis method should have some form of diagnostic to allow at least an informal assessment of goodness of fit to the assumed covariance structure.

Notation for the model:

Let there be $k$ subjects (or experimental units) each with with $n_i$ measurements on the $i$-th subject where $Y_{ij}$ denotes the $j$-th measure on the $i$-th subject.
The following assumptions are made:
\[ Y = (Y_1', \ldots, Y_K') \] is the entire matrix of observed data and has a multivariate normal distribution: \( Y \sim \text{GLMM}(\mu, \Sigma) \).

\[ Y_i = (Y_{i1}, \ldots, Y_{iN_i})' \] is the vector of \( n_i \) measurements for person \( i \).
Also, \( \mu_i = E[Y_i] \), and \( \mu = E[Y] \).

For the mean response vector \( \mu \), assume a linear model exists such that:
\[ \mu = X\beta, \] where design matrix \( X \) has dimension \( (N_i \times q) \) and \( \beta \) is \( (q \times 1) \),
\[ \Sigma_i = \text{var}(Y_i), \] and
\( \Sigma \) is block diagonal with \( \Sigma_i \) on the the diagonal.

### 2.4.2 Description of the Covariance Structure

Diggle suggests that \( \Sigma_i \) should accommodate certain features which are commonly found in repeated measures data. First, repeated measurements close in time are often not correlated perfectly, such as when subsampling of material is performed on the same unit. Therefore, a good model should account for measurement error. Second, average typical responses will vary randomly between units, some being high on average, some low, creating a positive correlation between any two measurements on the same unit. Third, the correlation between measurements on the same unit usually depends on the length of time between the two measures, with correlations typically decreasing as time separation increases.

According to Diggle a model that incorporates all of the above is
\[ Y_{ij} = \mu_{ij} + Z_{ij} + V_i + W_i(t_{ij}) \]
for the j-th measure on the i-th person where

- \( Z_{ij} \) is i.i.d. \( N(0, \tau^2) \) for subsampling variation within units,
- \( U_i \) is i.i.d. \( N(0, \omega^2) \) for variation between units, and
\(W_i(t_{ij})\) are independent stationary Gaussian processes, with \(E[W_i(t_{ij})] = 0\), and

\[
\text{cov}\{W_i(t_j), W_i(t_k)\} = \sigma^2 \rho(|t_j - t_k|)
\]

where function \(\rho(|t_j - t_k|) = \exp(-\alpha u), u = |t_j - t_k|, \alpha = -\ln(\rho)\).

The variance matrix\(^3\) is

\[
\Sigma_i = \tau_i^2 \mathbb{I} + \omega_i^2 \mathbb{I}', \quad \xi_i = (\xi_{i1}, \ldots, \xi_{iN_i})'.
\]

(2.4.1)

Note: if \(\nu = 1\) we have an AR-1 type of covariance structure.

if \(\sigma^2 = 0\) the model reduces to a uniform correlation model (compound symmetry).

if \(\sigma^2 = 0\) and \(\omega^2 = 0\) the model becomes the classical GLUM model.

2.4.3 Computing aspects for these Models

Estimates of parameters for this model can be obtained from BMDP-5V using a FORTRAN subroutine which is linked to the program and called once for each iteration. The elements of \(\Sigma_i\) must be defined along with the derivatives with respect to each of the covariance parameters. This covariance model has four parameters: \(\tau^2, \omega^2, \sigma^2\) and \(\alpha\). Diggle does not estimate the parameter\(^4\) \(\nu\); it is set to an integer constant such as 1 or 2.

Again, consider a longitudinal study with five time points, \(t = 1\) through \(t = 5\). For the \(i\)-th person the general framework for estimating the parameters of \(\Sigma_i\) and their derivatives when \(\nu\) is a constant is given below. The covariance for person \(i\) (omitting subscript \(i\)) is

\(\mathbb{I}_k\) will be used to indicate \(\mathbb{I}_k \mathbb{I}_k'\) hereinafter.

\(\text{This is possibly due to convergence problems for the covariance structure when } \nu \text{ is estimated as a parameter in the model, although no explanation is given.}\)
$$\Sigma_i = \tau^2 I + \omega^2 J + \sigma^2 \mathcal{R}(t_i),$$

(2.4.2)

where the j-th row and k-th column element of $\sigma^2 \mathcal{R}(t_i) = \text{cov}(e_j, e_k) = \sigma^2 \exp(-\alpha |t_j - t_k|^\nu), 0 \leq \alpha$, and the curvature parameter $\nu$ is set to a constant, say $c$, equal to 1 or 2. The model parameters are

$$\mathcal{R}(t_k) = \sigma^2 \begin{bmatrix}
    e^{-\alpha(t_1 - t_1)} e^{-\alpha(t_1 - t_2)} e^{-\alpha(t_1 - t_3)} e^{-\alpha(t_1 - t_4)} e^{-\alpha(t_1 - t_5)} \\
    e^{-\alpha(t_2 - t_2)} e^{-\alpha(t_2 - t_3)} e^{-\alpha(t_2 - t_4)} e^{-\alpha(t_2 - t_5)} \\
    e^{-\alpha(t_3 - t_3)} e^{-\alpha(t_3 - t_4)} e^{-\alpha(t_3 - t_5)} \\
    e^{-\alpha(t_4 - t_4)} e^{-\alpha(t_4 - t_5)} \\
    (\text{sym}) \\
    e^{-\alpha(t_5 - t_5)}
\end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix}
    1 & e^{-\alpha} & e^{-\alpha^2} & e^{-\alpha^3} & e^{-\alpha^4} \\
    1 & e^{-\alpha} & e^{-\alpha^2} & e^{-\alpha^3} & e^{-\alpha^4} \\
    1 & e^{-\alpha} & e^{-\alpha^2} \\
    1 & e^{-\alpha} \\
    (\text{sym}) \\
    1
\end{bmatrix}.$$
The derivatives are

\[ \frac{\partial \Sigma_i}{\partial \tau^2} = \mathbb{I}_5, \quad \frac{\partial \Sigma_i}{\partial \omega^2} = \mathbb{I}_5, \]

\[ \frac{\partial \Sigma_i}{\partial \alpha^2} = \begin{bmatrix} 1 & e^{-\alpha} & e^{-\alpha^2} & e^{-\alpha^3} & e^{-\alpha^4} \\ 1 & e^{-\alpha} & e^{-\alpha^2} & e^{-\alpha^3} \\ 1 & e^{-\alpha} & e^{-\alpha^2} \\ 1 & e^{-\alpha} \\ \text{(sym)} \end{bmatrix}, \text{ and} \]

\[ \frac{\partial \Sigma_i}{\partial \sigma^2} = \begin{bmatrix} 0 & -\sigma^2 e^{-\alpha} & -2c^2 \sigma^2 e^{-\alpha^2} & -3c^2 \sigma^2 e^{-\alpha^3} & -4c^2 \sigma^2 e^{-\alpha^4} \\ 0 & -\sigma^2 e^{-\alpha} & -2c^2 \sigma^2 e^{-\alpha^2} & -3c^2 \sigma^2 e^{-\alpha^3} \\ 0 & -\sigma^2 e^{-\alpha} & -2c^2 \sigma^2 e^{-\alpha^2} \\ 0 & -\sigma^2 e^{-\alpha} \\ \text{(sym)} \end{bmatrix}. \]
The parameters for Diggle's model when \( \nu = 2 \) are

\[
\Sigma_i = \tau^2 \mathbb{I} + \omega^2 \mathbb{J} + \sigma^2 \mathbb{R}(t_i)
\]

\[
= \tau^2 \mathbb{I} + \omega^2 \mathbb{J} + \sigma^2 \begin{bmatrix}
1 & e^{-\alpha} & e^{-4\alpha} & e^{-9\alpha} & e^{-16\alpha} \\
1 & e^{-\alpha} & e^{-4\alpha} & e^{-9\alpha} & e^{-16\alpha} \\
1 & e^{-\alpha} & e^{-4\alpha} & e^{-9\alpha} & e^{-16\alpha} \\
1 & e^{-\alpha} & e^{-4\alpha} & e^{-9\alpha} & e^{-16\alpha} \\
1 & e^{-\alpha} & e^{-4\alpha} & e^{-9\alpha} & e^{-16\alpha}
\end{bmatrix}.
\]

The derivatives are

\[
\frac{\partial \Sigma_i}{\partial \tau^2} = \mathbb{I}_5 , \quad \frac{\partial \Sigma_i}{\partial \omega^2} = \mathbb{I}_5 ,
\]

\[
\frac{\partial \Sigma_i}{\partial \sigma^2} = \begin{bmatrix}
1 & e^{-\alpha} & e^{-4\alpha} & e^{-9\alpha} & e^{-16\alpha} \\
1 & e^{-\alpha} & e^{-4\alpha} & e^{-9\alpha} & e^{-16\alpha} \\
1 & e^{-\alpha} & e^{-4\alpha} & e^{-9\alpha} & e^{-16\alpha} \\
1 & e^{-\alpha} & e^{-4\alpha} & e^{-9\alpha} & e^{-16\alpha} \\
1 & e^{-\alpha} & e^{-4\alpha} & e^{-9\alpha} & e^{-16\alpha}
\end{bmatrix}, \text{ and}
\]

\[
\frac{\partial \Sigma_i}{\partial \alpha} = \begin{bmatrix}
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha}
\end{bmatrix}.
\]

\[
\frac{\partial \Sigma_i}{\partial \alpha} = \begin{bmatrix}
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha}
\end{bmatrix}.
\]

\[
\frac{\partial \Sigma_i}{\partial \alpha} = \begin{bmatrix}
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -4\sigma^2 e^{-4\alpha} & -9\sigma^2 e^{-9\alpha} & -16\sigma^2 e^{-16\alpha}
\end{bmatrix}.
\]
The parameters for Diggle’s model when \( \nu = 1 \) are

\[
\Sigma_i = \tau^2 I + \omega^2 J + \sigma^2 R(x_i)
\]

\[
= \tau^2 I + \omega^2 J + \sigma^2 \begin{bmatrix}
1 & e^{-\alpha} & e^{-2\alpha} & e^{-3\alpha} & e^{-9\alpha} \\
1 & e^{-\alpha} & e^{-2\alpha} & e^{-3\alpha} \\
1 & e^{-\alpha} & e^{-2\alpha} \\
1 & e^{-\alpha} \\
(sym) & 1
\end{bmatrix}.
\]

The derivatives are

\[
\frac{\partial \Sigma_i}{\partial \tau^2} = I_5,
\quad \frac{\partial \Sigma_i}{\partial \omega^2} = I_5,
\]

\[
\frac{\partial \Sigma_i}{\partial \sigma^2} = \begin{bmatrix}
1 & e^{-\alpha} & e^{-2\alpha} & e^{-3\alpha} & e^{-9\alpha} \\
1 & e^{-\alpha} & e^{-2\alpha} & e^{-3\alpha} \\
1 & e^{-\alpha} & e^{-2\alpha} \\
1 & e^{-\alpha} \\
(sym) & 1
\end{bmatrix}, \text{ and}
\]

\[
\frac{\partial \Sigma_i}{\partial \alpha} = \begin{bmatrix}
0 & -\sigma^2 e^{-\alpha} & -2\sigma^2 e^{-2\alpha} & -3\sigma^2 e^{-3\alpha} & -4\sigma^2 e^{-4\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -2\sigma^2 e^{-2\alpha} & -3\sigma^2 e^{-3\alpha} \\
0 & -\sigma^2 e^{-\alpha} & -2\sigma^2 e^{-2\alpha} \\
0 & -\sigma^2 e^{-\alpha} \\
(sym) & 0
\end{bmatrix}.
\]
2.5 Mixed Model with Random Effects and and Autoregressive Error

2.5.1 Model Terminology and Equation See equation 2.3.1 for model description.

2.5.2 Description of the Covariance Structure

Another covariance model that can be fit involves a combination of the mixed model and Murray's generalization of the AR-1 model. The general linear model (1.2.1) is used to describe the fixed effects of the model. The covariance structure takes the form:

$$\Sigma_i = Z_i \Delta Z_i' + \sigma^2 \mathbb{R}(\bar{t}_i),$$  \hspace{1cm} (2.5.1)

where the $j$-th row and $k$-th column element of $\sigma^2 \mathbb{R}(\bar{t}_i) = \text{cov}(e_j, e_k) = \sigma^2 \rho^{|t_j - t_k|^{\nu}}, 0 \leq \rho < 1$, with 'curvature' parameter $\nu \geq 0$. The first part of the covariance model, $Z_i \Delta Z_i'$, is directly from a mixed model and involves random effects in $Z_i$ and $\Delta$. From Murray's generalized model come the parameters $\rho$, $\sigma^2$, and $\nu$.

This model incorporates important features from two different covariance structures. From the mixed model we have the ability to fit data when there is a large between subject variation in $\mathbf{Y}$. From Murray's model we get further flexibility for modeling exponentially decreasing correlation in $\mathbf{Y}$.

2.5.3 Computing aspects for these Models

Estimates for this combined model can be obtained from BMDP-5V using a FORTRAN subroutine which is linked to the program and called once for each iteration. The elements of $\Sigma$ must be defined along with the derivatives with respect to the $q$ covariance parameters. The covariance for person $i$ is

$$\Sigma_i = Z_i \Delta Z_i' + \sigma^2 \mathbb{R}(\bar{t}_i),$$

where the $j$-th row and $k$-th column element of $\sigma^2 \mathbb{R}(\bar{t}_i) = \text{cov}(e_j, e_k) = \sigma^2 \rho^{|t_j - t_k|^{\nu}}, 0 \leq \rho < 1$, with the 'curvature' parameter $\nu \geq 0$. This covariance
model has \( q(q+1)/2 \) parameters in \( \Delta \) contributed from the mixed model part of the covariance structure, plus parameters \( \rho, \sigma^2 \) and \( \nu \) from Murray's model. Due to convergence problems for models in which \( \nu \) is an unknown parameter, \( \nu \) is usually set to a constant such as 1 or 2. Also, only \( \Xi \) matrices which are the same for all subjects can be fit using BMDP-5V; that is, covariates that have different values for different individuals cannot be added to \( \Xi \).

Again, consider the same hypothetical longitudinal study with five time points, 1 through 5. If Time is the only covariate for \( \Xi \) then

\[
\text{Var}[Y_i] = \Sigma_i = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} + \sigma^2_e + \begin{bmatrix} 1 & \rho^{1\nu} & \rho^{2\nu} & \rho^{3\nu} & \rho^{4\nu} \\ 1 & \rho^{1\nu} & \rho^{2\nu} & \rho^{3\nu} \\ 1 & \rho^{1\nu} & \rho^{2\nu} \\ 1 & \rho^{1\nu} \end{bmatrix} + \begin{bmatrix} \text{(sym)} \\ 1 \end{bmatrix},
\]

(note: \( \rho^{X\nu} = \rho^{(X\nu)} \) for these models)

which when multiplied out and added is equal to:
Note: matrix is split after third column.

\[
\begin{bmatrix}
\delta_{11} + 2\delta_{12} + \delta_{22} + \sigma_{\epsilon}^2 & \delta_{11} + 3\delta_{12} + 2\delta_{22} + \rho\sigma_{\epsilon}^2 & \delta_{11} + 4\delta_{12} + 3\delta_{22} + \rho^2\sigma_{\epsilon}^2 \\
\delta_{11} + 4\delta_{12} + 4\delta_{22} + \sigma_{\epsilon}^2 & \delta_{11} + 5\delta_{12} + 6\delta_{22} + \rho\sigma_{\epsilon}^2 \\
\delta_{11} + 6\delta_{12} + 9\delta_{22} + \sigma_{\epsilon}^2 & & \\
\end{bmatrix}
\]

\[(\text{sym})\]

\[
\begin{bmatrix}
\delta_{11} + 5\delta_{12} + 4\delta_{22} + \rho^2\sigma_{\epsilon}^2 & \delta_{11} + 6\delta_{12} + 5\delta_{22} + \rho^4\sigma_{\epsilon}^2 \\
\delta_{11} + 6\delta_{12} + 8\delta_{22} + \rho^2\sigma_{\epsilon}^2 & \delta_{11} + 7\delta_{12} + 10\delta_{22} + \rho^3\sigma_{\epsilon}^2 \\
\delta_{11} + 7\delta_{12} + 12\delta_{22} + \rho\sigma_{\epsilon}^2 & \delta_{11} + 8\delta_{12} + 15\delta_{22} + \rho^2\sigma_{\epsilon}^2 \\
\delta_{11} + 8\delta_{12} + 16\delta_{22} + \sigma_{\epsilon}^2 & \delta_{11} + 9\delta_{12} + 20\delta_{22} + \rho\sigma_{\epsilon}^2 \\
\delta_{11} + 10\delta_{12} + 25\delta_{22} + \sigma_{\epsilon}^2 & & \\
\end{bmatrix}
\]

The derivatives are:

\[
\frac{\partial \Sigma_i}{\partial \delta_{11}} = J_5, \quad \frac{\partial \Sigma_i}{\partial \sigma_{\epsilon}^2} = \begin{bmatrix}
1 & \rho & \rho^{2\nu} & \rho^{3\nu} & \rho^{4\nu} \\
1 & \rho & \rho^{2\nu} & \rho^{3\nu} & \\
1 & \rho & \rho^{2\nu} & & \\
(\text{sym}) & 1 & \rho & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
\end{bmatrix},
\]

\[
\frac{\partial \Sigma_i}{\partial \delta_{12}} = \begin{bmatrix}
2 & 3 & 4 & 5 & 6 \\
4 & 5 & 6 & 7 & \\
6 & 7 & 8 & & \\
8 & 9 & & & \\
10 & & & & \\
\end{bmatrix}, \text{ and}
\]

\[
\frac{\partial \Sigma_i}{\partial \delta_{21}} = \begin{bmatrix}
\end{bmatrix}
\]
\[
\frac{\partial \Sigma_i}{\partial \delta_{22}} = \begin{bmatrix}
1 & 2 & 3 & 4 & 5 \\
4 & 6 & 8 & 10 \\
9 & 12 & 15 \\
16 & 20 \\
(sym) & 25
\end{bmatrix}
\]

As in Murray's model, \( \nu \) can be estimated as another covariance parameter or it can be set to a constant. The derivative matrix for \( \frac{\partial \Sigma_i}{\partial \rho} \) when \( \nu \) is set to a constant \( c \) is:

\[
\frac{\partial \Sigma_i}{\partial \rho} = \begin{bmatrix}
0 & \sigma_e^2 & \sigma_e^2(2^c)\rho^{(2^c-1)} & \sigma_e^2(3^c)\rho^{(3^c-1)} & \sigma_e^2(4^c)\rho^{(4^c-1)} \\
0 & \sigma_e^2 & \sigma_e^2(2^c)\rho^{(2^c-1)} & \sigma_e^2(3^c)\rho^{(3^c-1)} \\
0 & \sigma_e^2 & \sigma_e^2(2^c)\rho^{(2^c-1)} & \sigma_e^2 \\
(sym) & 0
\end{bmatrix}
\]

If \( \nu \) is estimated in the model the constant \( c \) is replaced with the current estimate of \( \nu \) during each iteration of the model fitting process:
\[
\begin{bmatrix}
0 & \sigma_e^2 & \sigma_e^2(2^\nu) \rho(2^\nu - 1) & \sigma_e^2(3^\nu) \rho(3^\nu - 1) & \sigma_e^2(4^\nu) \rho(4^\nu - 1) \\
0 & \sigma_e^2 & \sigma_e^2(2^\nu) \rho(2^\nu - 1) & \sigma_e^2(3^\nu) \rho(3^\nu - 1) & \\
0 & 0 & \sigma_e^2 & \sigma_e^2(2^\nu) \rho(2^\nu - 1) & \\
0 & 0 & 0 & \sigma_e^2 & \\
(sym) & 0 & 0 & 0 & 0
\end{bmatrix}
\]

For models in which \( \nu \) is estimated the derivative matrix \( \frac{\partial \Sigma_i}{\partial \nu} \) must be evaluated. For the same hypothetical longitudinal study with regular timing being considered, these derivatives have the identical form as in Murray's model:

\[
\begin{bmatrix}
0 & 0 & \sigma_e^2 \rho^2 \log \rho \cdot 2^\nu \cdot \log(2) & \sigma_e^2 \rho^3 \log \rho \cdot 3^\nu \cdot \log(3) & \sigma_e^2 \rho^4 \log \rho \cdot 4^\nu \cdot \log(4) \\
0 & 0 & \sigma_e^2 \rho^2 \log \rho \cdot 2^\nu \cdot \log(2) & \sigma_e^2 \rho^3 \log \rho \cdot 3^\nu \cdot \log(3) & \\
0 & 0 & 0 & \sigma_e^2 \rho^2 \log \rho \cdot 2^\nu \cdot \log(2) & \\
0 & 0 & 0 & 0 & \\
(sym) & 0 & 0 & 0 & 0
\end{bmatrix}
\]

The diagonal elements and first off-diagonal elements of \( \Sigma_i \) are not functions of \( \nu \) if the data are regularly timed with integer values such 1, 2, ..., 10.

For irregularly timed data: For a complete description of the derivative matrix \( \frac{\partial \Sigma_i}{\partial \nu} \) when the data are irregularly timed, see Section 2.3.3.
CHAPTER III

ANALYSIS OF TOTAL CHOLESTEROL DATA

3.1 Total Cholesterol Measures

There are added complications to analyzing longitudinal data with these covariance models since they require modeling not only the usual expected value part of the model in $\mathbf{X}$, but also the covariance matrix. Investigators have found that changing the expected value model can have a profound effect on the choice of covariance structure for the data. That is, dropping or adding predictor variables from the expected value model can affect the characteristics of the covariance matrix. As a result, whenever variables are added or deleted from the expected value model, the adequacy of fit of the covariance matrix needs to be reassessed. This complication to the overall modeling process creates a back-and-forth type of sequence to the analysis, as will be shown. During the course of the analysis, different covariance structures need to be fit for each expected value model considered. Therefore, because of the dependence of the two, the expected value model and the covariance structure model must be considered throughout the analysis as being closely related. The objective of this analysis is to demonstrate the use of the different covariance structure models described in Chapter 2 and does not represent a complete analysis of the data.

The data in this chapter are total cholesterol measures (mg/dl) from two groups of men scheduled to be made at nine time points, one year apart. Due to dropouts and the scheduled termination of the study, the mean duration of
participation was about 7.5 years. For the purposes of this chapter we will label the groups A \( (N = 197) \) and B \( (N = 196) \).

The missing data patterns of the total cholesterol measures, scheduled to be taken one year apart, are shown in Table 3.1. Note that only 27% of the subjects have data for all nine measurements; the data are complete for 84% of the subjects for the first eight data points, month 6 through month 90. The 106 subjects with a ninth year visit were the earliest recruits for the study. Most of the subjects were recruited later during the next year, and were followed for less time because the study was terminated on a specified date.

For the models in this section, the default algorithms in BMDP-5V (usually Newton-Raphson) were used to obtain maximum likelihood estimates using a change in the log-likelihood between iterations of 0.00001 as the stopping criteria. The algorithms usually required from 5 to 15 iterations for convergence of the estimates.

### 3.2 Assessment of the Unmodeled Data

#### 3.2.1 Introduction

A plot of population mean values versus year in study will help to reveal the relationship between total cholesterol and year in study. The plot of the mean values for the two groups (Figure 3.1) shows that while group A has fairly constant means across time at about 275 mg/dl, group B means increase slightly for a few years and then level off at a about 250 MG/dl. These plots suggest a linear relationship between total cholesterol and year in study with separate lines for each group. The actual means appear in Table 3.2.
3.2.2 Trends in the Correlations and Covariances

One way to detect trends in a correlation or covariance matrix is to plot the matrix element values (correlations or covariances) as a function of "lag", where lag is time between measures. The actual plots are constructed using a separate "line" or series of correlations for each row in the upper diagonal matrix starting with the first off-diagonal element. The correlation plots contain $t-2$ series of at least 2 correlation data points in length for a study with measures at $t$ time points. For example, a study with 9 time points would have seven series containing 2, 3, ..., 8 correlations. The series for the covariance plots begin with the variance and contain $t-1$ series of at least a variance and a covariance. A study with $t=9$ time points would have eight series containing 2, 3, ..., 9 variance-covariance data points. The last series for each type of plot can be plotted as a single point.

Below are two different plots which represent two different phenomenon in the data, each indicative of a different covariance structure. The straight line plot indicated by $\bullet$ is what one would expect to find in a data set where the correlations between measures are constant, regardless of the amount of time between measures. In this scenario a compound symmetric model would be appropriate and the correlation matrix would have all correlations equal, or nearly so. If all the correlations are approximately equal, the plotted series appear to be flat, horizontal lines at some value of $\rho$, which can be represented by a single line.

The line indicated by $\ast$ represents a different phenomenon and leads to a different kind of model for the covariance matrix. The way the correlations tend to decrease as time between measures increases is indicative of an autoregressive effect in the errors of the repeated outcome measures. Here, the correlations start out high for measures close together in time and drop off quickly to near zero for
measures made far apart in time. The constant decrease of the correlations over time or the "dropping off" effect is what indicates the autoregressive effect. While for this plot the amount of decrease in correlations is dramatic, actual data sets might produce plots with more subtle drop-off rates. A plot of this type would be typical of covariance structures which model autoregressive trends such as AR(1) models. If the assumptions of the AR(1) model held, the actual plot would contain $t-2$ lines of correlations, and each one would drop off with nearly the same rate, with the lines appearing to lie on top of each other.

![Plot of the correlations from a repeated measures study](image)

Difference in time points between measures

Consider three different scenarios for longitudinal data where each is exhibiting some form of autoregressive trend in the errors. In each case differences in the variances and correlation structure lead to different kinds correlation plots. The first case is where the variances are about equal and the correlation for any two measures separated in time by the same amount are equal. This means the correlations are a function of lag only, an assumption of the AR(1) model, and the series will be plotted near each other, giving the appearance of a single line. The second case is where the variances increase slightly over time.
but the correlations are still AR(1). Each series will have the same shape but will originate from a higher point on the Y axis as the series lengthen, with the shortest series at the top and the longest at the bottom. As a result of the increasing variances, the plot will no longer appear as a single line. Consequently, one would not want to use a covariance structure that forces the variances to be equal (i.e. forces the series to appear as a single line). Finally, consider data where the variances are equal but the correlations between points equidistant apart in time change during the study and are no longer the same (i.e. no longer AR(1)). For this case the series would all originate from the same point on the Y axis but would fan out at different rates. They might fan out in a pattern or they might not. If the correlations between time points equidistant apart in time deceased during the study period, the lines would gradually flatten. The longest series would descend sharply and each series would get flatter until the last ones were nearly parallel to the x-axis. These three different examples show the diversity of possible correlation patterns, all closely related to the AR(1) model, yet each producing a graphically different correlation plot.

Plots of the covariances are also useful, especially for identifying potential compound symmetric models. For these plots both the variances and the covariances are plotted. Their interpretation is similar to the correlation plots. For example, autoregression trends are identified by covariances that decrease over time. Compound symmetric models would be characterized by the variances taking on one value in the plot and the off-diagonal covariances forming a flat line at a different value.

These plots should be made with the fixed effects (such as treatment, sex, etc.) removed from the responses. Since patterns in the correlation matrix can be due to the effects of variables not taken into account, important effects should be removed from the model. If there are many fixed effects that need to be
considered, it may be best to reduce the number of effects by running preliminary OLS regressions. To quickly decide which variables need to be included in the models for these plots, the data can be 'stacked' ignoring the correlation structure and OLS regression can be used to find important predictors with p-values below 0.15. Of course, there is no need to carry out this procedure if the study has only a few important fixed effects.

Once the number of potentially important fixed effect variables has been reduced to a reasonable number, say three or four, there are basically two ways to obtain the estimated correlation and covariance matrix. One way to do this is to fit a general multivariate model (e.g. PROC REG in SAS) to \( Y_i \), remove the fixed effects in the design matrix \( X \), and compute the estimated correlation and covariance matrix (using Output data sets and PROC CORR).

The recommended way to obtain the estimated correlation and covariance matrix, and perhaps the easiest, is to fit an unstructured covariance model in BMDP-5V with the fixed effect variables included in the model. The estimated correlation and covariance matrix are provided in the output. The actual plots will need to be constructed using some appropriate software such as Harvard Graphics.

**Trends in the Correlations and Covariances for these data**

The easiest way to detect trends in a correlation or a covariance matrix from longitudinal data is to plot the individual values (correlations or covariances) as a function of "lag", as just described. These plots help establish the dependence of the measures from the same person over time during the study. Figure 3.2 is a plot of the correlations by lag and shows that the correlations are decreasing as time between measures increases. The same is generally true for the
covariances except for the last covariance of each series (Figure 3.3). This is the covariance of the last measure, month 102, with each of the other time points. The variance of month 102 was very high at 2818 (the next closest was 2018) which in turn causes covariances with this time point to be large. Note that only 30% of the subjects had data for this time point. If the last covariance from each series is ignored, the overall trend for the covariances is clear -- a general decrease as time between measures increases. Table 3.2 has the corresponding covariances and correlation matrices which show the actual values for these plots. Note that there is no real trend in the magnitudes of the variances.

3.3 Modeling the Expected Value Part of the Model with Unstructured Covariance

3.3.1 Initial Models

The first model fit uses a separate mean for each group and time pair (i.e. a cell mean model) and a common unstructured covariance matrix for the responses from each subject. Let \( Y_{git} \) denote the response for the \( i \)-th subject in the treatment group \( g = A \) or \( B \) at Time = 1, 2, ..., 9. The equation for the cell mean model model for the \( i \)-th person at time \( t \) is

\[
Y_{git} = \alpha_{git} + \epsilon_{git} \tag{3.1}
\]

If the data are balanced, the ML estimates for this model are the cell means \( \bar{Y}_{g \cdot t} \). Earlier, a plot of the \( \bar{Y}_{g \cdot t} \) versus year in study for both groups (Figure 3.1) suggested that the slopes are not equal. As a result, models which allow for separate slopes for each group were fit. The equation for the model with separate slopes is

\[
Y_{git} = \alpha_{lg} + \beta_{lg}[\text{Time}] + \epsilon_{git}, \tag{3.2}
\]

where \( l \) denotes a linear polynomial model. Since the plots also do not suggest a straight-line relationship, polynomial trends will be fit. The plots suggests that
polynomials up to and including a cubic term will be sufficient. Equation 3.2 is
easily expanded to accommodate polynomial trend contrasts for the nine time
points:
Linear: \[ X'_L = [ -4 \quad -3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \quad 4 ] \]
Quadratic: \[ X'_Q = [ \quad 28 \quad 7 \quad -8 \quad -17 \quad -20 \quad -17 \quad -8 \quad 7 \quad 28 ] \]
Cubic: \[ X'_C = [ \quad -14 \quad 7 \quad 13 \quad 9 \quad 0 \quad -9 \quad -13 \quad -7 \quad 14 ] \]
The three models with cubic, quadratic, and linear trends which allow for
different polynomial effects for each group are
1) \[ Y_{git} = \alpha_{cg} + \beta_{clg}[X_L] + \beta_{cgg}[X_Q] + \beta_{ccg}[X_C] + \epsilon_{git}, \] (3.3)
2) \[ Y_{git} = \alpha_{qg} + \beta_{qlg}[X_L] + \beta_{qqg}[X_Q] + \epsilon_{git}, \] and (3.4)
3) \[ Y_{git} = \alpha_{lg} + \beta_{llg}[X_L] + \epsilon_{git}, \] (3.5)
where the first subscript on an \( \alpha \) or \( \beta \) indicates the degree of the model (\( l = \) linear,
\( q = \) quadratic, and \( c = \) cubic) and the second subscript on a \( \beta \) indicates whether
the coefficient corresponds to a linear, quadratic or cubic polynomial. The last
subscript \( g \) on an \( \alpha \) or \( \beta \) is the group indicator. Below is a table of some of the
results from the three linear effects models and their likelihood ratio (LR) tests for
comparison to the cell mean model or the higher order model.

<table>
<thead>
<tr>
<th>Model</th>
<th>(-2\lambda)</th>
<th>Number of Parameters</th>
<th>Model Comparison</th>
<th>(\chi^2)</th>
<th>df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Cell Mean</td>
<td>670.1</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) L, Q, and C</td>
<td>702.7</td>
<td>8</td>
<td>1</td>
<td>32.6</td>
<td>10</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>3) L and Q</td>
<td>718.2</td>
<td>6</td>
<td>2</td>
<td>15.5</td>
<td>2</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>3) Linear</td>
<td>789.7</td>
<td>4</td>
<td>3</td>
<td>71.5</td>
<td>2</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Note: 28,000 has been subtracted from \(-2\lambda\) for easier comparisons; i.e. for the
cell mean model \(-2\lambda = 28,670.1\).

All models were fit with an unstructured covariance matrix.
The high significance of the $\chi^2$ values indicates that on the basis of the LR tests alone, one is not justified to use any of the above polynomial models. Any departure from the cell mean model is statistically significant, as is any lower order polynomial model from its higher order alternative.

3.3.2 Mean Regression Plots

Figures 3.4 to 3.7 are the plots of the model-based estimates of group means from the cell mean model and the three polynomial models. Aside from the LR tests, these plots should help establish the extent to which the polynomial models depart from the cell mean model and from each other.

The estimated means from the cell mean model are very similar to the simple group means. A somewhat smoother representation of this process appears to arise from the cubic polynomial model, which seems to capture the essence of the cell mean plot, but with 10 less parameters. It models the slight increase for group B in years 1 to 3 and the leveling off of the two groups at measures after year five. The plot of the means from the quadratic model is similar to the cell mean plot for years 1 to 5. However, after year five the means for the two groups start decreasing as the quadratic term dominates. In the population mean plots, these means appear to be leveling off after year five. The model with a linear term only forces the group B means to increase at a constant rate, which seems unreasonable when considering the original cell mean plot.

From these plots it appears the model with a cubic, quadratic and linear term (i.e. cubic model) for each group captures the linear relationship of the cell means both successfully and adequately. Its attractiveness is not only that it fits the data well, but by doing so, supports the hypothesis that a cubic relationship exists between cholesterol and year in study, for years 0.5 to 8.5. It accomplishes
this using ten fewer parameters than the cell mean model. Note that one would not use the cubic model to extrapolate past 8.5 years. Below are the model-based estimates of group means from the cell mean and cubic models and their differences for comparisons.

<table>
<thead>
<tr>
<th></th>
<th>Group A Models</th>
<th></th>
<th></th>
<th>Group B Models</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Cell Mean</td>
<td>Cubic Model</td>
<td>Difference</td>
<td>Cell Mean</td>
<td>Cubic Model</td>
<td>Difference</td>
</tr>
<tr>
<td>1</td>
<td>273.46</td>
<td>273.23</td>
<td>0.23</td>
<td>224.43</td>
<td>224.06</td>
<td>0.37</td>
</tr>
<tr>
<td>2</td>
<td>276.21</td>
<td>277.00</td>
<td>-0.79</td>
<td>243.87</td>
<td>238.26</td>
<td>5.61</td>
</tr>
<tr>
<td>3</td>
<td>277.15</td>
<td>279.00</td>
<td>-1.85</td>
<td>244.07</td>
<td>247.18</td>
<td>-3.11</td>
</tr>
<tr>
<td>4</td>
<td>282.58</td>
<td>279.60</td>
<td>2.98</td>
<td>252.48</td>
<td>251.94</td>
<td>0.54</td>
</tr>
<tr>
<td>5</td>
<td>278.63</td>
<td>279.15</td>
<td>-0.52</td>
<td>255.99</td>
<td>253.66</td>
<td>2.33</td>
</tr>
<tr>
<td>6</td>
<td>276.07</td>
<td>278.03</td>
<td>-1.96</td>
<td>255.33</td>
<td>253.44</td>
<td>1.89</td>
</tr>
<tr>
<td>7</td>
<td>275.32</td>
<td>276.61</td>
<td>-1.29</td>
<td>251.89</td>
<td>252.42</td>
<td>-0.53</td>
</tr>
<tr>
<td>8</td>
<td>276.75</td>
<td>275.24</td>
<td>1.51</td>
<td>252.68</td>
<td>251.69</td>
<td>0.99</td>
</tr>
<tr>
<td>9</td>
<td>271.83</td>
<td>274.29</td>
<td>-2.46</td>
<td>253.83</td>
<td>252.37</td>
<td>1.46</td>
</tr>
</tbody>
</table>

The magnitudes of the differences are quite small, on the order of 1%, indicating that any bias introduced to the model by using the cubic polynomial model instead of the cell mean model will be quite small. Consequently, despite the results of the LR tests, which to some extent rule out its use as a "valid" model, the cubic polynomial model appears "reasonable" and will be used for the expected value part of the model while modeling the covariance structure. As discussed in Chapter 2, frequent checks will be made on the expected value part of the model to see if different covariance structures affect the relationship of cholesterol and time.
3.3.3 Checking Trends in the Covariances and Correlations

Now that preliminary investigation of the expected value part of the model has established that a cubic model fits the data well, the trends in the correlation and covariance matrices need to re-evaluated. The reason for this is that once the fixed effects of group and the polynomial trends have been ‘removed’ from the data, the correlation and covariance trends are subject to change. Figures 3.8 and 3.9 show the correlation and covariance plots for the cubic model using an unstructured covariance. Figures 3.10 and 3.11 show corresponding plots for the cell mean model. These plots show that: (1) the correlations and the covariances exhibit the same trends for the cell mean and the cubic fixed effect models with unstructured covariance, and (2) these trends are also consistent with the simple correlation and covariance plots from the unmodeled data. From this we conclude that the correlations continue to show a decreasing trend as the time between measures increases, and that different fixed effects models do not affect this process. The estimated covariance matrix for the cell mean model with unstructured covariance is displayed in Table 3.3 and shows that the variances are increasing slightly over time.

Also of note is Figure 3.12 which is a plot of the covariances estimated by OLS, i.e. ‘stacking’ an individual’s data, ignoring the correlations between measures for a subject, and using OLS to remove group effects and obtain residuals, which are subsequently analyzed. This is a short-cut method that does not require using BMDP-5V. (See Section 3.2.2 for details on estimating correlations and covariances using OLS). Of interest is that the same trends prevail when using this method, which ignores within-subject correlations, when removing the mean from each group separately.
3.3.4 Homogeneity of the Variances of Repeated Outcome Measures

Some of the methods described in this paper require homogeneity of the variances of the repeated outcome measures. In some covariance models, such as the compound symmetry model, the need for equality of variances is implicit in their formulae:

\[ \Sigma = \nu^2 I + \sigma^2 J, \]

where \( J = I - I' \). In the case where \( \Sigma \) is unstructured, we usually assume the diagonal elements of \( \Sigma \) are equal, or nearly so, while the off diagonal covariances are unspecified and allowed to vary.

No exact test exists for testing the equality of correlated, dependent variances. Tests are available for testing equality of variances from several independent populations but not when the variances are from the same subjects at different time points. One way to test for homogeneity of variance is to compare the likelihoods from a model with unstructured covariance to one with the variances restrained to be equal, with each having the same expected value model. Andrade and Helms (1984) showed that this likelihood ratio (LR) test is distributed as a \( \chi^2 \) with the degrees of freedom equal to the difference in the number of parameters for the two models. For example, if the fully parametrized model had nine estimated variance parameters, then the reduced model would have just one variance estimate, and the degrees of freedom for the LR test would be 8.

For the mixed model the variances will never be equal, except in two cases: (1) when \( \Sigma_i = I \), the case of compound symmetry or (2) the extreme case where \( \Sigma_i \) takes on an unnatural form, described in detail by Murray and Helms (1990, Chapter 3). They show that \( \Sigma_i \) will be a Toeplitz matrix (symmetric matrix with equal diagonal elements) only under extreme conditions. Since in a mixed model
the variances are functions of polynomials and depend on the random effects in the model, they can functionally increase or decrease over time or take on a parabola shape. For example, consider a study with four time points, \( t = 1 \) through \( t = 5 \). If the covariate Time is the only random effect in the model, one can see below how the variances can increase over time. The variance of \( Y_i \) is

\[
\text{Var}[Y_i] = \Sigma_i = Z_i D Z_i' + \sigma^2 Y_i ,
\]

(3.6)

with \( Z_i' = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \\ 3 \\ 1 \\ 4 \\ 1 \\ 5 \end{bmatrix} \), \( \mathcal{D} = \begin{bmatrix} 500 & 1 \\ 1 & 5 \end{bmatrix} \), \( \sigma^2 = 50 \), \( Y_i = I_5 \), and

\[
\Sigma_i = \begin{bmatrix} T1 & T2 & T3 & T4 & T5 \\ 557 & 513 & 519 & 525 & 531 \\ 574 & 535 & 546 & 557 & \end{bmatrix}
\]

Even if the Time trends are counteracted by another variable in the reverse magnitude, the variances result in a parabola-type of function, not equality. For example, a third column of \( Z_i \) with values that decreases over time (i.e. down the column), such as decrease in cigarettes usage during a study, can counter the natural increase of the variances caused by having Time as a random effects covariate. However, this tends to impose a parabolic shape on the magnitudes of the variances. In the naive example below, a perfect parabolic
shape results from columns of $Z_i$ with reverse order of magnitude. From (3.6) we get

$$Z_i' = \begin{bmatrix} 1 & 1 & 5 \\ 1 & 2 & 4 \\ 1 & 3 & 3 \\ 1 & 4 & 2 \\ 1 & 5 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 500 & 1 & 1 \\ 1 & 5 & 1 \\ 1 & 1 & 5 \end{bmatrix}, \quad \sigma^2 = 50 \quad Y_i = I_5,$$

and

$$\Sigma_i = \begin{bmatrix} \text{sym} & \text{sym} \\ \text{sym} & \text{sym} \end{bmatrix}.$$ 

Although this example is overly simplistic, it shows that even when the naturally increasing Time variable is offset in $Z_i$ by another variable that decreases over time, the variances are still unequal.

Assessment of Homogeneity for these Data

Upon having settled, at least temporarily, on a cubic model for the expected values, it is of interest now to model the covariance matrix. Most of the covariance structures under consideration require homogeneity of variance for the repeated outcome measures. Recall that no exact method exists to test correlated measures for homogeneity of variance. Before any covariance models are considered, the homogeneity of the variances should be assessed.
One way to test for homogeneity of variance is to compare the likelihoods from a model with unstructured covariance to one with the variances restrained to be equal, with each having the same expected value model. For the cubic and cell mean model, the likelihood ratio tests are given below:

<table>
<thead>
<tr>
<th>Expected Value Model</th>
<th>Covariance Structure</th>
<th>No. of Param.</th>
<th>$-2\lambda$</th>
<th>$\chi^2$</th>
<th>d.f.</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Mean</td>
<td>Unstructured</td>
<td>45</td>
<td>670.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell Mean</td>
<td>Equal Variances</td>
<td>37</td>
<td>699.55</td>
<td>29.44</td>
<td>8</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Cubic</td>
<td>Unstructured</td>
<td>45</td>
<td>702.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cubic</td>
<td>Equal Variances</td>
<td>37</td>
<td>730.68</td>
<td>28.00</td>
<td>8</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

These LR tests indicate that for either expected value model there may be a lack of homogeneity in the variances for these data. The estimated covariance matrices appear in Table 3.3 and 3.4 for the cell mean model employing the two different covariance matrices.

The range of the unstructured variances is 1375-2091 which does not seem to be an extreme violation of the assumption of equal variances. With the understanding that these models are fairly robust to small violations of assumptions, we will continue to model the covariance matrix, noting possible heterogeneity.
3.3.5 Conclusions

From the preliminary results obtained thus far, the following statements about modeling the covariance structure can be made.

1. A compound symmetric structure is not likely to fit the data well due to the variation in the variances and in the covariances.

2. The fairly consistent trend of decreasing correlations over time suggests that an autoregressive trend can be modeled.

3. Subject knowledge that cholesterol measures vary considerably from person to person suggest that a mixed model, which is useful for modeling between-subject variation, is a favorable choice for the covariance structure.

4. Models capable of modeling between subject variation and autoregressive trends simultaneously may work well given (2) and (3).

5. As discussed in Chapter 3, some choices of $Z$ in a mixed model impose a structure on the covariance matrix in which the variances increase over time. This may be advantageous since an increasing trend in the variances has been observed.
3.4 Modeling the Covariance Matrix Using the Mixed Model, Compound Symmetry, and AR(1) Covariance Structures

3.4.1 Introduction

The next step is to fit the mixed model, compound symmetric model, and the AR(1) model. These three basic models will provide a broad range of covariance structures. These models will be fit for both the cell mean model and the cubic model to assure a continuing check on the expected value part of the model.

There are several things to be learned from the mixed, compound symmetric and AR(1) models and their comparisons to the unstructured case. If there are not large amounts of missing data or mistimed data, the unstructured covariance matrix should display the trends present in the data. However, if these data are sparse, trends need to be interpreted with consideration given to the amount of missing and mis-timed data. Recall that with sparse data, unstructured covariance matrices may contain some poor estimates, in which case trends will need to be interpreted with caution. In many cases, however, it is mostly general trends that we want to observe and these should be present even if the covariance matrix contains some poorly estimated parameters.

In the mixed model each subject has its own covariance matrix; we need some special terminology to describe the general form of $\mathbf{\Sigma}$. For all mixed models in Section 3.4, $\mathbf{Z}'$ takes the form

$$\mathbf{Z}' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{bmatrix},$$

and for the compound symmetry model, $\mathbf{Z}'$ takes the form

$$\mathbf{Z}' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \end{bmatrix}.$$
For the i-th subject, \( Z_i = \mathbf{E}_i Z \) where \( \mathbf{E}_i \) is a 9 \( \times \) 9 matrix of 0's and 1's that picks off the rows of \( Z \) corresponding to the observed times for the i-th subject. If a subject has complete data for all nine visits, \( \mathbf{E}_i \) is a 9 \( \times \) 9 identity matrix.

Sections 3.4.2 and 3.4.3 contain parallel analyses using the same covariance models in combination with the cell mean and cubic expected value models, respectively. This will allow direct comparison of the two expected value models for the three covariance structures under consideration.

3.4.2 Evaluation of the Covariance Structure for the Cell Mean Model

If we compare the three covariance matrix estimates in Table 3.5 from the mixed, compound symmetric and AR(1) models with the unstructured covariance matrix in Table 3.3, the one from the mixed model most closely models the unstructured case. The mixed model picks up the two trends in the data: (1) the slightly increasing variances and (2) the extent to which the covariances tend to decrease over time. The magnitudes are also similar for the two models as demonstrated by the range of variance estimates (1375 - 2091) and (1418 - 2144); and the covariance estimates (843 - 1587) and (990 - 1629), for the unstructured and mixed model, respectively. The AR(1) model gives constant variance estimates of 1642 and covariance estimates that fall off faster than what appears to be occurring in the data. The compound symmetry model gives constant variance estimates of 1626 and does not model the decreasing correlations over time, with all covariance estimates equal to 1127.

The correlation matrix estimates in Table 3.6 also indicate similar results. The mixed model correlation matrix closely resembles that from the unstructured case in Table 3.3. The range of correlation estimates in the unstructured model (0.54 - 0.81) is closer to the mixed model (0.56 - 0.79) than the AR(1) model.
\(0.11 - 0.76\). The compound symmetric or 'constant correlation' model gives a constant \(\widehat{\rho} = 0.69\), which ignores any correlation trends in the data. The drawback of the AR(1) model is that it seems to have modeled a trend of decreasing correlations that is stronger and more consistent than what is actually occurring in the data.

The actual covariance parameter estimates, \(\widehat{\Sigma}\), from these models can be used to help decide which covariance models appear to fit the data best. The model parameters for the three covariance structures are provided in Table 3.7. The mixed model parameter \(\widehat{\delta}_{22} = 11.9\) is the estimated variance of the random coefficient for Time (Z = 7.6, \(p < 0.0001\)). The high significance of the Time parameter basically means that the model cannot be justifiably reduced to the compound symmetric model. However, the high significance of the parameter \(\widehat{\rho} = 0.76\) for the AR(1) model (Z = 71.3, \(p < 0.0001\)) does not necessarily demonstrate that there is a strong autoregressive trend in the data.

The \(-2\lambda\) values\(^1\) also reflect in terms of likelihoods what is seen visually in the covariance and correlation matrices and noted again in their model parameters. The unstructured model has \(-2\lambda\) of 670.1 followed by the mixed model (772.4), compound symmetry (938.6) and the AR(1) model (1148.8). Since low values of \(-2\lambda\) are generally indicative of favorable fit compared to larger ones, one can see that of the three structured covariance models, the mixed model has the 'best fit' followed by the compound symmetric model and the AR(1) model. These values generally imply that the mixed model is fitting the data better than the AR(1) and compound symmetry models but perhaps not as well as the unstructured model.\(^2\)

\(^1\) For simplicity of presentation, the constant 28,000 has been subtracted from each \(-2\lambda\) value in this chapter.

\(^2\) Formal Likelihood Ratio tests will be conducted in Section 3.8.
Figures 3.13 to 3.16 show the graphs of the covariances and correlations for the mixed model and the AR(1) model. These same plots for the unstructured model are in Figures 3.10 and 3.11. The graphs reveal that the mixed model has the covariance structure most similar to the model with unstructured covariance.

**Evaluation of Fixed Effects**

Although not shown, the estimates of expected value parameters, in this case the cell means, remain essentially the same regardless of which covariance structure is used.

**Conclusions**

After fitting these three models and comparing them to each other and to the unstructured covariance, it seems that the mixed model provides the closest structured fit to the unstructured covariance matrix in terms of overall trends and magnitudes of the covariances and correlations.

**3.4.3 Evaluation of the Covariance Structure for the Cubic Model**

At this point we want to fit the same three covariance models, the mixed model, the compound symmetric model and the AR(1) model, with an expected value model containing cubic, quadratic and linear effects (i.e. cubic model). The reason for this is to compare the results from the three covariance structures for the cubic model with the results from the cell mean model. By comparing the estimates of the parameters from both the expected value part of the model and the covariance model we will determine the extent to which the two expected value models depart. If the estimates are similar, it will further justify use of the cubic model when using more complicated structures to model the covariance structure.
Tables 3.8 and 3.9 are the estimated covariance and correlation matrices for the three covariance structures with the cubic model. The unstructured estimates of the covariance matrix and correlation matrix are also provided in Table 3.10. These cubic model based covariance and correlation matrices are quite similar to those from the cell mean model in Tables 3.5 and 3.6. The model parameters are also alike. This supports the adequacy of the cubic model as a good model for the data and is consistent across all three covariance structures, and the unstructured case. The model parameters for the three covariance structures are provided in Table 3.11.

The same is true for the plots of the covariances and correlations in Figures 3.17 to 3.20, which reveal the same trends for the cubic model as for the cell mean model for all three covariance structures.

Evaluation of Fixed Effects

When making a comparison between the cell mean model and the cubic model of the model-based means at each time point, the estimates are essentially the same for a given covariance structure. That is, there is almost no difference between the two expected value models for, say, the mixed model. The model-based means do vary, however, if comparisons are made between covariance structures for the same expected value model. An example would be the regression population means from an AR(1) versus the compound symmetric model for the cell mean model.

An illustration of how well the model-based group means from the general mixed model compare to those from the cell mean model can be made by comparing Figure 3.21 to Figure 3.4.
Conclusions

The results obtained in this section from the cubic models fit with the four different covariance structures all compare favorably to their cell mean model counterparts. This is true for both (1) the expected value part of the model in terms of comparisons of the model-based means and (2) the estimated correlation and covariance matrices. In terms of overall fit among the structured covariance models, the mixed model is doing the 'best' job of modeling the data -- i.e. it is providing model-based means and a covariance matrix that are very similar to the cell mean model with unstructured covariance, but using less parameters in each part of the model to accomplish this.
3.5 Model With Stationary Variance and Exponentially Decreasing Correlation

Three other types of covariance structures will be investigated, structures which give the analyst the ability to model covariance matrices for data whose trends do not fall under the strict realm of the mixed model, compound symmetry or an AR(1) type of covariance. These covariance structures, although technically more complicated and computationally more difficult to fit, offer the analyst more flexibility than the previous models; one is a more general version of the AR(1) model and the other two incorporate some of the advantages of the mixed model and autoregressive-type structures. The three models which will be used for the remainder of this chapter are (1) a model with stationary variance and exponentially decreasing correlation, (2) a mixed model with compound symmetry plus an autoregressive error, and (3) a mixed model with general random effects plus an autoregressive error. Each of these models was described in detail in Chapter 2. Sections 3.5, 3.6, and 3.7 will present and discuss results from fitting these three covariance structures for the cell mean and cubic model. Section 3.8 discusses choosing the best model and contains comments on the overall process of modeling the covariance for this data.

3.5.1 Introduction

The models in this section are generalizations of the AR(1) model (Murray and Helms, 1990) which include a curvature parameter to model the strength and pattern of the correlations over time (see equation 2.3.3). As noted in Section 2.3.2, the AR(1) model from Section 3.4 is a special case of this model when \( \nu = 1 \). Values of \( \nu < 1 \) indicate a decreasing correlation that is more steady over time than when \( \nu = 1 \). This is the case for the total cholesterol measures where there is evidence of a trend in the decreasing correlations, but where the AR(1) model seemed to 'over-model' that trend (as indicated by the model-based correlations
dropping to 0.11). Therefore, we might expect that the autoregressive model with a value of \( \nu < 1 \) should fit the data better than the special case AR(1) with \( \nu = 1 \). The estimates from the AR(1) model with \( \nu = 1 \) are used to determine starting values for the model when \( \nu \) is estimated.

3.5.2 Evaluation of the Covariance Structure

Model with \( \nu \) set to 1

This is the model discussed in the Section 3.4.2 and 3.4.3, and as mentioned above is a special case of the more general model. Table 3.12 shows the estimated covariance and correlation matrices for the cell mean and cubic models with \( \nu = 1 \). As discussed earlier, this covariance model forces the correlations to decrease at a faster rate than observed in the simple correlations and in the unstructured covariance models. Note how similar the matrices are for the different models.

Model with \( \nu \) set to 2

Table 3.13 shows the model covariance and correlation matrices from the cubic and cell mean model \( \nu = 2 \). As suspected, this covariance structure with \( \nu > 1 \) is forcing the correlations and covariances to decrease at a dramatic rate. They approach zero quickly, and are equal to zero for any measures taken more than three years apart. The data are not consistent with this fast drop-off of correlation.

Model with \( \nu \) estimated as \( \widehat{\nu} \)

The covariance and correlation matrices from the cubic and cell mean model are in Table 3.14. Immediately one can see that this covariance structure with parameter \( \nu \) estimated as \( \widehat{\nu} \) fits the data much better than when \( \nu \) is set to an arbitrary constant (\( \nu \geq 1 \)). These estimates compare favorably to those from
the cell mean model with unstructured covariance (Table 3.3). The range of the covariances and the correlations for both expected value models are similar. Recall that this model forces the variances to be equal and restrains any two measures separated in time by the same amount to have the same covariance (or correlation).

The covariance estimates obtained from this model might be considered as 'averages' of those obtained from the unstructured model. For example, consider the cell mean model with parameter estimate $\nu = 0.29$. The correlation of any two measures one year apart (i.e. all the first off-diagonal elements) is 0.76. This is almost equal to the average of the eight first off-diagonal elements from the unstructured, cell mean model (Table 3.3) which are $(\times 0.01)$ 70, 71, 73, 76, 77, 78, 79, and 80, with $\bar{X} = 75.5$. The variance estimate is 1628, which is close to the average of the nine variances of 1400, 1495, 1375, 1799, 1776, 1856, and 2091 with $\bar{X} = 1663$. The model parameter estimates for the different covariance structures are provided in Table 3.15.

Conclusions

The covariance structure with parameter $\nu$ estimated as $\hat{\nu}$ models the covariance matrix fairly well -- especially considering it only uses three parameters. The one drawback is that it forces the estimated variances to be equal when for these data, in the unstructured, cell mean model, they seem to increase with time. The issue of equal variances was addressed in the section on homogeneity of variance; recall the significance of the LR test against the null hypothesis of equal variances. This causes some concern that bias maybe added to the model if the variances are forced to be equal.
3.6 Mixed Model with Compound Symmetry Plus an Autoregressive Error

3.6.1 Introduction

The model proposed by Diggle (1988) and described in Section 2.4 combines aspects of the mixed model with compound symmetry and an autoregressive type of model. The advantage of this model over the AR(1) model is that it has the added ability to model between-subject variation. Its advantage over a compound symmetry model is the additional ability to model decreasing correlations over time. Since both of these characteristics occur in the data, we should expect this model to fit the data well. One drawback to consider is that for this model the variances are constrained to be equal, which may introduce some unwanted bias into the model.

The variance equation (2.4.2) for this model is

$$\Sigma_i = \tau^2 I + \omega^2 J + \sigma^2 R(t_i),$$

where the elements of \( \sigma^2 R(t_i) \) are of the form \( \exp(-\alpha|t_j - t_k|^{\nu}) \), the curvature parameter \( \nu \) is set to a constant, and the parameter \( \alpha \) can be thought of as \(-\ln\rho\). The four covariance parameters to be estimated are: \( \tau^2, \omega^2, \sigma^2 \) and \( \alpha \). Diggle does not estimate \( \nu \) but rather suggests fitting models where \( \nu \) is set to an integer constant. Estimation using this covariance structure was done with \( \nu = 1 \), \( \nu = 2 \), and \( \nu = 3 \).

3.6.2 Evaluation of the Covariance Structure

Model with \( \nu \) set to 1

The estimation algorithm for the cubic model with \( \nu = 1 \) converged but the algorithm did not converge for the cell mean model, indicting possible lack of fit. Three of the parameter estimates for the cubic model (Table 3.20) have large standard errors in comparison to their respective estimates. Despite the poor
precision for this model, the actual covariance and correlation matrices in Table 3.16 are similar to those from earlier models in Sections 3.4 and 3.5. The fact that one expected value model did not converge and the one that did has poor estimates leads us to try other values for $\nu$.

Model with $\nu$ set to 2

With $\nu$ set to 2, both the cell mean and the cubic model converged to final estimates, and all four parameter estimates have small standard errors (Tables 3.19 and 3.20). The estimated covariance and correlation matrices in Table 3.17 are very similar to those when $\nu = 1$ for the cubic model, although the model with $\nu = 2$ would be preferred since the standard errors are much smaller. The range of the covariances for the cell mean model ($933 - 1224, \hat{\sigma}^2 = 1632$) is close to that for the cubic model ($931 - 1224, \hat{\sigma}^2 = 1635$). The range of the correlations, $0.57 - 0.75$, is the same for both models.

Model with $\nu$ set to 3

There is little change from the models with $\nu = 2$ to models with $\nu = 3$ for the estimated covariance and correlation matrices. The matrices for both expected value models remain similar (Table 3.18). The range of the covariances for the cell mean model is $961$ to $1208$, with $\hat{\sigma}^2 = 1629$; for the cubic model the range is $959$ to $1206$, with $\hat{\sigma}^2 = 1633$. The correlations ranged from $0.59$ to $0.74$ for both models. The parameter estimates (Tables 3.19 and 3.20) are similar to those from the model with $\nu = 2$.

Figures 3.23 to 3.26 are the four plots of interest for this section: the estimated covariance and correlation matrices for the cell mean and cubic model for $\nu = 2$. (The plots for $\nu = 3$ are almost indistinguishable from Figures 3.23-
3.26). Recall that this model forces the variances to be equal and restrains any two measures separated in time by the same amount to have the same covariance (or correlation). These single line plots show a steady decrease in both the correlations and the covariances over time, and exhibit almost no discernible differences between expected value models. Note that the autoregressive part of the covariance model forces the covariances to continually decrease, while in the unstructured models (Figures 3.9 and 3.11), the covariances increase slightly over time after a strong initial decrease. This continual decrease might be viewed as a model-induced deviation from the true underlying process observed in the unstructured case, i.e. a lack of fit.

Conclusions

The cell mean model and the cubic model behave very similarly with this covariance structure. The covariance model with $\nu = 2$ or $\nu = 3$ is preferred over the model with $\nu = 1$ because (1) the standard errors for the covariance parameter estimates are considerably smaller when $\nu = 2$ or $\nu = 3$, and (2) the covariance parameter estimates are consistent across both models $\nu = 2$ and $\nu = 3$, while those from the model with $\nu = 1$ are strikingly different. Despite the substantial differences in estimates of covariance parameters in the model with $\nu = 1$, the resulting estimates of covariance and correlation matrices compare closely to those when $\nu = 2$ or $\nu = 3$, suggesting possible robustness of the model. Although the $-2\lambda$ values indicate the model with $\nu = 1$ has the best overall fit, the models with $\nu = 2$ or 3 seem the better choice due to reasons (1) and (2) above. Between the models with $\nu = 2$ or $\nu = 3$, the model with $\nu = 2$ has a slight edge in terms of the likelihood $-2\lambda$ values. The overall trends seen in the plots of the covariances and the correlations are encouraging but this model may be forcing the covariances to decrease further than the observed process. Formal LR tests will be conducted at the end of this chapter to compare the models in this section.
3.7 General Mixed Model With an Autoregressive Error

3.7.1 Introduction

The models in this section also incorporate features from two different covariance structures. From the mixed model comes the ability to model between subject variation, but without the constraint of equal variances. From Murray’s generalized autoregressive model (1990) we gain flexibility for modeling exponentially decreasing correlation. The covariance structure takes the form:

\[ \Sigma_i = Z_i \Delta Z'_i + \sigma^2 R(t_i), \]

where the j-th row and k-th column element of \( \sigma^2 R(t_i) = \text{cov}(e_j, e_k) = \sigma^2 \rho |t_j - t_k|^\nu \), \( 0 \leq \rho \leq 1 \), with ‘curvature’ parameter \( \nu \geq 0 \). The first part of the covariance model, \( Z_i \Delta Z'_i \), is directly from a mixed model. From Murray’s generalized model come the parameters \( \rho, \sigma^2 \), and \( \nu \). Since these total cholesterol measures exhibit both decreasing correlations over time and large between subject variation, this model ought to provide a favorable fit to the data. Three models were fit with \( \nu \) set to an integer constant of 1 or 2. Models where \( \nu \) was estimated as \( \hat{\nu} \) were also fit.

3.7.2 Evaluation of the Covariance Structure

Model with \( \nu \) set to 1

The models with \( \nu = 1 \) produced covariance and correlation matrices that are very similar for both expected value models. The variances for this model increase from 1399 to 2111 for the cubic model and 1395 to 2112 for the cell mean model (Table 3.21). The range of covariances was alike for both models: (1) cell mean: 1012 to 1639 and (2) cubic model: 1010 to 1632. The correlations were 0.72 to \( \sim 0.8 \) for both models.

For the cell mean model and the cubic model, the covariance and
correlation matrices compare very closely to each other. Excepting the covariance parameter $\delta_{12}$ for $\varphi$, all the parameters for both expected value models are highly significant indicating that their standard errors are small (Tables 3.24 and 3.25).

Model with $\nu$ set to 2

No noteworthy changes occur in the covariance parameters or the covariance and correlation matrices when $\nu$ is changed from 1 to 2. Table 3.22 contains the covariance and correlation matrices and the covariance parameter estimates are in Tables 3.24 and 3.25.

Model with $\nu$ estimated as $\widehat{\nu}$

When $\nu$ is estimated as $\widehat{\nu}$ there are notable changes in the parameter estimates compared to when $\nu$ is set to 1 or 2. These parameter estimates are given in Tables 3.24 and 3.25 and the covariance and correlation matrices are in Table 3.23. The general trends and magnitudes of the covariances do not change. The parameter estimates from this model have larger standard errors than when $\nu = 1$ or $\nu = 2$. Also, the parameter estimate $\widehat{\nu}$ has a Z-score of 1.2 which indicates that addition of this parameter to the model may not be necessary or even helpful.

Figures 3.27 to 3.30 and Figure 3.22 are the five plots of interest for this section: the estimated covariance and correlation matrices for the cell mean and cubic model with $\nu = 2$ and model-based means for the cubic model, $\nu = 2$. These plots show a sharp initial decrease in the covariances and a slight increase as lag times increase. The correlations have a sharp initial decrease followed by a very slight decreases. These plots are the closest of any model to the unstructured models (Figures 3.9 and 3.11).
Conclusions

The covariance and correlation matrices are quite similar for both values of \( \nu \) and are similar for both expected value models. Overall, the model with \( \nu = 2 \) would be a slight favorite over the model with \( \nu = 1 \) in terms of \( -2\lambda \) values. The model with estimated parameter \( \hat{\nu} \) has convergence problems that need to further investigation. It is possible that the model becomes over-parametrized when \( \hat{\nu} \) is included as the sixth covariance parameter to be estimated.

3.8 Choosing the Best Model

Until now comparisons have been made between expected values models within a certain covariance structure type. Also, comparisons which depend on the choice of parameter for a covariance type have been made such as whether \( \nu \) is set to an integer or estimated. Now comparisons will be made between the different covariance structures for each expected value model in the form of LR tests. Tables 3.27 and 3.28 provide some results and LR tests for the cell mean and the cubic expected value models.

The \( -2\lambda \) values and the accompanying LR tests are another way to help decide which models fit the data best. Models 1 through 4 are basic models that are often seen in the literature and are easily fit using BMDP-5V. The low \( -2\lambda \) value for the unstructured model is to be expected considering the large number of parameters it uses to estimate the covariance matrix. The mixed model provides the next best fit but the LR test is significant suggesting lack of fit. Despite that, the mixed model is better, as measured by \( -2\lambda \), than the compound symmetric and the AR(1) models.

Models 5 through 11 are seen less often in the literature and require extra effort to obtain parameter estimates. Effective use of these models requires a
good understanding of the data, as well as the models themselves.

A quick glance at either Table 3.27 or 3.28 shows that models 5 and 6, the
general mixed models plus autoregressive (AR) errors, are providing the closest
model to the unstructured covariance. The model with $\nu = 2$ fits slightly better
than when $\nu = 1$; The plots of the covariances and correlations from Section 3.7
reveal a smoothed process of the trends observed in the fully parametrized
covariance matrix.

The set of mixed models with compound symmetry plus an AR error
(models 7, 8 & 9) provides the next best overall fit. The model with $\nu = 2$ was
considered the best of the three models because its parameter estimates were
consistent with those from the $\nu = 3$ model. Although the $-2\lambda$ values are lowest
for the model with $\nu = 1$, its parameters estimates were inconsistent with all
previous models.

The models with stationary variance and exponentially decreasing
correlation (Models 10 & 11) offer a competitive model when $\nu = \tilde{\nu}$. When $\nu$ is set
to 1 or 2 the model provides poor fit.

Some general statements about the different models are:

1. The models that allow the variances to increase, i.e. the mixed model and
the general mixed model plus an AR error, provide the closest $-2\lambda$ values
to the models with unstructured covariances.

2. The models that constrain the variances to be equal may only be effective
when the variances are very homogeneous.

3. The Murray and Helms (1990) generalized model (stationary variance with
decreasing correlation, Models 10 & 11) is seen to be a substantial
improvement over the well known but restrictive AR(1) model. It was also
seen to be competitive with general mixed models (Models 2, 7, 8, and 9)
in terms of $-2\lambda$, and uses just three covariance parameters.

4. The three basic structured covariance models, the mixed model, compound symmetry, and AR(1), while perhaps adequate for some data analyses, clearly are not the best structured covariance models for this data set.

One issue purposely not mentioned until now is the fact that with the high number of subjects in each group, there may be too much power in this study to rely solely on LR tests to decide which model best fits the data, if at all. Recall in Section 3.2 and 3.3 that any departure from the cell mean model was statistically significant in terms of LR tests. This is where the plots of the model-based covariances, correlations, and group means are helpful. These plots show that many of the covariance structures are reasonable models, in comparison to the unstructured case. The best model appears to be the general mixed model plus an AR error with $\nu = 2$ and the cubic model for the expected value function. This model seems to be a smoother process of what is actually seen in the cell mean model with unstructured covariance.

The cubic model uses 8 parameters compared to 18 cell means and the mixed model with AR error uses 5 covariance parameters compared to 45 for the fully parametrized matrix. In terms of the expected values the cubic model (Figure 3.22) removes the random 'bumps' in the plot of the cell means. The structured covariance model provides a smooth process for the covariances and correlations. It removes the random variation that is seen in the plots for the unstructured matrix (Figures 3.9 and 3.11). By removing the random fluctuations through the modeling process, the trends are smoothed into a more believable pattern.
Some Overall Comments:

The usefulness of these models is clear from the plots and tables. The researcher must decide if fitting specialty models (with FORTRAN subroutines) is worth the effort or if the basic covariance structures offered in BMDP-5V (or future releases of SAS) provide sufficient fit to their data. If interest lies mostly in the expected value part of the model, then the built-in structures may be sufficient. However, if there is inherent interest in the covariance matrix and the dependence of the measures over time, these alternative models offer information about the covariances not available from basic covariance structures.
CHAPTER IV
INCLUSION OF BASELINE VALUES

4.1 The Lipid Research Clinics Coronary Primary Prevention Trial

4.1.1 Description of the Study

The Lipid Research Clinics (LRC) Coronary Primary Prevention Trial (CPPT) was a clinical trial sponsored by the National Heart, Lung, and Blood Institute. Although it was long believed that high levels of blood cholesterol lead to Coronary Heart Disease (CHD), no studies had demonstrated convincingly that reducing cholesterol led to a lower rate of onset of heart disease in humans. Cross-sectional studies had established that statistical correlations existed between high levels of total cholesterol and CHD, but findings from studies up to that time were inconclusive. For example, many had not used rigorous experimental designs or had used poor statistical methods. Some did not show a large enough difference between treatment groups; others showed a significant mortality and morbidity associated with cholesterol lowering drugs (LRCP, 1979).

The LRC-CPPT used a randomized, double-blind design dividing patients equally among two treatment groups within eight well defined prognostic strata at each of 12 clinics. The prognostic strata were based on low and high risk categories of positive exercise ECG, LDL cholesterol, and a combined risk measure based on blood pressure, age and number of cigarettes smoked. Participants were to be followed for a minimum of seven years.
The primary objective of the trial was to test the hypothesis that long-term reduction of serum cholesterol in hypertensive men initially free of clinical coronary heart disease would lead to lowered incidence of CHD. A total of 3806 asymptomatic males aged 35-39 with primary Type II hyperlipoproteinemia were randomized at 12 North American clinical centers. At the time of entry to the study, these men were in good general health and free of any overt symptoms of CHD. Middle aged men were picked because they showed the most potential for the study: older men might have irreversible lesions and younger men are at insufficient risk in terms of study power (LRCP I, 1984).

It was determined that an intervention based on diet alone would be prohibitively long and expensive to conduct. The chosen mode of intervention included a cholesterol-lowering pharmacologic agent and a moderate cholesterol-lowering diet. All participants were prescribed a standard diet allowing daily intake of approximately 400 mg. of cholesterol with a polyunsaturated-to-saturated fat ratio of approximately 0.8. The diet was imposed on both groups to standardize their lipid intake over the course of the study. In addition, 24 grams per day of a bile acid sequestrant, cholestyramine, was prescribed to half the men; the other half received a placebo. Cholestyramine is not metabolized or absorbed in the gut but rather acts as a site to which bile acids bind and impedes their reabsorption. This was seen as an advantage over drugs which need to be absorbed. Its main disadvantages were the inconvenience of taking 6 packets per day and the gastro-intestinal side-effects which included bloating and constipation.

The general results of the study showed that for participants assigned to cholestyramine, the average plasma total and LDL cholesterol levels were reduced by 13.4% and 20.3%, respectively, which were 8.5% and 12.6% greater than the placebo group (LRCP I, 1984). In terms of health benefits, the cholestyramine group experienced a 19% risk reduction in definite CHD and/or definite CHD
death (LRCP II, 1984). Adherence to protocol was measured by the number of unused packets that were returned at each clinic visit. This will not be taken into account here, although it was in the LRC analysis report (LRCP II, 1984).

Recall that the data in Chapter 3 were all post-randomization measures of total cholesterol. In this chapter, the data include the two pre-diet, pre-randomization visits, three pre-randomization visits on diet alone, and the nine post-randomization ‘six-month visits’ (same as Chapter 3) which were one year apart. Data from the Minnesota clinic were chosen for two reasons: (1) this was one of the larger clinics (N = 393 subjects) and (2) this was considered to be one of the most successful clinics in terms of compliance and clinic-patient relationship. The treatment group has 197 subjects and the placebo group has 196 subjects. The outcome variables to be analyzed is total cholesterol measured at several baseline values and nine time points one year apart and the covariates are Time and Treatment group.

Table 3.1 shows the missing data patterns of the total cholesterol measures, taken one year apart starting at visit 7, which was approximately 6 months after the first visit. For visits 1, 2, 3, and 4 there was no missing data. Visit five had six subjects with missing data (visit 6 is not analyzed). Note that only 27% of the data are complete for all nine measurements; the data are complete for 84% of the subjects for the first eight data points, month 6 through month 90. The 106 subjects with a ninth mid-year visit were the earliest recruits for the study. Most of the subjects were recruited later during the next year and were followed for less time.

In this chapter we analyze the cholesterol data as data from a clinical trial; consideration of pre-randomization measures is a key issue. These measures are important since they characterize a starting point, or baseline, against which
individual change can be measured. This is a central theme of the clinical trial: to what extent does the drug/placebo treatment alter an individual’s cholesterol level, over the course of the study, when compared to baseline values?

There are three basic evaluations (or hypotheses) that investigators are typically interested in when a clinical trial involves pre and post measures (Stanek, 1988). (Since this trial involves cholesterol, we will refer directly to cholesterol as the dependent variable, although the ideas are applicable to many types of dependent variables.) The first issue is the difference in the pre-trial measures of total cholesterol for the two treatment groups. The second is the comparison of the change measures, post-treatment cholesterol minus pre-treatment cholesterol, for the two treatment groups. The third is determining if the difference of the change for the two groups is the same.

Each of these hypotheses addresses a different issue; together, they summarize the principal results of a clinical trial. The first hypothesis addresses the idea of having an equal distribution of “disease” (hypercholesterolemia) in the drug and the placebo group. The relevant question is ‘was the randomization of patients to drug treatment successful in balancing the groups with respect to important correlates?’ Senn (1989a) and others (Altman, 1985; Senn, 1989b) have argued that if randomization was used to allocate patients to treatment groups then this test ought not to be performed. They claim investigators should be assured of the validity of their own study by realizing that on average, across all studies, randomization will be effective. Senn (1989b) urges investigators who cannot accept this to use analysis of covariance to correct for baseline disparities among groups. Laird (1983) agrees that if randomization is used for allocating patients to treatments, then comparing pre-treatment groups is of little interest (presumably to statisticians). She goes on to note, however, that investigators are always interested in testing this hypothesis.
The second hypothesis addresses mean changes from baseline after the intervention. In this study, this hypothesis will investigate the extent to which cholestyramine lowers total cholesterol, beyond a regimented diet given to all patients. An additional issue is whether the control treatment (placebo and diet) lowers mean total cholesterol.

The third hypothesis tests if the two groups experienced the same changes, i.e. are the mean change scores the same for both groups? This can be interpreted as an interaction of amount of change by treatment group. If the drug is effective in lowering cholesterol, the average change scores (baseline – post-randomization) should be larger for the treatment group than the placebo group. For the present study, this hypothesis is the essence of the trial results (although this was an important issue in the CPPT, the primary issue was about lowering incidence of CHD).

4.2 Incorporating Baseline Data into the Analysis
4.2.1 Choice of Analysis Variables

Table 4.1 shows the scheduled time sequence of visits for the first year of the study and describes the timing of diet and treatment allocation. Randomization to drug or placebo was done at visit 5, making visit 6 the first post-randomization visit. Visit 3 is the first visit with patients on diet; this means visits 1 and 2 were pre-diet, pre-randomization and visits 3, 4, and 5 were with-diet, pre-randomization. Recall that all patients, those on drug and placebo, were put on a common diet at visit 3 and were expected to remain on this diet throughout the study.

Several alternative strategies for incorporating baseline (pre-randomization) data into an analysis are listed below.
1. Treat the baseline values as additional time points in the series of measures.

2. Treat the baseline measure or an average baseline measure as a covariate in the analysis.

3. Create difference scores (gain scores, change scores, etc.) of (post-randomization – baseline) scores and analyze them as a series.

4. Create percent change scores by dividing the difference scores by a subject’s baseline value and analyze as a series.

5. Analyze the difference scores treating the baseline measures as a covariate in the analysis (Laird, 1983).

Each of these strategies is a valid approach to analyzing longitudinal data. The appropriateness of one versus another depends upon factors such as the subject matter being studied and the type of study design. For example Bock (1975) claims that the choice of analysis, change scores versus treating the baseline measures as covariates, should be based on the study design. He suggests analyzing change scores when the assignment to drug groups is done without randomization and treating the baseline measures as covariates when randomization is used. Laird (1983) points out that under some circumstances, either analysis will give the same valid results. Crager (1987) also discusses the use of analysis of covariance in parallel-group clinical trials.

For this study we will be interested in not only the analysis of the fixed effects, but also how the choice of analysis affects the covariance models. That is, choices of response variables will be made in conjunction with choices for structural models for the covariance.
4.2.2 Choice of Analysis of the Repeated Measures

There are numerous ways to analyze repeated measures data which includes information about baseline. The choice of baseline value for this study is complicated by whether the choice includes averaged baseline measures or single measures of baseline cholesterol. Outlined below are different analysis strategies which include treating baseline values as additional time points, as covariates, and as part of difference scores. Valid options for analyses that treat baseline values as additional time points are:

1. Include all 5 baseline measures: Visits 1, 2, 3, 4, and 5.

2. Use first visit and the last pre-randomization visit: Visits 1 and 5.

3. Use mean $\bar{X}_1$ of the two pre-diet visits and the last pre-randomization visit: $\bar{X}_1$ and Visit 5.

4. Use mean $\bar{X}_1$ of the two pre-diet visits and mean $\bar{X}_2$ of the three on-diet, pre-randomization visits.

Valid options for selecting a baseline covariate are:

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Dependent-Variable Time Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Visit 1</td>
<td>Visits 2, 3, 4, 5, 7, ..., 55</td>
</tr>
<tr>
<td>2. Visit 1</td>
<td>$\bar{X}_2$ and Visits 7, ..., 55</td>
</tr>
<tr>
<td>3. $\bar{X}_1$</td>
<td>Visits 3, 4, 5, 7, ..., 55</td>
</tr>
<tr>
<td>4. $\bar{X}_1$</td>
<td>$\bar{X}_2$ and Visits 7, ..., 55</td>
</tr>
</tbody>
</table>
Options for analysis with difference scores equal to \([\text{cholesterol measure} - \text{baseline value}]\) are:

<table>
<thead>
<tr>
<th>Baseline Value</th>
<th>Dependent-Variable Cholesterol Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Visit 1</td>
<td>Visits 2, 3, 4, 5, 7, ..., 55</td>
</tr>
<tr>
<td>2. Visit 1</td>
<td>(\bar{X}_2) and Visits 7, ..., 55</td>
</tr>
<tr>
<td>3. (\bar{X}_1)</td>
<td>Visits 3, 4, 5, 7, ..., 55</td>
</tr>
<tr>
<td>4. (\bar{X}_1)</td>
<td>(\bar{X}_2) and Visits 7, ..., 55</td>
</tr>
</tbody>
</table>

Percent differences, which are usually \(100 \times \frac{\text{difference score}}{\text{baseline}}\), offer the same variety of analyses as the difference scores. Senn (1989a) claims "ratios (or percentages) are less tractable than differences". Kaiser (1989) points out that division of a difference score by the baseline makes no sense unless the difference score and baseline value are (highly) correlated. He also mentions that dividing differences by baseline supports the notion that a given change is more important for subjects with low rather than high baselines. In fact, for patients with high cholesterol the opposite may be true when it comes to lowering risk of CHD; it may be more important for patients with higher baseline values to lower their levels than for patients with lower baseline values. For these reasons percent differences will not be considered.

The remainder of this chapter will focus on the analysis of four models that deal with treating the baseline values as time points or as covariates in the analysis. The models are (1) A model in which all 5 baseline visits are included in the analysis as additional time points; (2) A model related to (1) except where the 5 baseline visits are used to form 2 averages, 1 pre-diet and 1 on-diet, and the averages become additional time points; (3) A model in which the cholesterol measure at visit 1 is treated as a covariate in the analysis and the remaining 4 baseline visits are included as additional time points; and (4) A model similar to
(3) except where the average of the 2 pre-diet baseline values is treated as a covariate in the analysis and the average of the 3 on-diet measures is an additional time point.

These models seem like the ones most likely to be chosen from the above lists for a typical analysis of longitudinal cholesterol measures. Generally, they represent two important ways to treat the data: as added time points or as covariates in the analysis. Additionally, for each model type, the baseline values will be evaluated in their original form and then averages are taken and the data re-analyzed. This will allow evaluation of the effects, if any, that averaging baseline values has on the covariance structure. Models which contain difference scores will not be considered since they do seem appropriate for these data. Reasons for this are discussed in detail in Section 4.4. For reasons mentioned earlier, percent differences will also not be analyzed.
4.3 Treating Baseline Values as Additional Time Points

4.3.1 Introduction - Unstructured $\mathcal{E}$.

When averages of the baseline values are made (e.g. the average of visits 1 and 2 and of visits 3, 4, and 5) the variance of these new average measures is typically lower than that of the individual measures, as might be expected. Incorporation of these average baseline scores into the analysis introduces 'artificially' low variances into $\mathcal{E}$. Table 4.2 shows the unstructured covariance matrix for the model in which the averages of visits 1 and 2 (pre-diet, pre-randomization) and visits 3, 4, and 5 (on diet, pre-randomization) are concatenated to the nine post-randomization six-month measures from Chapter 3. The fixed effects comprise a cell mean model with a mean for each Time and Treatment group combination (22 estimated parameters). Note that the first two variances of the averaged baseline values are considerably smaller than those of the single measures at visits 7 through 55.

Another way to treat the baseline values is to simply concatenate them directly to the individual response vector $Y_i$ without averaging, resulting in more time points. Consider the addition of all five pre-randomization baseline values (two pre-diet and three on-diet visits) to the analysis with the nine post-randomization measures from Chapter 3, resulting in 14 time points. For this analysis averages are not taken; the single measures are used. Fitting an unstructured covariance matrix with 28 cell means gives $\mathcal{E}$ in Table 4.3. The variances for visits 1 through 5 appear lower, as a group, than the post-randomization visits 7 through 55. However, they are not as low as the variances of the two mean value measures seen in Table 4.2. The average of these five variances is 1088 compared to the average of the 9 post-randomization visit variances of 1660. Although the range of variances is about the same for both models (1000–2100 for the unaveraged values vs 900–2100 for mean values), the
change in variance size is more abrupt for the model with mean values and any modeling (or smoothing) of these variances would seem more difficult than for the unaveraged measures.

There are two issues that arise from averaging baseline values. First is that averaging baseline values is an effort to guarantee the fixed effects are as accurate as possible. When there are several values that represent the same phenomenon, averaging reduces the effects of measurement error and natural human fluctuation. In our case, averaging the two pre-diet and the three on-diet baseline measures helps to more accurately describe an individual’s cholesterol at each of the pre-randomization study points. If one is interested mainly in the analysis of expected values, then averaging these values is a desirable feature of the analysis.

The second issue is that averaging baseline values affects the variance associated with that time point, resulting in some variances and covariances in \( \Sigma \) being ‘artificially’ low. Consider the variances in Table 4.2. From the magnitudes of the variances, it would seem inappropriate to fit any structural model which requires that the variances be equal. For example, fitting an equal variance structure to this data such as an autoregressive model with stationary variance, or a model with compound symmetry, would clearly introduce bias (or lack of fit) into the covariance structure. Although the same is probably true for the covariance matrix in Table 4.3, the effect is not as large. Recall from Chapter 3, that only two of the structural models being considered allow the variances to vary. They were (1) the mixed model with random effects and (2) the mixed model with random effects plus an autoregressive error. This does not leave many choices of covariance structures when the variances are not equal.

Model-based estimates of group means are plotted for both models in
Figures 4.1 and 4.2. These plots show the large drop in total cholesterol in the first few months of the study, followed by a gradual increase. The two plots are similar except that Figure 4.1 has more time points in the first six months. The plots, as expected, are identical after six months. The variances and covariances are plotted for each model in Figures 4.3 and 4.4. The same trends that were seen in plots from Chapter 3 are seen here: a sharp initial decrease in the covariance followed by smaller decreases or slight increases. (Note that due to software limitations, some of the shorter series had to be left off the graph.) Generally, these two plots show similar trends.

The effects of induced variance reduction by averaged baseline values is an important consideration for modeling the covariance matrix, and merits further attention. This issue will be investigated in detail in Section 4.3.3. The next section, 4.3.2, will address the effects of fitting the covariance structures described in Chapter 2 to $Y_i$ when the 5 baseline visits are treated as additional time points.

4.3.2 Using Individual Baseline Measures as Additional Time Points

Including baseline measures in the analysis as additional time points is one acceptable strategy for analyzing longitudinal data. In this section all five baseline measures are included in the individual observation vector $Y_i$. The expected value model has been expanded to fit cell means for the baseline measures and to fit cubic polynomials to the post-randomization visits. The reasons for this expected value model are two-fold: (1) the five pre-randomization visits would be difficult to model (See Figures 4.1 and 4.2), making the cell mean model for the baseline measures an appropriate choice; and (2) from Chapter 3 we know that cubic, quadratic, and linear polynomials appear to model the treatment effects quite successfully. This combined expected value model of the cell mean model and cubic polynomials will be the only expected value model considered.
The design matrix for this model (without the 8 treatment × [cell mean and cubic trends] columns) is below.

\[
\begin{bmatrix}
\text{Int.} & \text{Trt.} & \text{Cell Means} & \text{Cubic Trends} & \text{Time (years)} \\
\hline
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0.08 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0.16 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0.24 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0.32 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0.40 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & -4 & 28 & -14 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & -3 & 7 & 7 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & -2 & -8 & 13 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & -17 & 9 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -20 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & -17 & -9 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 2 & -8 & -13 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 3 & 7 & 7 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 4 & 28 & 14 \\
1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0.08 \\
1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0.16 \\
1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0.24 \\
1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0.32 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & -4 & 28 & -14 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & -3 & 7 & 7 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & -2 & -8 & 13 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & -17 & 9 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -20 & 0 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & -17 & -9 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & 2 & -8 & -13 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & 3 & 7 & 7 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & 4 & 28 & 14 \\
\end{bmatrix}
\]

For this model the time points are not equidistant apart in time as they were in Chapter 3. The first six visits were scheduled one month apart and the remainder are one year apart in time. This effects fitting some models, such as the autoregressive and mixed models which involve the values of time in their model equations. For the mixed model, the individual random effects design matrix $Z_i$ usually contains a time parameter and autoregressive models have the absolute value of time differences in the exponent of $\rho$. Conceptually, this is not a
problem as these models provide maximum likelihood estimates for unequal time intervals. Computationally, however, these models were more difficult. For example, estimates from the mixed model with random effects and autoregressive error, which were estimated easily for the yearly data of Chapter 3, could not be obtained from the estimation algorithms for this data. In general, the models with unequal spacing of time took longer to converge and sometimes required more accurate starting values than when the model had equal spacing.

A Note on Unequal Spacing and BMDP-5V

There is no safeguard or warning in BMDP-5V to prevent the program from obtaining estimates from unequally spaced data for the 'AR(1)' covariance structure, which is a built-in covariance structure option. The program assumes that the data are equally spaced if the AR(1) option is selected. A careless user with irregularly-timed data could be presented with meaningless estimates of covariance covariance parameters and not realize that the estimates were incorrect. The 'random' or mixed model option requires specification of either a general $\mathbf{Z}$ or an individual $\mathbf{Z}_i$ in the model MODEL paragraph. If a time parameter is to be included as an individual random effect, the actual values must be input directly. Of course, for an unstructured covariance or compound symmetry, two of the other available options, the spacing of the time points is not a factor in obtaining covariance parameter estimates.

Evaluation of the Models

The $-2\lambda$ likelihood values for several models are provided in Table 4.4. The most noticeable change in likelihood patterns for these models, if compared to those in Chapter 3 (see Table 3.28), is that the general mixed model is not fitting the data as well as the pure autoregressive models\(^1\). The adjusted $-2\lambda$ values\(^2\)
for the mixed model \((-2\lambda = 576.0)\) is considerably higher than Murray's model with \(\nu\) estimated as \(\bar{\nu} = 0.222\) \((-2\lambda = 406.5)\). As expected, the model with compound symmetry has the worst fit \((-2\lambda = 1156.7)\). The model with the best fit was Diggle's\(^1\) when the parameter \(c\) was set to 0.50, with \(-2\lambda = 398.0\). Oddly though, models with \(c < 0.50\) produce negative variance estimates.

Table 4.5 provides a closer look at Diggle's mixed model with compound symmetry plus autoregressive error. All four estimates \(\tau^2, \phi^2, \sigma^2\) and \(\alpha\) behave nicely for \(0.60 \leq c \leq 2.0\). Note that the \(-2\lambda\) value for each model continues to decrease as \(c\) becomes smaller. In fact, it continues to decrease for values of \(c < 0.50\) (not shown), at which point the some variance estimates are negative. Strict interpretation of the likelihood criterion would mean choosing the model with \(c = 0.50\) as the best fitting model since it has the lowest \(-2\lambda\) value. The parameter estimates, however, are not stable in this area. For a model with more locally stable estimates it might be wiser to pick a model with a larger \(c\), perhaps near 0.80, even though according to the likelihood ratio criterion it has greater lack of fit than when \(c = 0.50\). The abrupt change in the magnitude of the variance parameters around \(c = 0.50\) indicates possible lack of fit of the model in this area.

Models with stationary variance and exponentially decreasing correlation (Murray’s model) are described in more detail in Table 4.6. The parameter estimates for these models are much better behaved. For example, note that the minimum \(-2\lambda\) value is at \(\bar{\nu} = 0.222\), the MLE. The smooth increases (and

\(^1\) Models with stationary variance and exponentially decreasing correlation (Murray and Helms, 1990). See Section 2.2 or equation 2.3.4.

\(^2\) All \(-2\lambda\) values have been adjusted by subtracting 46,126.16 (See Table 4.4.)

\(^3\) Mixed model with compound symmetry plus autoregressive error. See Section 2.4 for explanation of model parameters.
decreases) of the magnitudes of the parameter estimates as \( c \) changes are an indication that this model provides good fit to the data. Also of interest are the parameter estimates for the classical AR(1) model with \( c = 1.0 \). Use of this special case with \( \nu = 1 \) would have meant introducing considerable bias to the model.

**Graphical Assessment of the Covariance Structures**

In addition to the likelihood ratios, graphical methods can be used to judge goodness of fit of the covariance structures. Plots of individual covariance values vs lag time were introduced in Chapter 3. Figures 4.5 through 4.8 contain covariance plots for unstructured \( \hat{\Sigma} \), the general mixed model, Diggle’s model and Murray’s model, respectively. Visually, none of the three covariance structures seem to model unstructured \( \hat{\Sigma} \) perfectly. The \(-2\lambda\) values indicate that the models from Diggle and Murray, whose plots are nearly identical, provide better fit than the mixed model. This is supported visually by the graphs. Note the plot from the mixed model has consistently increasing covariances, which are not necessarily evident in the unstructured \( \hat{\Sigma} \) plot. The autoregressive models do a better job of modeling the decreasing correlations, which tend to become nearly constant at larger lag values. Figures 4.9 and 4.10 show that the fixed effects for the two models, unstructured \( \hat{\Sigma} \) and Murray’s model, are indistinguishable from each other. This supports the notion that, in some cases, these covariance structures have little impact on the expected value (fixed effect) part of the model.
4.3.3 Using Average Baseline Measures as Additional Time Points

In this section the five baseline values are divided according to whether they were pre-diet or on-diet and then averages are taken. This results in a mean cholesterol value for the two pre-diet measures and a mean value for the three on-diet measures. These two mean values are then concatenated to the individual response vector $\mathbf{Y}_i$, resulting in 11 possible time points. To further develop some earlier thoughts, the idea of averaging baseline values is appealing for several reasons: (1) the baseline visits are closer together in time (1 month) than the post-randomization visits selected for analysis (1 year), so using average measures is intuitively attractive; (2) they clearly represent two distinct phases in the pre-randomization phase of the study: before being placed on a regimented diet and after being placed on diet, and so taking averages accurately represents a patients condition at each phase; and (3) using average measures reduces the number of time points, which simplifies the analysis. This may especially be useful for studies with multiple baseline values and many post-baseline time points; averaging reduces the total number of time points, which helps to assure convergence of the estimation algorithms.

The expected value model for these data is the same as that used with the five-baseline model, except that now there are only two cell means to estimate instead of five. The desired design matrix $\mathbf{X}$ can be thought of as a sub-matrix of the one shown in Section 4.3.2, with the rows and columns corresponding to the 3rd, 4th, and 5th cell means deleted. The design matrix for this model (without the 5 treatment $\times$ [cell mean and cubic trends] columns) is below.
Design Matrix

\[
\begin{bmatrix}
\begin{array}{cccc|c}
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & -4
\end{array}
\end{bmatrix}
\]

The time sequences shown with the design matrix indicate that the data are not equally spaced. It should be noted that even if the first measure at visit 1 was the sole baseline value in the analysis, the data would still be irregularly timed since the yearly visits begin at about five months and are one year apart thereafter. The two-average models experienced about the same amount of computational difficulties (or a little less) as the models which utilized all 5 baseline values, when compared to their equally spaced counterparts of Chapter 3. That is, compared to equally spaced data, they generally took longer to converge and one model, the random effects mixed model with autoregressive error, never converged. Compared to the five-baseline model, these two-average baseline models were less sensitive to starting values and usually converged in fewer iterations.
Evaluation of the Models

The adjusted $-2\lambda$ likelihood values for selected models are shown in Table 4.7. These models follow the same pattern as the five-baseline model of Section 4.3.2. Again, we see that the autoregressive models are fitting the data better than the general mixed model. The adjusted $-2\lambda$ values for the mixed model ($-2\lambda = 394.5$) is higher than that from Murray’s model with $\nu$ estimated as $\hat{\nu} = 0.2727$ ($-2\lambda = 330.1$). If $\nu$ is set to the constant 1.0, Murray’s model provides worse fit ($-2\lambda = 1275.5$) than the compound symmetry model ($-2\lambda = 716.7$). Diggle’s model with parameter $c = 0.35$ provides the best fit with $-2\lambda = 325.6$. As was the case for the the five-baseline model, Diggle’s models with low values of $c$ ($c \leq 0.30$) produced negative variance estimates.

Table 4.8 provides parameter estimates and $-2\lambda$ values for Diggle’s mixed model with compound symmetry plus autoregressive error for different values of $c$. As was seen in Table 4.5, the parameter estimates $\tau^2$, $\omega^2$, $\sigma^2$ and $\alpha$ are most stable for higher values of $c$ and break down when $c \leq 0.30$. The $-2\lambda$ values continue to decrease despite the negative variance estimates, which makes choosing a ‘best’ fitting model difficult. The model with $c = 0.50$, where the estimates are stable, might be a good choice for a final model. Examination of $\hat{\Sigma}$ for these models for $c = 0.35$ and $c = 0.50$ in Table 4.10 shows that both models produce essentially the same covariance matrix. Note that these models have constant variances and equal covariances for measures equidistant apart in time.

It is interesting to compare the covariance matrix from Diggle’s models to that of the mixed model in Table 4.11 and the unstructured $\hat{\Sigma}$ in Table 4.2. In the unstructured matrix, the covariances decrease swiftly at first and then increase slightly as the amount of time between visits increases. The autoregressive component of Diggle’s structure is not able to model this increase,
and neither is Murray's model. The mixed model covariance matrix for this data has covariances that increase as time between measures increases, which is the observed pattern in the unstructured $\hat{\Sigma}$. In terms of modeling the variances, the mixed model covariance matrix has increasing covariances over time. The autoregressive models have stationary, or constant, variances.

We are presented with the mixed model which has two advantageous characteristics for modeling $\hat{\Sigma}$ over the autoregressive structures. The mixed model allows for increasing variance, where Diggle's and Murray's model have stationary variance. The autoregressive models have constantly decreasing correlation (and covariance) structures while the mixed model covariance structure being used here allows for increasing covariances over time. Apparently, these two advantages of the mixed model do not outweigh the ability of the autoregressive structures, which according to $-2\lambda$ values, are providing better fit for $\hat{\Sigma}$. By site the mixed model $\hat{\Sigma}$ seems to be a better fit to the unstructured $\hat{\Sigma}$. The $-2\lambda$ values, however, favor the autoregressive models as providing better fit.

Parameter estimates from Murray's model (models with stationary variance and exponentially decreasing correlation) are presented in more detail in Table 4.9. The parameter estimates for these models are well behaved and the MLE for the model with $\nu = 0.2727$ clearly has the minimum adjusted $-2\lambda = 330.15$. As was the case for the five-baseline model, the smooth patterns of the parameter estimates as $c$ varies are an indication that this model provides good fit to the data. It should be noted that fitting the classical AR(1) model (except with irregular timing) with $c = 1.0$ would introduce considerable bias to the model.

---

4 The random effects matrix $Z'$ is shown in Table 11. The shape of the variances and covariances are influenced by the covariates in $Z'$. The covariance matrix will take on different patterns with inclusion of different covariates.
Graphical Assessment of the Covariance Structures

Covariance plots are provided for unstructured $\hat{\Sigma}$, the general mixed model, and the models from Diggle and Murray in Figures 4.11 through 4.14, respectively. These plots generally take the same shape as those in Section 4.3.4. Again, the mixed model seems to over-model the increasing trends of the covariances over time. The autoregressive models, which are providing the best overall fit, are very similar to each other and can be thought of as an average of the individual series seen in the unstructured $\hat{\Sigma}$ in Figure 4.11.

To Average or Not to Average

We have investigated two models: one in which all five baseline values were included in the analysis and one in which two averages from those same five baseline values were included in the analysis.

In terms of fixed effects, one model allows for greater discrimination among the baseline values; the other shortens the data vector and hence simplifies the analysis. The differences are clear and have been noted.

In terms of the covariance matrix, it appears that, at least for these data, the chosen models perform about the same relative to each other for goodness of fit, in the case of average or individual baseline measures. That is, in both cases the autoregressive models fit better than the mixed model; specifically for the autoregressive models, Diggle's model fits slightly better that Murray's model.
4.4 Treating Baseline Values as Covariates

4.4.1 Introduction

A preliminary analysis of covariance using the measure at Visit 1 as a covariate was performed to evaluate how baseline cholesterol related to blood levels after intervention. A separate, univariate model analysis was performed at each time point in which the cholesterol measure at visit 1 was included in the model as a covariate. In particular, it was of interest to test if the regression coefficient was equal to 1.0. The model equation for the post-baseline cholesterol measure from the \(i\)-th subject from treatment group \(j\) is

\[
Y_{ij} = \alpha + \tau_j + \beta_j X_{ij} + e_i \tag{4.1}
\]

where \(\alpha\) is the intercept,

\(\tau_j\) is the fixed effect for the \(j\)-th drug group,

\(\beta_j\) is the group regression coefficient (slope) parameter,

\(X_{ij}\) is the baseline measure at visit 1, and

\(e_i\) is the random error term,

where \(j = 1\) (for drug) or 2 (for placebo) and \(i = 1, 2, ..., n\). The slope parameters were found to be the same for both treatment groups indicating a common slope adequately fits the data. The model equation then simplifies to

\[
Y_{ij} = \alpha + \tau_j + \beta X_{ij} + e_i
\]

where \(\beta\) is the common slope parameter for both treatment groups. If \(\beta = 1.0\), the covariance model reduces to an analysis of variance model:

\[
Y_{ij} = \alpha + \tau_j + X_{ij} + e_i.
\]

In this case, analysis of the change-from-baseline data is a reasonable alternative to including the covariate in the model. The equation for analyzing difference scores (post-randomization − baseline) is:

\[
D_{ij} = \alpha + \tau_j + e_i \tag{4.2}
\]

where \(D_{ij} = (Y_{ij} - X_{ij})\). If difference scores were analyzed and the covariate was
also included in the model, the covariance model would take the form

\[ Y_{ij} - X_{ij} = \alpha + \tau_j + (\beta - 1)X_{ij} + e_i \]  \hspace{1cm} (4.3)

In this case, the test of \( H_0: \beta = 0 \), routinely provided in regression programs, has the same interpretation as the test of \( H_0: \beta = 1 \) in (4.1). Rejection of either null hypothesis means the slope in (4.1) is not equal to 1.0 and that model (4.1) is preferred. Below are the univariate-model results, separately by visit, comparing models (4.1) and (4.2).

<table>
<thead>
<tr>
<th>Visit</th>
<th>( \hat{\beta} )</th>
<th>Model (4.1) Covariate Model MSE</th>
<th>Model (4.2) Difference Score Model MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.78</td>
<td>534</td>
<td>581</td>
</tr>
<tr>
<td>3</td>
<td>0.68</td>
<td>515</td>
<td>619</td>
</tr>
<tr>
<td>4</td>
<td>0.69</td>
<td>560</td>
<td>655</td>
</tr>
<tr>
<td>5</td>
<td>0.75</td>
<td>641</td>
<td>702</td>
</tr>
<tr>
<td>7</td>
<td>0.64</td>
<td>982</td>
<td>1110</td>
</tr>
<tr>
<td>13</td>
<td>0.62</td>
<td>1074</td>
<td>1217</td>
</tr>
<tr>
<td>19</td>
<td>0.53</td>
<td>1171</td>
<td>1296</td>
</tr>
<tr>
<td>25</td>
<td>0.55</td>
<td>1192</td>
<td>1396</td>
</tr>
<tr>
<td>31</td>
<td>0.60</td>
<td>1438</td>
<td>1599</td>
</tr>
<tr>
<td>37</td>
<td>0.66</td>
<td>1185</td>
<td>1306</td>
</tr>
<tr>
<td>43</td>
<td>0.61</td>
<td>1369</td>
<td>1525</td>
</tr>
<tr>
<td>49</td>
<td>0.70</td>
<td>1339</td>
<td>1432</td>
</tr>
<tr>
<td>55</td>
<td>0.70</td>
<td>1868</td>
<td>1995</td>
</tr>
</tbody>
</table>

\( \dagger \) The associated F-tests for \( H_0: \beta = 1 \) with (1, \( N - p - 1 \)) d.f. were all significant at the 0.001 level; \( N \) is the total number of subjects and \( p = 2 \) is the number of model parameters (in addition to the intercept).

The \( \hat{\beta} \) values above are all less than 1.0 and the average is about 0.65. If the slopes were close to 1.0, it would imply that analysis of difference scores was essentially equivalent to using the measure at visit 1 as a covariate. If this were
the case, the mean square error (MSE) from the two models would be expected to be about the same. Notice that in each case the MSE from the model with the baseline covariate is less than the difference score models, which indicates that the covariate provides a small, yet consistent, gain in precision. These results, although performed at separate time points, indicate that an analysis with the visit 1 cholesterol measure as the covariate is probably a better choice of analysis compared to analyzing difference scores. This is also follows the recommendations by Bock (1975), who encourages a covariate analysis for studies with randomization of treatments.

4.4.2 Choosing a Covariate

The idea of using an average baseline value as a covariate is appealing. The two pre-diet measures, visit 1 and 2, seem especially appropriate to average since they are essentially pre-intervention of any kind. The next three measures, from visits 3, 4 and 5, are on-diet, pre-randomization and the average of these three visits also seems like an appropriate baseline value. The problem with these averaged or mean values is that their variances are smaller than individual measures, as noted in Section 4.2. The alternative to averaged values is to use an individual measure such as total cholesterol at visit 1.

Two approaches have arisen from these discussions that offer equally valid ways for treating baseline values as covariates. The first is to analyze the individual measures, using cholesterol at visit 1 as the covariate and to concatenate the other four baseline cholesterol measures to the series of nine post-randomization visits (visit 7 through visit 55 from Chapter 3). The second is to use the average of visits 1 and 2 (pre-diet, pre-randomization) as the covariate and include the average of visits 3, 4 and 5 (on-diet, pre-randomization) as an additional time point in the series of visits. By fitting structural models to those
two different sets of measures, we will gain insight to the effects of averaged and unaveraged baseline values on the covariance matrix. The next two sections will involve fitting structural models to these two models.

Preliminary Plots

The model-based estimates of group means for the cell mean model with unstructured $\hat{\Sigma}$ and the accompanying covariance plots can be seen in Figures 4.15 through 4.18 for two preliminary models. Note that both have the same nine post-randomization visits (7, 13, ..., 55) for time points while one has cholesterol at visit 1 as a covariate and the other has the mean cholesterol from visits 1 and 2 as the covariate. The closeness of the two estimated group mean plots indicates that both expected value models are about the same regardless of which covariate is used. The covariance plots are very similar for all of the individual series except the longest one -- the one that contains the first off-diagonal covariances. These are covariances of each visit with the 'first six-month visit', i.e. visit 7, and they are considerably lower than the other series. Note that the covariances for the model with covariate 'mean cholesterol from visits 1 and 2' are slightly smaller in magnitude than when the covariate is 'cholesterol at visit 1'. This is to be expected since the mean value acts to reduce the overall variation more than the individual value.
4.4.3 Models with Cholesterol Measure at Visit 1 Treated as a Covariate with Other Baseline Measures Treated as Additional Time Points

When including a covariate in the model it is important to assess whether separate slope parameter estimates are needed for the two drug treatment groups. Below are $-2\lambda$ values from two sets of models. One has a cell mean model for all the expected responses and another has a cell mean model for the baseline visits with cubic contrasts for the nine post-randomization visits. All models employ an unstructured covariance matrix. The $\chi^2$ values show that the common slope model provides adequate fit to the data and that separate slope parameters are not necessary. This holds for both expected value models. All subsequent models will use a common slope for the covariate, ‘cholesterol measure at visit 1’.

### Comparison of Separate and Common Slope Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Covariance Structure</th>
<th>Expected Value Model</th>
<th>Baseline Covariate</th>
<th>$-2\lambda$</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unstructured</td>
<td>Cell Mean</td>
<td>Separate slopes</td>
<td>42280.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Unstructured</td>
<td>Cell Mean</td>
<td>Common slope</td>
<td>42282.2</td>
<td>1</td>
<td>1.6</td>
</tr>
<tr>
<td>3</td>
<td>Unstructured</td>
<td>Cell Mean &amp; Cubic</td>
<td>Separate slopes</td>
<td>42313.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Unstructured</td>
<td>Cell Mean &amp; Cubic</td>
<td>Common slope</td>
<td>42315.1</td>
<td>1</td>
<td>1.6</td>
</tr>
</tbody>
</table>

### Evaluation of the Models

Table 4.12 provides a summary of the models that were successfully fit to the data. Note that two models, Diggle’s mixed model with compound symmetry plus autoregressive error and the mixed model with random effects plus autoregressive error, failed to converge to final estimates despite repeated attempts with different starting values. The only ‘specialty’ model from Chapter 2 that provided valid parameter estimates was Murray’s model with exponentially
decreasing correlation and stationary variance.

The pattern of best fitting models is the same as that seen for the previous two models. That is, Murray's autoregressive model with $\bar{\nu} = 0.2302$ is providing the best fit ($-2\lambda = 528.7$) followed by the general mixed model ($-2\lambda = 535.4$) and the compound symmetry model ($-2\lambda = 1094.1$). The classical autoregressive model with $\nu$ set to the constant 1.0 (and also with unequal spacing) provides poor fit, with $-2\lambda = 2329.4$.

The adjusted $-2\lambda$ values for Murray's model are provided for different values of $\nu$ in Table 4.13. The model with MLE $\bar{\nu} = 0.2302$ provides the best fit of any model as indicated by the lowest $-2\lambda$ value. The lower values of $\bar{\sigma}^2$ and $\bar{\rho}$ compared to those in sections 4.3.3 and 4.3.4 reflect the overall reduction of variances and covariances in the model. These values are reduced from about $\sigma^2 = 1500$ and $\rho = 0.70$ to $\sigma^2 = 1100$ and $\rho = 0.57$. The covariance matrices from the unstructured model, mixed model and Murray's model with $\bar{\nu} = 0.2302$, shown in Table 4.14, all reflect the overall variance/covariance reduction in $\mathcal{H}$.

Graphical Assessment of the Covariance Structures

The covariance plots for the unstructured $\mathcal{H}$, the mixed model, and Murray's model are shown in Figures 4.19 through 4.21, respectively. Murray's model is providing the best fit according to $-2\lambda$ values, and appears to be a smoothed average of the individual series seen in the unstructured $\mathcal{H}$. The mixed model seems to over-model a slight increasing covariance trend in the unstructured $\mathcal{H}$.
4.4.4 Models with Mean Baseline Measure from Visits 1 and 2 Treated as a Covariate, Mean Baseline Measure from Visits 3 - 5 is an Additional Time Point

This section deals with averaged baseline values, where the first average (pre-diet, pre-randomization) is used as a covariate and the second average (on-diet, pre-randomization) is treated as an additional time point. All expected value models in this section have a single cell mean for the second averaged baseline visit and cubic contrasts for the nine post-randomization visits. All models also use a common slope for the covariate, 'average on-diet, pre-randomization measure from visits 3, 4 and 5'.

Evaluation of the Models

A summary of the models with adjusted $-2\lambda$ values is provided in Table 4.15. The same two models as in section 4.4.3, Diggle’s mixed model with compound symmetry plus autoregressive error and the mixed model with random effects plus autoregressive error, failed to converge to final estimates despite repeated attempts with different starting values. Murray’s model with exponentially decreasing correlation and stationary variance consistently provided model estimates for the cases when $\nu$ is set to a constant and when $\tilde{\nu}$ was estimated.

For these data, the general mixed model is providing the best fit ($-2\lambda = 344.7$), with Murray’s autoregressive model with $\tilde{\nu} = 0.2562$ fitting second best ($-2\lambda = 388.0$), followed by the compound symmetry model ($-2\lambda = 612.9$ 1150.5). This is the first set of models in this chapter in which the mixed model has provided the best fit to the data. The classical autoregressive model with $\nu$ set to the constant 1.0 (and also with unequal spacing) provides poor fit, with $-2\lambda = 1150.5$. 

110
The adjusted $-2\lambda$ values for Murray’s model are provided for different values of $\nu$ in Table 4.16. The model with MLE $\hat{\nu} = 0.2562$ provides the best fit of any model as indicated by the lowest $-2\lambda$ value. Table 4.17 has the covariance matrices from the unstructured model, mixed model and Murray’s model with estimated $\hat{\nu}$. Note that for unstructured $\hat{\Sigma}$ the $\text{Var}(\bar{X}) = 302$, which is considerably lower than the other variances, which start at 900. Neither the mixed model, which has $\text{Var}(\bar{X}) = 786$ or Murray’s model with $\text{Var}(\bar{X}) = 1052$, are able to effectively model this low variance. The covariances associated with $\bar{X}$ in the unstructured $\hat{\Sigma}$ (see top row of unstructured $\hat{\Sigma}$) are also quite low; the covariances with $\bar{X}$ in the two structured covariance matrices are not seen to be nearly as low as in the unstructured case.

Graphical Assessment of the Covariance Structures

The covariance plots for unstructured $\hat{\Sigma}$, the mixed model, and Murray’s model can be seen in Figures 4.22 through 4.24, respectively. These plots are very similar to Figures 4.19 through 4.21., where the covariate was ‘cholesterol at visit 1’. Note the longest series, that of the covariances with the ‘mean of visits 3, 4 and 5’ is considerably lower than the rest of the series, which tend to be clumped together. The mixed model, which is providing the best fit according to $-2\lambda$ values, seems to successfully model a slight increasing covariance trend seen in the unstructured $\hat{\Sigma}$. 

111
4.5 Conclusions

4.5.1 When Baseline Values are Treated as Additional Time Points

The main emphasis of this chapter has been evaluating alternative ways to incorporate baseline data into an analysis. Two approaches were taken although several possibilities were discussed in the introduction. The first approach, which will be described in this section, was to simply concatenate the baseline values to $\mathbf{Y}$. This was demonstrated in two ways: (1) all five individual baseline values were concatenated to $\mathbf{Y}$ and (2) the two averaged baseline values (pre-diet and on-diet) were added to $\mathbf{Y}$.

In general, neither of these two models is highly preferred over the other. For each covariance structure, the mixed model, Diggle's model and Murray's model, the covariance plots were quite similar for both approaches. The only real difference was that the covariance plots for the five-baseline models tended to be more complicated because there are more measures made at irregular time points. For both choices of baseline, the autoregressive models, i.e. those from Diggle and Murray, fit about equally well, and the compound symmetry and mixed models fit poorly. For both approaches the estimation algorithm never converged for the mixed model with random effects plus autoregressive error.

The tables below show how the expected values from each model compare for Murray's covariance structure evaluated at the MLE for $\nu$. Notice that the expected mean values are very similar for both models. Therefore, at least for these data, the choice of baseline values has little impact on the expected value part of the model.
Model-based estimates of group means:

<table>
<thead>
<tr>
<th>Model:</th>
<th>Five Baseline Values</th>
<th>Two Mean Baseline Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Visit</td>
<td>Group Mean</td>
</tr>
<tr>
<td>Placebo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>291.89</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td>278.38</td>
<td>293.61</td>
</tr>
<tr>
<td>3</td>
<td>278.71</td>
<td>NA</td>
</tr>
<tr>
<td>4</td>
<td>279.02</td>
<td>279.82</td>
</tr>
<tr>
<td>5</td>
<td>281.75</td>
<td>NA</td>
</tr>
<tr>
<td>7</td>
<td>273.13</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>276.91</td>
<td>276.92</td>
</tr>
<tr>
<td>19</td>
<td>278.84</td>
<td>278.84</td>
</tr>
<tr>
<td>25</td>
<td>279.34</td>
<td>279.36</td>
</tr>
<tr>
<td>31</td>
<td>278.82</td>
<td>278.85</td>
</tr>
<tr>
<td>37</td>
<td>277.67</td>
<td>277.71</td>
</tr>
<tr>
<td>43</td>
<td>276.32</td>
<td>276.34</td>
</tr>
<tr>
<td>49</td>
<td>275.16</td>
<td>275.14</td>
</tr>
<tr>
<td>55</td>
<td>274.61</td>
<td>274.46</td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>293.47</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td>292.00</td>
<td>292.77</td>
</tr>
<tr>
<td>3</td>
<td>279.29</td>
<td>NA</td>
</tr>
<tr>
<td>4</td>
<td>280.02</td>
<td>279.39</td>
</tr>
<tr>
<td>5</td>
<td>278.77</td>
<td>NA</td>
</tr>
<tr>
<td>7</td>
<td>225.22</td>
<td>225.21</td>
</tr>
<tr>
<td>13</td>
<td>239.29</td>
<td>239.37</td>
</tr>
<tr>
<td>19</td>
<td>248.05</td>
<td>248.16</td>
</tr>
<tr>
<td>25</td>
<td>252.63</td>
<td>252.74</td>
</tr>
<tr>
<td>31</td>
<td>254.17</td>
<td>254.26</td>
</tr>
<tr>
<td>37</td>
<td>253.83</td>
<td>253.90</td>
</tr>
<tr>
<td>43</td>
<td>252.74</td>
<td>252.81</td>
</tr>
<tr>
<td>49</td>
<td>252.04</td>
<td>252.16</td>
</tr>
<tr>
<td>55</td>
<td>252.88</td>
<td>253.11</td>
</tr>
</tbody>
</table>
4.5.2 When Baseline Values are Treated as Covariates

The second way in which baseline values were incorporated into the analysis was by including one in the model as a covariate and considering the rest as additional time points. This analysis was also demonstrated in two ways: (1) the individual measure at visit 1 was used as a covariate and the other four individual baseline measures were concatenated to $Y$ and (2) the pre-diet average value was used as the covariate and the on-diet average value was concatenated to $Y$. As in the previous section, no clear choice arises as the ‘best’ method for these data. The estimation algorithm failed to converge for Diggle’s model and the mixed model with random effects plus autoregressive error. Murray’s model produced valid estimates, but provided the best fit for only the model with ‘measure at visit 1’ as the covariate; when the ‘average value of visits 1 and 2’ was the covariate, the mixed model provided better fit.

The tables below show how the expected values from each model compare for Murray’s covariance structure evaluated at the MLE for $\nu$. As was the case in the previous set of tables, the expected mean values are quite similar for both models. The different choices for a covariate baseline value had little impact on the estimates of expected group means.
Covariate model-based estimates of group means:

<table>
<thead>
<tr>
<th>Model Covariate:</th>
<th>Measure at Visit 1</th>
<th>Mean Value of Visits 1 and 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit 1</td>
<td>Group Mean</td>
<td>Group Mean</td>
</tr>
<tr>
<td>1</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td>291.89</td>
<td>NA</td>
</tr>
<tr>
<td>3</td>
<td>278.22</td>
<td>NA</td>
</tr>
<tr>
<td>4</td>
<td>278.52</td>
<td>281.16</td>
</tr>
<tr>
<td>5</td>
<td>281.26</td>
<td>NA</td>
</tr>
<tr>
<td>7</td>
<td>272.63</td>
<td>273.54</td>
</tr>
<tr>
<td>13</td>
<td>276.44</td>
<td>277.35</td>
</tr>
<tr>
<td>19</td>
<td>278.41</td>
<td>279.32</td>
</tr>
<tr>
<td>25</td>
<td>278.95</td>
<td>279.86</td>
</tr>
<tr>
<td>31</td>
<td>278.45</td>
<td>279.36</td>
</tr>
<tr>
<td>37</td>
<td>277.31</td>
<td>278.21</td>
</tr>
<tr>
<td>43</td>
<td>275.93</td>
<td>276.82</td>
</tr>
<tr>
<td>49</td>
<td>274.71</td>
<td>275.57</td>
</tr>
<tr>
<td>55</td>
<td>274.05</td>
<td>274.86</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model Covariate:</th>
<th>Measure at Visit 1</th>
<th>Mean Value of Visits 1 and 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit 1</td>
<td>Group Mean</td>
<td>Group Mean</td>
</tr>
<tr>
<td>1</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td>292.50</td>
<td>NA</td>
</tr>
<tr>
<td>3</td>
<td>279.80</td>
<td>NA</td>
</tr>
<tr>
<td>4</td>
<td>280.52</td>
<td>280.46</td>
</tr>
<tr>
<td>5</td>
<td>279.28</td>
<td>NA</td>
</tr>
<tr>
<td>7</td>
<td>225.75</td>
<td>226.35</td>
</tr>
<tr>
<td>13</td>
<td>239.76</td>
<td>240.39</td>
</tr>
<tr>
<td>19</td>
<td>248.50</td>
<td>249.15</td>
</tr>
<tr>
<td>25</td>
<td>253.10</td>
<td>253.74</td>
</tr>
<tr>
<td>31</td>
<td>254.65</td>
<td>255.30</td>
</tr>
<tr>
<td>37</td>
<td>254.31</td>
<td>254.96</td>
</tr>
<tr>
<td>43</td>
<td>253.18</td>
<td>253.84</td>
</tr>
<tr>
<td>49</td>
<td>252.40</td>
<td>253.07</td>
</tr>
<tr>
<td>55</td>
<td>253.06</td>
<td>253.79</td>
</tr>
</tbody>
</table>
4.5.3 Comparing the Covariate Method to the Additional Time Point Method

Figures 4.25 and 4.26 contain plots of the model-based estimates of group means comparing the two different methods of utilizing baseline values. Each plot has the group means from the covariate method and from the method that treats the baselines as additional time points, for both treatment groups. These plots show that there is virtually no difference in the expected values according to which type of analysis one chooses to use. It should be noted then, that analyses which have (1) included baseline values as extra time points in $Y$ and (2) utilized a covariate type of strategy, have resulted in essentially the same estimates of expected values.

Comparisons of Treatment Differences

Comparisons of the form $\theta = \mathcal{C} \hat{\beta}$ were defined to compare differences of treatment groups at visit 43 (i.e. the seventh 'sixth-month visit'). The idea was to compare both the $\hat{\theta}$ values and their variances from the four models used in this chapter at a time point near the end of the study. These treatment differences can be seen in Figures 4.25 and 4.26. The comparisons from the models are described below. For the models which included baseline values as extra time points in $Y$, the comparisons take the form

$$\hat{\theta} = \mathcal{C} \hat{\beta}$$

$$= [(\hat{\mu}_{p,7} - \hat{\mu}_{p,b}) - (\hat{\mu}_{t,7} - \hat{\mu}_{t,b})],$$

where $\hat{\mu}_{i,j}$ is the model based mean from the $i$-th treatment group ($p =$ placebo, $t =$ treatment) and the $j$-th time point ($7 =$ the seventh sixth-month visit, $b =$ baseline). Models which use a covariate type of strategy have comparisons

$$\hat{\theta} = \mathcal{C} \hat{\beta}$$

$$= [\hat{\mu}_{p,7} - \hat{\mu}_{t,7}],$$

where $\hat{\mu}_{i,j}$ is the model based mean from the $i$-th treatment group ($p =$ placebo,
$t = \text{treatment}$) and the $j$-th time point is the seventh sixth-month visit.

<table>
<thead>
<tr>
<th>Strategy Type</th>
<th>$\hat{\theta}$</th>
<th>Std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Models with Baseline Values Added to $X$:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Individual Baselines</td>
<td>22.20</td>
<td>3.242</td>
</tr>
<tr>
<td>2 Mean Baselines</td>
<td>22.69</td>
<td>3.221</td>
</tr>
<tr>
<td>Models Utilizing Covariates:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measure at Visit 1</td>
<td>22.75</td>
<td>2.926</td>
</tr>
<tr>
<td>Mean Value of Visits 1 &amp; 2</td>
<td>22.98</td>
<td>2.891</td>
</tr>
</tbody>
</table>

All four models produce similar values for $\hat{\theta}$. The models which utilize covariates produced secondary parameter estimates with considerably smaller standard errors.
CHAPTER V
SUMMARY AND SUGGESTIONS FOR FUTURE DEVELOPMENTS

5.1 An Analysis Paradigm

After months of fitting structural covariance models to several different data sets, a strategy has evolved that can be used to analyze incomplete longitudinal data. A series of steps are proposed as a guide for fitting different structural models to a given set of data. It involves the structural covariance models described in Chapter 2 and applied in Chapters 3 and 4. Experience with fitting these models has lead to the proposed analysis paradigm for analyzing data with these structural covariance models.

5.1.1 Analysis Paradigm

1. Fit an unstructured covariance matrix, say $\hat{\Sigma}_U$, to the data. Use a cell mean model for the fixed effects part of the model for simplification. Obtain the within-person covariance and correlation matrices. The statistical software package BMDP-5V can easily provide this information.

2. Make plots of the data to look at the shapes and trends of the covariances, correlations and the estimated group means. The plots are:
   - Covariance plot vs lag time
   - Correlation plot vs lag time
   - Plots of the model-based estimates of group means vs time
3. Evaluate the three plots for:

(1) The behavior of the variances and covariances of $\hat{\Sigma}_U$.

Are the variances equal?

If YES: possible models are Murray's model, Diggle's model and the compound symmetry model.

Do the covariances remain constant with increasing lag?

If Yes: consider the compound symmetry model

Are the covariances decreasing with increasing lag times?

If Yes: consider Murray's and Diggle's model

Are the covariances without any pattern?

If Yes: $\hat{\Sigma}_U$ may not be amenable to modeling

If NO: possible models are the general mixed model and the mixed model with random effects plus autoregressive error.

Are the covariances decreasing with increases lag?

If Yes: consider the mixed model with random effects plus autoregressive error

After an initial sharp decrease in the covariance for early lags, do the covariances begin to increase for longer lags?

If Yes: consider the general mixed model
(2) The behavior of the correlations of $\hat{\Sigma}_U$.

Are the correlations approximately equal?

If **YES**: the compound symmetry model should be considered.

Are the correlations decreasing with increasing lag time?

If **YES**: consider Murray's and Diggle's model

Are the correlations without any pattern?

If **YES**: $\hat{\Sigma}_U$ may not be amenable to modeling

(3) The treatment effects in the group mean plots.

Are there trends that need to be modeled?

If **YES**: consider adding orthogonal contrasts to the design matrix $X$

to model trends.

If **NO**: the cell mean model may be adequate or try reducing $X$ to a
simpler form.

4. Select several models that are likely to fit the data according to the steps
outlined above. Obtain plots of the covariances and correlations for each
model and compare their trends and magnitudes to those found in plots from
$\hat{\Sigma}_U$. Also, compare the $-2\lambda$ values from each model to each other and to
$\hat{\Sigma}_U$.

5. Refit the fixed effects part of the model (if necessary) with the best structural
models from step 4. Re-evaluate the model (with graphs, $-2\lambda$ values, etc.)
to make certain that any changes to the fixed effects in the model have not
resulted in changes in the goodness of fit of the covariance structure.
This paradigm is meant to guide the analyst through the different model choices and can be used to help predict which models should fit a data set. It outlines which types of models are most likely to accommodate certain characteristics of the data typically found in the covariance and correlation matrices. These steps have proven to be empirically accurate and consistent for data sets leading up to, and included in, this dissertation.

The analyst will, however, usually try fitting several different models and make comparisons between these models. This is recommended since reliance on any single analysis paradigm could lead to false conclusions about choice of structural models. Typically, an analysis should start out evaluating models that are easy to obtain and proceed to more esoteric structures.¹ It may turn out that the simple models (also more popular and usually easier to obtain) provide adequate fit. In BMDP-5V, these would be the built-in covariance structures: unstructured \( \mathbb{Q} \), the mixed model, compound symmetry and (if the data are regularly spaced) the AR-1 model. Comparisons of likelihood ratios is convenient and recommended for these models.

The plots described in this paradigm are the same ones used throughout this dissertation to graph estimated covariances, correlations, and model-based means. The plots of covariances and correlations provide a graphical way to judge goodness of fit for any covariance structure. It should be evident from the examples provided in Chapters 3 and 4 that evaluation of goodness of fit of these models is difficult from viewing the actual covariance and correlation matrices alone. We have found it is very difficult to accurately assess trends from these matrices, especially if they are large. Likelihood ratios provide another way to

¹ Murray's and Diggle's model and the mixed model with random effects plus autoregressive error fall into this category.
evaluate model fit, and for these data, they have tended to agree with graphical assessments. Evaluation of covariance model parameter estimates is also recommended to be sure all parameters in the model are necessary. This was discussed in detail in Chapter 3. It should be noted that it is important to use the same expected value model for any two covariance models that are being compared for goodness of fit. This was the strategy used for all model comparisons in this dissertation. Also, each time the fixed effects part of the model is changed, the covariance models under consideration should be refit. This is due to the fact that different covariance models can fit the data better than others for certain fixed effects models. In conclusion, evaluation of goodness of fit using these graphs alone is not recommended, but rather they should be used in conjunction with likelihood ratio comparisons and evaluation of parameter estimates.

5.2 Future Developments

Access to some of the covariance structures used in this dissertation is limited since they are not available to the public in standard software packages. The models from Diggle, Murray, and the mixed model with random effects plus autoregressive error exist in few standard software packages. With user-written FORTRAN subroutines, these models are accessible through BMDP-5V (see also SAS procedure PROC MIXED for other models). Murray’s and Diggle’s model can be specified completely for any data set. The mixed model with random effects plus autoregressive error is somewhat limited because of the difficulty of defining the random effect derivatives for $\mathbf{Z}$. The elements of $\mathbf{Z}$ in this dissertation were limited to an intercept and the time variable, which was the same for all subjects. Individual random effects such as the number of cigarettes
smoked at each time point could not be used for this model.

It would be easy for Murray's model with the curvature parameter $\nu$ estimated, and Diggle's model with $\nu$ a user-supplied constant, to be incorporated into standard software packages. BMDP-5V currently offers a special case of Murray's model, called the 'AR(1)' covariance, with $\nu = 1$ and the assumption that all measures are one unit apart in time. Murray's and Diggle's model could easily be added to BMDP-5V. For irregularly timed data, a vector of times could be supplied by the user. This is precisely how the time points were supplied to the FORTRAN subroutine for this dissertation. The mixed model with random effects plus autoregressive error could also be added to 5V, with the user supplying the random effects for the model, in addition to the vector of time points.

Plots of the covariances and correlations would be a valuable addition to BMDP-5V. These plots could be a user option to aid in model fitting, much like the residual plots provided by regression programs. We have found these plots to be essential in understanding the relationship between the covariances and correlations with lag time between measures. This understanding, in turn, lends insight to choosing the correct covariance structure for the data.

These plots also complement the use of likelihood ratio tests for judging goodness of fit; these tests can be over-discriminating when the number of subjects is large, which can result in an excess of power. Recall in Chapter 3 that the modeled covariance actually provided more believable patterns of the covariance trends than those observed in the unstructured $\hat{G}$. This was despite the fact that likelihood ratio tests concluded the modeled covariance matrix was statistically inferior to the unstructured covariance matrix.

Plots of the model-based predicted means at each time point by grouping
factor level, which are computed directly from $\hat{Y} = X\hat{\beta}$, would also be helpful for modeling the expected value part of the model. The user-supplied time vector could be used to define the x-axis. If plots are not feasible, the model-based predicted means from $\hat{Y}$ could be printed in list format, as is done with predicted values in regression programs.
5.2.1 Some Notes on BMDP-5V

Some peculiarities have been noticed in this software worthy of note to potential analysts. Users fitting non-standard covariance structures which require FORTRAN subroutines should beware that logical errors are not recognized by 5V. For example, if the data have ten time points and are defined as such in the BMDP code, but the FORTRAN statements are written for data with five time points, 5V will proceed to try to fit the model. The resulting error message will typically read ‘too many step halvings’, which is a general message indicative of poor model fit. Mistakes in the BMDP code are generally recognized and identified, although sometimes cryptically. An example would be if the data are defined to have ten repeated measures and contrasts were defined with, say, eight levels. The BMDP editor would indicate that there was a problem in the contrast command line and the user would search the program for the error.

Other Notes:

- For a given covariance structure, some estimation algorithms may fail to converge while others may converge successfully, for the same data. Try several algorithms in the case of non-convergence.

- Most built-in and user-supplied FORTRAN subroutine models converge swiftly (less than 10 iterations) if the model fits properly. If a model is taking 30 or more iterations with step halvings not equal to 0 or 1, users should be concerned with lack of fit and consider other covariance models. One exception is for unstructured $\hat{\Sigma}$ for models in which the number of time points is large, making the number of covariance parameters large. The number of iterations can sometimes be high for these models with unstructured $\hat{\Sigma}$.

- Data which are incomplete are handled well, as long as each time point has (roughly) 1/2 to 3/4 complete data. Studies with excessive mistiming, which will have some time points with almost all missing data,
are not handled well, if at all, by 5V.

- Models which utilize FORTRAN subroutines require user-supplied initial estimates. For simple models (e.g. Murray's model) with regular timing (e.g. models in Chapter 3), the choice of initial estimates was not important for convergence of the estimation algorithms. For complicated models (e.g. mixed model with random effects plus autoregressive error) with irregular timing (e.g. models in Chapter 4), the correct choice of initial estimates was sometimes crucial for convergence.
Table 3.1  
Pattern of Missing Data for Total Cholesterol

<table>
<thead>
<tr>
<th>Month of Measure</th>
<th>Number of Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>106</td>
</tr>
<tr>
<td>18</td>
<td>223</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>42</td>
<td>3</td>
</tr>
<tr>
<td>54</td>
<td>3</td>
</tr>
<tr>
<td>66</td>
<td>6</td>
</tr>
<tr>
<td>78</td>
<td>4</td>
</tr>
<tr>
<td>90</td>
<td>4</td>
</tr>
<tr>
<td>102</td>
<td>1</td>
</tr>
</tbody>
</table>

N 386 380 375 368 361 360 364 360 117 Total N = 393
Table 3.2

Simple Covariance and Correlation Matrices, and Group Means

<table>
<thead>
<tr>
<th>Month</th>
<th>6</th>
<th>18</th>
<th>30</th>
<th>42</th>
<th>54</th>
<th>66</th>
<th>78</th>
<th>90</th>
<th>102</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2018</td>
<td>1372</td>
<td>1245</td>
<td>1174</td>
<td>1239</td>
<td>1226</td>
<td>1109</td>
<td>1249</td>
<td>1636</td>
</tr>
<tr>
<td></td>
<td>1715</td>
<td>1315</td>
<td>1239</td>
<td>1354</td>
<td>1252</td>
<td>1259</td>
<td>1247</td>
<td>1807</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1617</td>
<td>1321</td>
<td>1323</td>
<td>1254</td>
<td>1286</td>
<td>1252</td>
<td>1778</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1719</td>
<td>1329</td>
<td>1232</td>
<td>1356</td>
<td>1320</td>
<td>1836</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covariances</td>
<td>1945</td>
<td>1467</td>
<td>1408</td>
<td>1487</td>
<td>2187</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1756</td>
<td>1451</td>
<td>1459</td>
<td>2095</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1882</td>
<td>1551</td>
<td>2338</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1983</td>
<td>2375</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(sym)</td>
<td>2818</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1.0</th>
<th>.75</th>
<th>.70</th>
<th>.64</th>
<th>.63</th>
<th>.65</th>
<th>.57</th>
<th>.62</th>
<th>.64</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.0</td>
<td>.78</td>
<td>.72</td>
<td>.73</td>
<td>.72</td>
<td>.69</td>
<td>.67</td>
<td>.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>.79</td>
<td>.75</td>
<td>.76</td>
<td>.73</td>
<td>.69</td>
<td>.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>.73</td>
<td>.73</td>
<td>.75</td>
<td>.71</td>
<td>.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlations</td>
<td>1.0</td>
<td>.79</td>
<td>.73</td>
<td>.75</td>
<td>.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>.80</td>
<td>.78</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>.80</td>
<td>.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(sym)</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Group Means for Month In Study

<table>
<thead>
<tr>
<th>Group</th>
<th>6</th>
<th>18</th>
<th>30</th>
<th>42</th>
<th>54</th>
<th>66</th>
<th>78</th>
<th>90</th>
<th>102</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>272.7</td>
<td>274.6</td>
<td>276.1</td>
<td>282.4</td>
<td>278.0</td>
<td>275.1</td>
<td>273.8</td>
<td>275.1</td>
<td>278.5</td>
</tr>
<tr>
<td>N =</td>
<td>198</td>
<td>194</td>
<td>187</td>
<td>185</td>
<td>179</td>
<td>176</td>
<td>177</td>
<td>177</td>
<td>56</td>
</tr>
<tr>
<td>B</td>
<td>224.7</td>
<td>243.5</td>
<td>244.6</td>
<td>251.7</td>
<td>254.6</td>
<td>253.4</td>
<td>250.7</td>
<td>251.3</td>
<td>251.9</td>
</tr>
<tr>
<td>N =</td>
<td>199</td>
<td>196</td>
<td>195</td>
<td>190</td>
<td>185</td>
<td>185</td>
<td>188</td>
<td>184</td>
<td>61</td>
</tr>
</tbody>
</table>
Table 3.3

Estimated Covariance and Correlation Matrices From the Cell Mean Model

Unstructured Covariance Matrix

<table>
<thead>
<tr>
<th>Month</th>
<th>6</th>
<th>18</th>
<th>30</th>
<th>42</th>
<th>54</th>
<th>66</th>
<th>78</th>
<th>90</th>
<th>102</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1400</td>
<td>1023</td>
<td>871</td>
<td>843</td>
<td>944</td>
<td>948</td>
<td>856</td>
<td>959</td>
<td>940</td>
</tr>
<tr>
<td></td>
<td>1495</td>
<td>1066</td>
<td>1024</td>
<td>1163</td>
<td>1101</td>
<td>1080</td>
<td>1056</td>
<td>1089</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1375</td>
<td>1108</td>
<td>1160</td>
<td>1126</td>
<td>1116</td>
<td>1077</td>
<td>1092</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1522</td>
<td>1179</td>
<td>1127</td>
<td>1201</td>
<td>1159</td>
<td>1203</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Covariances

\[
\begin{pmatrix}
1799 & 1338 & 1274 & 1342 & 1429 \\
1653 & 1361 & 1350 & 1398 &     \\
1776 & 1423 & 1484 &     &     \\
1856 &     & 1587 &     & 2091
\end{pmatrix}
\]

(sym)

Correlations

\[
\begin{pmatrix}
1.0 & .71 & .63 & .58 & .59 & .62 & .54 & .59 & .55 \\
1.0 & .74 & .68 & .71 & .70 & .66 & .63 & .61 &     \\
1.0 & .77 & .74 & .75 & .71 & .67 & .64 &     &     \\
1.0 & .71 & .71 & .73 & .69 & .67 &     &     &     \\
1.0 & .77 & .71 & .73 & .74 &     &     &     &     \\
1.0 & .79 & .77 & .75 &     &     &     &     &     \\
1.0 & .78 & .77 &     &     &     &     &     &     \\
1.0 & .81 &     &     &     &     &     &     &     \\
\end{pmatrix}
\]

(sym)

1.0
Table 3.4

Estimated Covariance and Correlation Matrices From the Cell Mean Model

Equal Variances and Unstructured Covariances

<table>
<thead>
<tr>
<th>Month</th>
<th>6</th>
<th>18</th>
<th>30</th>
<th>42</th>
<th>54</th>
<th>66</th>
<th>78</th>
<th>90</th>
<th>102</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1624</td>
<td>1189</td>
<td>1073</td>
<td>980</td>
<td>979</td>
<td>1037</td>
<td>890</td>
<td>965</td>
<td>859</td>
</tr>
<tr>
<td></td>
<td>1624</td>
<td>1241</td>
<td>1130</td>
<td>1153</td>
<td>1151</td>
<td>1074</td>
<td>1017</td>
<td>966</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1624</td>
<td>1270</td>
<td>1194</td>
<td>1224</td>
<td>1153</td>
<td>1075</td>
<td>994</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1624</td>
<td>1150</td>
<td>1157</td>
<td>1177</td>
<td>1097</td>
<td>1040</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covariances</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1624</td>
<td>1242</td>
<td>1132</td>
<td>1161</td>
<td>1147</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1624</td>
<td>1271</td>
<td>1222</td>
<td>1167</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1624</td>
<td>1240</td>
<td>1209</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1624</td>
</tr>
<tr>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|       | 1.0  | .73  | .66  | .60  | .60  | .64  | .55  | .59  | .53  |
|       | 1.0  | .76  | .70  | .71  | .71  | .66  | .63  | .60  |      |
|       | 1.0  | .78  | .74  | .75  | .71  | .66  | .66  | .61  |      |
|       | 1.0  | .71  | .71  | .72  | .72  | .68  | .68  | .64  |      |
| Correlations |      |      |      |      |      |      |      |      |      |
|       | 1.0  | .77  | .70  | .71  | .71  |      |      |      |      |
|       | 1.0  | .78  | .75  | .72  |      |      |      |      |      |
|       | 1.0  | .76  | .74  |      |      |      |      |      |      |
|       |      |      |      |      |      |      |      |      | 1.0  |
| (sym) |      |      |      |      |      |      |      |      |      |

130
Table 3.5
Estimated Covariance Matrices Using the Cell Mean Model

<table>
<thead>
<tr>
<th>Mixed Model</th>
<th>1418</th>
<th>995</th>
<th>992</th>
<th>990</th>
<th>988</th>
<th>985</th>
<th>983</th>
<th>981</th>
<th>979</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1426</td>
<td>1014</td>
<td>1023</td>
<td>1033</td>
<td>1043</td>
<td>1052</td>
<td>1062</td>
<td>1071</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1457</td>
<td>1057</td>
<td>1078</td>
<td>1100</td>
<td>1121</td>
<td>1143</td>
<td>1164</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1512</td>
<td>1124</td>
<td>1157</td>
<td>1190</td>
<td>1224</td>
<td>1257</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1591</td>
<td>1214</td>
<td>1260</td>
<td>1305</td>
<td>1350</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1693</td>
<td>1329</td>
<td>1386</td>
<td>1443</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1819</td>
<td>1467</td>
<td>1536</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1969</td>
<td>1629</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2144</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Compound Symmetry</th>
<th>1626</th>
<th>1127</th>
<th>1127</th>
<th>1127</th>
<th>1127</th>
<th>1127</th>
<th>1127</th>
<th>1127</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1626</td>
<td>1127</td>
<td>1127</td>
<td>1127</td>
<td>1127</td>
<td>1127</td>
<td>1127</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1626</td>
<td>1127</td>
<td>1127</td>
<td>1127</td>
<td>1127</td>
<td>1127</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1626</td>
<td>1127</td>
<td>1127</td>
<td>1127</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1626</td>
<td>1127</td>
<td>1127</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1626</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ar(1)</th>
<th>1642</th>
<th>1253</th>
<th>955</th>
<th>728</th>
<th>556</th>
<th>424</th>
<th>323</th>
<th>247</th>
<th>188</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1642</td>
<td>1253</td>
<td>955</td>
<td>728</td>
<td>556</td>
<td>424</td>
<td>323</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1642</td>
<td>1253</td>
<td>955</td>
<td>728</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1642</td>
<td>1253</td>
<td>955</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1642</td>
</tr>
</tbody>
</table>
Table 3.6
Estimated Correlation Matrices Using the Cell Mean Model

<table>
<thead>
<tr>
<th>Mixed Model</th>
<th>1.0</th>
<th>.70</th>
<th>.69</th>
<th>.68</th>
<th>.66</th>
<th>.64</th>
<th>.61</th>
<th>.59</th>
<th>.56</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>.70</td>
<td>.70</td>
<td>.69</td>
<td>.67</td>
<td>.65</td>
<td>.63</td>
<td>.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.71</td>
<td>.71</td>
<td>.70</td>
<td>.69</td>
<td>.67</td>
<td>.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.72</td>
<td>.72</td>
<td>.72</td>
<td>.71</td>
<td>.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.72</td>
<td>.74</td>
<td>.74</td>
<td>.74</td>
<td>.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.76</td>
<td>.76</td>
<td>.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.77</td>
<td>.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(sym)

For compound symmetry model, \( \hat{\rho} \) is fixed at 0.69.

<table>
<thead>
<tr>
<th>AR(1)</th>
<th>1.0</th>
<th>.76</th>
<th>.58</th>
<th>.44</th>
<th>.34</th>
<th>.26</th>
<th>.20</th>
<th>.15</th>
<th>.11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>.76</td>
<td>.58</td>
<td>.44</td>
<td>.34</td>
<td>.26</td>
<td>.20</td>
<td>.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.76</td>
<td>.58</td>
<td>.44</td>
<td>.34</td>
<td>.26</td>
<td>.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.76</td>
<td>.58</td>
<td>.44</td>
<td>.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.76</td>
<td>.58</td>
<td>.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(sym)</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

132
Table 3.7

Model Parameters for the Cell Mean Model

Model 1: Mixed Model

Maximum log Likelihood = −14386.2

\[
\hat{\Sigma} = \mathbf{Z} \begin{bmatrix} 1013.3 & -14.2 \\ -14.2 & 11.9 \end{bmatrix} \mathbf{Z}' + 421.6 \mathbf{I}_9
\]

Covariance Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \delta_{11} )</td>
<td>1013.3</td>
<td>91.08</td>
<td>11.1</td>
</tr>
<tr>
<td>( \delta_{12} )</td>
<td>-14.2</td>
<td>9.00</td>
<td>-1.6</td>
</tr>
<tr>
<td>( \delta_{22} )</td>
<td>11.9</td>
<td>1.57</td>
<td>7.6</td>
</tr>
<tr>
<td>( \sigma^2_e )</td>
<td>421.6</td>
<td>12.42</td>
<td>34.0</td>
</tr>
</tbody>
</table>

Model 2: Compound Symmetry

Maximum log Likelihood = −14469.3

\[
\hat{\Sigma} = 498.7 \mathbf{I} + 1127.5 \mathbf{I}
\]

Covariance Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma^2 )</td>
<td>498.7</td>
<td>13.62</td>
<td>36.6</td>
</tr>
<tr>
<td>( \sigma^2 )</td>
<td>1127.5</td>
<td>85.52</td>
<td>13.2</td>
</tr>
</tbody>
</table>

Model 3: AR(1)

Maximum log Likelihood = −14574.4

\[
\hat{\Sigma} = \begin{bmatrix} < 1642.2 (0.76) \big| t_j - t_k \big| > \end{bmatrix}_{jk}, \text{for the } j\text{-th row and } k\text{-th column element.}
\]

Covariance Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma^2 )</td>
<td>1642.2</td>
<td>69.57</td>
<td>23.6</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.76</td>
<td>0.01</td>
<td>71.3</td>
</tr>
<tr>
<td>Mixed Model</td>
<td>1422 993 991 989 987 985 983 982 980</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1430 1013 1022 1032 1042 1052 1062 1072</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1461 1056 1077 1099 1120 1142 1163</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1516 1122 1155 1189 1222 1255</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1594 1212 1257 1302 1347</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1696 1326 1382 1439</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1821 1462 1531</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1970 1623</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2142</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compound</td>
<td>1630 1127 1127 1127 1127 1127 1127 1127 1127</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1630 1127 1127 1127 1127 1127 1127 1127</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1630 1127 1127 1127 1127</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1630 1127 1127</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1630 1127</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1630</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1630</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symmetry</td>
<td>(sym)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ar(1)</td>
<td>1643 1247 946 718 545 413 314 238 181</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1643 1247 946 718 545 413 314 238</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1643 1247 946 718</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1643 1247 946</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1643</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(sym)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1643</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.9
Estimated Correlation Matrices Using the Cubic Model

Mixed Model

\[
\begin{bmatrix}
1.0 & .70 & .69 & .67 & .66 & .63 & .61 & .59 & .56 \\
1.0 & .70 & .69 & .68 & .67 & .65 & .63 & .61 \\
1.0 & .71 & .71 & .70 & .69 & .67 & .66 \\
1.0 & .72 & .72 & .72 & .71 & .70 \\
1.0 & .74 & .74 & .73 & .73 \\
1.0 & .75 & .76 & .76 \\
1.0 & .77 & .78 \\
1.0 & .79 \\
\end{bmatrix}
\]

(sym)

For compound symmetry model, \( \hat{\rho} \) is fixed at 0.69.

AR(1)

\[
\begin{bmatrix}
1.0 & .76 & .58 & .44 & .33 & .25 & .19 & .14 & .11 \\
1.0 & .76 & .58 & .44 & .33 & .25 & .19 & .14 \\
1.0 & .76 & .58 & .44 & .33 & .25 \\
1.0 & .76 & .58 & .44 \\
1.0 & .76 \\
\end{bmatrix}
\]

(sym)

1.0
Table 3.10

Estimated Covariance and Correlation Matrices
Using the Cubic Model with Unstructured Covariance

Covariances

\[
\begin{bmatrix}
1401 & 1023 & 871 & 842 & 945 & 949 & 858 & 959 & 937 \\
1510 & 1060 & 1022 & 1169 & 1106 & 1078 & 1054 & 1090 \\
1384 & 1105 & 1157 & 1126 & 1121 & 1073 & 1089 \\
1524 & 1176 & 1121 & 1200 & 1158 & 1192 \\
1800 & 1340 & 1274 & 1340 & 1422 \\
1656 & 1362 & 1347 & 1395 \\
1780 & 1421 & 1484 \\
1854 & 1569 & & & & & & 2076
\end{bmatrix}
\]

Correlations

\[
\begin{bmatrix}
1.0 & .70 & .62 & .58 & .59 & .62 & .54 & .59 & .55 \\
1.0 & .73 & .67 & .71 & .70 & .66 & .63 & .62 \\
1.0 & .76 & .73 & .74 & .71 & .70 & .64 \\
1.0 & .71 & .71 & .73 & .69 & .67 \\
1.0 & .77 & .71 & .73 & .74 \\
1.0 & .79 & .77 & .75 \\
1.0 & .78 & .77 & 1.0 & .80 \\
1.0 & 1.0 & & & & & & 1.0
\end{bmatrix}
\]
Table 3.11

Model Parameters for the Cubic Model

Model 1: Mixed Model

Maximum log Likelihood = $-14400.5$

$$\hat{\Sigma} = \hat{\Sigma} \left[ \begin{array}{cc} 1010.4 & -13.6 \\ -13.6 & 11.7 \end{array} \right] \hat{\Sigma} + 427.1 \mathbb{I}_9$$

Covariance Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{11}$</td>
<td>1010.4</td>
<td>91.11</td>
<td>11.1</td>
</tr>
<tr>
<td>$\delta_{12}$</td>
<td>-13.6</td>
<td>8.99</td>
<td>-1.5</td>
</tr>
<tr>
<td>$\delta_{22}$</td>
<td>11.7</td>
<td>1.56</td>
<td>7.5</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>451.1</td>
<td>12.58</td>
<td>33.9</td>
</tr>
</tbody>
</table>

Model 2: Compound Symmetry

Maximum log Likelihood = $-14481.0$

$$\hat{\Sigma} = 503.1 \mathbb{I} + 1126.5 \mathbb{J}$$

Covariance Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>503.1</td>
<td>13.74</td>
<td>36.6</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>1126.5</td>
<td>85.49</td>
<td>13.2</td>
</tr>
</tbody>
</table>

Model 3: AR(1)

Maximum log Likelihood = $-14593.1$

$$\hat{\Sigma} = \left[ < 1642.6 (0.76) | t_j - t_k > jk \right], \text{ for the } j\text{-th row and } k\text{-th column element.}$$

Covariance Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>1642.6</td>
<td>69.19</td>
<td>23.7</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.76</td>
<td>0.01</td>
<td>70.3</td>
</tr>
</tbody>
</table>
## Table 3.12

Estimated Covariance and Correlation Matrices

Model With Stationary Variance and Exponentially Decreasing Correlation, \( \nu = 1 \)

### Cubic Model

<table>
<thead>
<tr>
<th></th>
<th>1643</th>
<th>1247</th>
<th>946</th>
<th>718</th>
<th>545</th>
<th>413</th>
<th>314</th>
<th>283</th>
<th>181</th>
</tr>
</thead>
<tbody>
<tr>
<td>1643</td>
<td>1247</td>
<td>946</td>
<td>718</td>
<td>545</td>
<td>413</td>
<td>314</td>
<td>283</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>545</td>
<td>413</td>
<td>314</td>
<td>283</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>718</td>
<td>1643</td>
<td>1247</td>
<td>946</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1.0</th>
<th>.76</th>
<th>.58</th>
<th>.44</th>
<th>.33</th>
<th>.25</th>
<th>.19</th>
<th>.14</th>
<th>.11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>.76</td>
<td>.58</td>
<td>.44</td>
<td>.33</td>
<td>.25</td>
<td>.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.33</td>
<td></td>
<td></td>
<td>.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.44</td>
<td></td>
<td></td>
<td>.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
<td>.76</td>
</tr>
</tbody>
</table>

### Cell Mean

<table>
<thead>
<tr>
<th></th>
<th>1642</th>
<th>1253</th>
<th>955</th>
<th>729</th>
<th>556</th>
<th>424</th>
<th>323</th>
<th>247</th>
<th>188</th>
</tr>
</thead>
<tbody>
<tr>
<td>1642</td>
<td>1253</td>
<td>955</td>
<td>729</td>
<td>556</td>
<td>424</td>
<td>323</td>
<td>247</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>556</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>729</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1642</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1.0</th>
<th>.76</th>
<th>.58</th>
<th>.44</th>
<th>.34</th>
<th>.26</th>
<th>.20</th>
<th>.15</th>
<th>.11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>.76</td>
<td>.58</td>
<td>.44</td>
<td>.34</td>
<td>.26</td>
<td>.20</td>
<td>.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.34</td>
<td></td>
<td></td>
<td>.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
<td>.76</td>
</tr>
<tr>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>
Table 3.13

Estimated Covariance and Correlation Matrices
Model With Stationary Variance and Exponentially Decreasing Correlation, $\nu = 2$

**Cubic Model**

\[
\begin{bmatrix}
1382 & 675 & 79 & 2 & 0 & 0 & 0 & 0 & 0 \\
1382 & 675 & 79 & 2 & 0 & 0 & 0 & 0 & 0 \\
& & & & & & & & \\
& & & & & & & & \\
& & & & & & & & \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
\text{(sym)} \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1.0 & .448 & .057 & .002 & 0 & 0 & 0 & 0 & 0 \\
1.0 & .448 & .057 & .002 & 0 & 0 & 0 & 0 & 0 \\
& & & & & & & & \\
& & & & & & & & \\
& & & & & & & & \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
\text{(sym)} \\
\end{bmatrix}
\]

**Cell Mean**

\[
\begin{bmatrix}
1300 & 677 \\
& & & & & & & & \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
\text{(sym)} \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1.0 & .76 & .58 & 0 & 0 & 0 & 0 & 0 & 0 \\
1.0 & .76 & .58 & 0 & 0 & 0 & 0 & 0 & 0 \\
& & & & & & & & \\
& & & & & & & & \\
& & & & & & & & \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
\text{(sym)} \\
\end{bmatrix}
\]
Table 3.14

Estimated Covariance and Correlation Matrices
Model With Stationary Variance and Exponentially Decreasing Correlation, $\nu = \tilde{\nu}$

Cubic Model, $\tilde{\nu} = 0.28$

\[
\begin{bmatrix}
1630 & 1230 & 1159 & 1112 & 1078 & 1049 & 1026 & 1005 & 987 \\
1630 & 1230 & 1159 & 1112 & 1078 & 1049 & 1026 & 1005 & 1026 \\
& & & & & & & 1049 & 1078 \\
& & & & & & & 1112 & 1159 \\
& & & & & & & 1630 & 1230 \\
& & & & & & & & 1630 \\
\end{bmatrix}
\]

Covariances (sym)

\[
\begin{bmatrix}
1.0 & .75 & .71 & .68 & .66 & .64 & .63 & .62 & .60 \\
1.0 & .75 & .71 & .68 & .66 & .64 & .63 & .62 & .63 \\
& & & & & & & .64 & .66 \\
& & & & & & & .68 & 1.0 \\
& & & & & & & .71 & .75 \\
& & & & & & & 1.0 & .75 \\
\end{bmatrix}
\]

Correlations (sym)


Cell Mean, $\tilde{\nu} = 0.29$

\[
\begin{bmatrix}
1628 & 1234 & 1161 & 1114 & 1079 & 1050 & 1025 & 1004 & 985 \\
1628 & 1234 & 1161 & 1114 & 1079 & 1050 & 1025 & 1004 & 1025 \\
& & & & & & & 1050 & 1079 \\
& & & & & & & 1114 & 1161 \\
& & & & & & & 1628 & 1234 \\
& & & & & & & 1628 & 1234 \\
\end{bmatrix}
\]

Covariances (sym)

\[
\begin{bmatrix}
1.0 & .76 & .71 & .68 & .66 & .64 & .63 & .62 & .60 \\
1.0 & .76 & .71 & .68 & .66 & .64 & .63 & .62 & .63 \\
& & & & & & & .64 & .66 \\
& & & & & & & .68 & 1.0 \\
& & & & & & & .71 & .76 \\
& & & & & & & 1.0 & .76 \\
\end{bmatrix}
\]

Correlations (sym)
Table 3.15
Covariance Parameter Estimates and Standard Errors for Models with Stationary Variance and Exponential Decreasing Correlation

**Cubic Model**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>AR(1), $\nu = 1$</th>
<th>AR(1), $\nu = 2$</th>
<th>AR(1), $\nu$ estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>NA (69.20)$^1$</td>
<td>1642.6 (800)</td>
<td>1382.5 (40.20)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>NA (0.011)</td>
<td>0.76 (0.50)</td>
<td>0.49 (0.008)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>1.0* 1.0</td>
<td>2.0* 2.0</td>
<td>0.50 0.28 (0.028)</td>
</tr>
</tbody>
</table>

**Cell Mean Model**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>AR(1), $\nu = 1$</th>
<th>AR(1), $\nu = 2$</th>
<th>AR(1), $\nu$ estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>NA (69.57)$^1$</td>
<td>1642.2 (1000)</td>
<td>1376.4 (40.05)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>NA (0.011)</td>
<td>0.76 (0.25)</td>
<td>0.49 (0.008)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>1.0* 1.0</td>
<td>2.0* 2.0</td>
<td>0.5 0.29 (0.028)</td>
</tr>
</tbody>
</table>

* $\nu$ is set to a constant for these models

$^1$ (Standard Error)

Note: all covariance parameters with $p < 0.001$
Table 3.16
Estimated Covariance and Correlation Matrices
Models With Compound Symmetry Plus an Autoregressive Error, $\nu = 1$

<table>
<thead>
<tr>
<th>Cubic Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariances</td>
</tr>
</tbody>
</table>
| \[
\begin{bmatrix}
1634 & 1228 & 1179 & 1132 & 1086 & 1042 & 998 & 956 & 915 \\
1634 & 1228 & 1179 & 1132 & 1086 & 1042 & 998 & 956 & 915 \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \ddots & \ddots \\
1634 & 1228 & 1179 & 1132 & 1086 & 1042 & 998 & 956 & 915 \\
\end{bmatrix}
\]

| (sym) |
| Correlations |
| \[
\begin{bmatrix}
1.0 & .75 & .72 & .69 & .66 & .64 & .61 & .59 & .56 \\
1.0 & .75 & .72 & .69 & .66 & .64 & .61 & .59 & .56 \\
\vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\
1.0 & .75 & .72 & .69 & .66 & .64 & .61 & .59 & .56 \\
\end{bmatrix}
\]

| (sym) |

Cell Mean
The estimation algorithm did not converge.
Table 3.17

Estimated Covariance and Correlation Matrices
Models With Compound Symmetry With an Autoregressive Error, $\nu = 2$

**Cubic Model**

\[
\begin{bmatrix}
1635 & 1224 & 1188 & 1141 & 1088 & 1036 & 991 & 956 & 931 \\
1635 & 1224 & 1188 & 1141 & 1088 & 1036 & 991 & 956 & 911 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1635 & 1224 & 1188 & 1141 & 1088 & 1036 & 991 & 956 & 1036 \\
1635 & 1224 & 1188 & 1141 & 1088 & 1036 & 991 & 956 & 911 \\
\end{bmatrix}
\]

**Covariances**

\[
\begin{bmatrix}
1.0 & 0.75 & 0.73 & 0.70 & 0.67 & 0.63 & 0.61 & 0.58 & 0.57 \\
1.0 & 0.75 & 0.73 & 0.70 & 0.67 & 0.63 & 0.61 & 0.58 & 0.66 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1.0 & 0.75 & 0.73 & 0.70 & 0.67 & 0.63 & 0.61 & 0.58 & 0.66 \\
1.0 & 0.75 & 0.73 & 0.70 & 0.67 & 0.63 & 0.61 & 0.58 & 1.0 \\
\end{bmatrix}
\]

**Correlations**

\[
\begin{bmatrix}
1632 & 1224 & 1191 & 1142 & 1088 & 1035 & 991 & 956 & 933 \\
1632 & 1224 & 1191 & 1142 & 1088 & 1035 & 991 & 956 & 911 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1632 & 1224 & 1191 & 1142 & 1088 & 1035 & 991 & 956 & 1035 \\
1632 & 1224 & 1191 & 1142 & 1088 & 1035 & 991 & 956 & 911 \\
\end{bmatrix}
\]

**Cell Mean**

\[
\begin{bmatrix}
1.0 & 0.75 & 0.73 & 0.70 & 0.67 & 0.63 & 0.61 & 0.59 & 0.57 \\
1.0 & 0.75 & 0.73 & 0.70 & 0.67 & 0.63 & 0.61 & 0.59 & 0.66 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1.0 & 0.75 & 0.73 & 0.70 & 0.67 & 0.63 & 0.61 & 0.59 & 0.66 \\
1.0 & 0.75 & 0.73 & 0.70 & 0.67 & 0.63 & 0.61 & 0.59 & 1.0 \\
\end{bmatrix}
\]

**Correlations**

\[
\begin{bmatrix}
1632 & 1224 & 1191 & 1142 & 1088 & 1035 & 991 & 956 & 933 \\
1632 & 1224 & 1191 & 1142 & 1088 & 1035 & 991 & 956 & 911 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1632 & 1224 & 1191 & 1142 & 1088 & 1035 & 991 & 956 & 1035 \\
1632 & 1224 & 1191 & 1142 & 1088 & 1035 & 991 & 956 & 911 \\
\end{bmatrix}
\]
Table 3.18

Estimated Covariance and Correlation Matrices

Models With Compound Symmetry With an Autoregressive Error, \( \nu = 3 \)

**Cubic Model**

\[
\begin{bmatrix}
1633 & 1206 & 1190 & 1151 & 1094 & 1034 & 989 & 967 & 959 \\
1633 & 1206 & 1190 & 1151 & 1094 & 1034 & 989 & 967 & 959 \\
& & & & 1094 & 1034 & 989 & 967 & 959 \\
& & & & & 1151 & 1094 & 1034 & 989 & 967 \\
& & & & & & 1633 & 1206 & 1190 & 1151 \\
& & & & & & & 1633 & 1206 & 1190 \\
& & & & & & & & 1633 & 1633 \\
\end{bmatrix}
\]

\((\text{sym})\)

**Correlations**

\[
\begin{bmatrix}
1.0 & .74 & .73 & .71 & .67 & .63 & .61 & .59 & .59 \\
1.0 & .74 & .73 & .71 & .67 & .63 & .61 & .59 & .59 \\
& & & & & .61 & .63 & .67 & .71 \\
& & & & & & .67 & .71 & .73 \\
1.0 & .74 & .73 & .71 & .67 & .63 & .61 & .59 & .59 \\
1.0 & .74 & .73 & .71 & .67 & .63 & .61 & .59 & .59 \\
& & & & & & .61 & .63 & .67 \\
& & & & & & & .71 & .73 \\
\end{bmatrix}
\]

\((\text{sym})\)

**Cell Mean**

\[
\begin{bmatrix}
1629 & 1208 & 1192 & 1153 & 1094 & 1034 & 990 & 968 & 961 \\
1629 & 1208 & 1192 & 1153 & 1094 & 1034 & 990 & 968 & 961 \\
& & & & & 1094 & 1034 & 990 & 968 \\
& & & & & & 1153 & 1094 & 1034 \\
& & & & & & & 1629 & 1208 \\
& & & & & & & & 1629 \\
\end{bmatrix}
\]

\((\text{sym})\)

**Correlations**

\[
\begin{bmatrix}
1.0 & .74 & .73 & .71 & .67 & .63 & .61 & .59 & .59 \\
1.0 & .74 & .73 & .71 & .67 & .63 & .61 & .59 & .59 \\
& & & & & & .61 & .63 & .67 \\
& & & & & & .67 & .71 & .73 \\
1.0 & .74 & .73 & .71 & .67 & .63 & .61 & .59 & .59 \\
1.0 & .74 & .73 & .71 & .67 & .63 & .61 & .59 & .59 \\
& & & & & & .61 & .63 & .67 \\
& & & & & & & .71 & .73 \\
\end{bmatrix}
\]

\((\text{sym})\)
The estimation algorithm did not converge for the model with $\nu = 1$.

**Model with $\nu = 2$, $-2\lambda = 778.3$**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Starting Estimate</th>
<th>Parameter Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau^2$</td>
<td>400</td>
<td>395.7</td>
<td>12.99</td>
<td>30.5</td>
</tr>
<tr>
<td>$\omega^2$</td>
<td>800</td>
<td>899.8</td>
<td>106.66</td>
<td>8.4</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>300</td>
<td>336.6</td>
<td>69.84</td>
<td>4.8</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.5</td>
<td>0.04</td>
<td>0.011</td>
<td>3.2</td>
</tr>
</tbody>
</table>

**Model with $\nu = 3$, $-2\lambda = 786.9$**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Starting Estimate</th>
<th>Parameter Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau^2$</td>
<td>400</td>
<td>419.1</td>
<td>12.36</td>
<td>30.9</td>
</tr>
<tr>
<td>$\omega^2$</td>
<td>100</td>
<td>959.3</td>
<td>90.29</td>
<td>10.6</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>300</td>
<td>251.0</td>
<td>38.05</td>
<td>6.6</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.5</td>
<td>0.01</td>
<td>0.002</td>
<td>4.3</td>
</tr>
</tbody>
</table>
Table 3.20
Covariance Parameter Estimates for Models with Compound Symmetry
Plus an Autoregressive Error – Cubic Model

Model with $\nu = 1$, $-2\lambda = 804.9$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Starting Estimate</th>
<th>Parameter Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau^2$</td>
<td>400</td>
<td>355.9</td>
<td>18.85</td>
<td>18.9</td>
</tr>
<tr>
<td>$\omega^2$</td>
<td>900</td>
<td>-479.8</td>
<td>4262.78</td>
<td>-0.1</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>300</td>
<td>1757.9</td>
<td>4253.84</td>
<td>0.4</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.1</td>
<td>0.03</td>
<td>0.076</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Model with $\nu = 2$, $-2\lambda = 808.3$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Starting Estimate</th>
<th>Parameter Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau^2$</td>
<td>400</td>
<td>402.4</td>
<td>13.18</td>
<td>30.5</td>
</tr>
<tr>
<td>$\omega^2$</td>
<td>900</td>
<td>895.2</td>
<td>108.87</td>
<td>8.2</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>300</td>
<td>337.1</td>
<td>73.17</td>
<td>4.6</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.1</td>
<td>0.03</td>
<td>0.011</td>
<td>3.1</td>
</tr>
</tbody>
</table>

Model with $\nu = 3$, $-2\lambda = 816.6$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Starting Estimate</th>
<th>Parameter Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau^2$</td>
<td>400</td>
<td>424.7</td>
<td>12.73</td>
<td>33.9</td>
</tr>
<tr>
<td>$\omega^2$</td>
<td>100</td>
<td>956.6</td>
<td>90.57</td>
<td>10.6</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>300</td>
<td>251.3</td>
<td>38.88</td>
<td>6.5</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.1</td>
<td>0.01</td>
<td>0.002</td>
<td>4.2</td>
</tr>
</tbody>
</table>
Table 3.21

Estimated Covariance and Correlation Matrices
Mixed Models With Random Effects Plus an Autoregressive Error, $\nu = 1$

**Cubic Model**

\[
\begin{bmatrix}
1399 & 1010 & 961 & 961 & 967 & 974 & 981 & 989 & 996 \\
1423 & 1043 & 1003 & 1013 & 1028 & 1045 & 1061 & 1078 & 1078 \\
1446 & 1095 & 1065 & 1083 & 1108 & 1033 & 1159 & 1159 & 1159 \\
1527 & 1165 & 1146 & 1172 & 1206 & 1241 & 1241 & 1241 & 1241 \\
1607 & 1254 & 1243 & 1279 & 1323 & 1362 & 1360 & 1405 & 1405 \\
1705 & 1362 & 1360 & 1405 & 1459 & 1459 & 1459 & 1459 & 1459 \\
1822 & 1488 & 1488 & 1488 & 1488 & 1488 & 1488 & 1488 & 1488 \\
1957 & 1632 & 1632 & 1632 & 1632 & 1632 & 1632 & 1632 & 1632 \\
\end{bmatrix}
\]

(sym)

\[
\begin{bmatrix}
1.0 & .72 & .67 & .66 & .65 & .63 & .61 & .60 & .58 \\
1.0 & .72 & .68 & .67 & .66 & .65 & .64 & .62 & .62 \\
1.0 & .73 & .69 & .68 & .68 & .68 & .67 & .66 & .66 \\
1.0 & .74 & .71 & .70 & .70 & .70 & .69 & .69 & .69 \\
1.0 & .76 & .73 & .72 & .72 & .72 & .72 & .72 & .72 \\
1.0 & .77 & .74 & .74 & .74 & .74 & .74 & .74 & .74 \\
1.0 & .79 & .76 & .76 & .76 & .76 & .76 & .76 & .76 \\
1.0 & .80 & .80 & .80 & .80 & .80 & .80 & .80 & .80 \\
\end{bmatrix}
\]

(sym)

\[
\begin{bmatrix}
1395 & 1012 & 967 & 961 & 967 & 974 & 981 & 989 & 996 \\
1419 & 1045 & 1004 & 1013 & 1028 & 1045 & 1061 & 1078 & 1078 \\
1462 & 1097 & 1066 & 1084 & 1108 & 1135 & 1160 & 1160 & 1160 \\
1523 & 1168 & 1146 & 1173 & 1207 & 1243 & 1243 & 1243 & 1243 \\
1604 & 1258 & 1245 & 1282 & 1325 & 1325 & 1325 & 1325 & 1325 \\
1703 & 1366 & 1363 & 1409 & 1409 & 1409 & 1409 & 1409 & 1409 \\
1820 & 1494 & 1500 & 1500 & 1500 & 1500 & 1500 & 1500 & 1500 \\
1957 & 1639 & 2112 & 2112 & 2112 & 2112 & 2112 & 2112 & 2112 \\
\end{bmatrix}
\]

(sym)

\[
\begin{bmatrix}
1.0 & .72 & .67 & .66 & .65 & .63 & .62 & .60 & .58 \\
1.0 & .73 & .68 & .67 & .66 & .65 & .64 & .62 & .62 \\
1.0 & .74 & .70 & .69 & .68 & .67 & .66 & .66 & .66 \\
1.0 & .75 & .71 & .70 & .70 & .69 & .69 & .69 & .69 \\
1.0 & .76 & .73 & .72 & .72 & .72 & .72 & .72 & .72 \\
1.0 & .78 & .75 & .74 & .74 & .74 & .74 & .74 & .74 \\
1.0 & .79 & .76 & .76 & .76 & .76 & .76 & .76 & .76 \\
1.0 & .81 & .81 & .81 & .81 & .81 & .81 & .81 & .81 \\
\end{bmatrix}
\]

(sym)
Table 3.22

Estimated Covariance and Correlation Matrices
Mixed Models With Random Effects Plus an Autoregressive Error, \( \nu = 2 \)

### Cubic Model

<table>
<thead>
<tr>
<th></th>
<th>1401</th>
<th>1011</th>
<th>957</th>
<th>964</th>
<th>970</th>
<th>976</th>
<th>982</th>
<th>989</th>
<th>995</th>
</tr>
</thead>
<tbody>
<tr>
<td>1423</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1465</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1525</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covariances</td>
<td>[          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1.0</th>
<th>.72</th>
<th>.67</th>
<th>.66</th>
<th>.65</th>
<th>.63</th>
<th>.61</th>
<th>.60</th>
<th>.58</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlations</td>
<td>[          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Cell Mean

<table>
<thead>
<tr>
<th></th>
<th>1397</th>
<th>1013</th>
<th>958</th>
<th>964</th>
<th>970</th>
<th>976</th>
<th>982</th>
<th>988</th>
<th>995</th>
</tr>
</thead>
<tbody>
<tr>
<td>1419</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1461</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1522</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covariances</td>
<td>[          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1.0</th>
<th>.72</th>
<th>.67</th>
<th>.66</th>
<th>.65</th>
<th>.63</th>
<th>.62</th>
<th>.60</th>
<th>.58</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlations</td>
<td>[          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ] &amp; [          ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ] &amp; [         ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

148
Table 3.23

Estimated Covariance and Correlation Matrices
Mixed Models With Random Effects Plus an Autoregressive Error, $\nu = \bar{v}$

**Cubic Model, $\bar{v} = 0.50$**

\[
\begin{bmatrix}
1392 & 1008 & 971 & 962 & 962 & 968 & 977 & 987 & 998 \\
1428 & 1051 & 1021 & 1019 & 1028 & 1049 & 1057 & 1074 & \\
1478 & 1108 & 1086 & 1092 & 1108 & 1128 & 1152 & \\
1544 & 1181 & 1167 & 1180 & 1203 & 1231 & \\
1624 & 1269 & 1262 & 1282 & 1313 & \\
1719 & 1371 & 1371 & 1399 & \\
1828 & 1488 & 1496 & \\
1953 & 1620 & \\
\end{bmatrix}
\]

\[
\text{(sym)}
\]

\[
\begin{bmatrix}
1.0 & .71 & .68 & .66 & .64 & .63 & .61 & .60 & .58 \\
1.0 & .72 & .69 & .67 & .66 & .64 & .63 & .62 & \\
1.0 & .73 & .70 & .69 & .67 & .66 & .65 & \\
1.0 & .75 & .72 & .70 & .69 & .68 & \\
1.0 & .76 & .73 & .72 & .71 & \\
1.0 & .77 & .75 & .74 & \\
1.0 & .79 & .76 & \\
1.0 & .80 & \\
\end{bmatrix}
\]

\[
\text{(sym)}
\]

**Cell Mean, $\bar{v} = 0.44$**

\[
\begin{bmatrix}
1380 & 1003 & 966 & 956 & 957 & 962 & 970 & 980 & 992 \\
1417 & 1047 & 1018 & 1015 & 1023 & 1035 & 1051 & 1068 & \\
1469 & 1106 & 1084 & 1088 & 1103 & 1023 & 1145 & \\
1535 & 1179 & 1164 & 1176 & 1197 & 1224 & \\
1615 & 1266 & 1259 & 1277 & 1306 & \\
1710 & 1368 & 1368 & 1394 & \\
1819 & 1485 & 1491 & \\
1942 & 1615 & \\
2080 & \\
\end{bmatrix}
\]

\[
\text{(sym)}
\]

\[
\begin{bmatrix}
1.0 & .72 & .68 & .66 & .64 & .63 & .62 & .60 & .59 \\
1.0 & .73 & .69 & .67 & .66 & .64 & .63 & .62 & \\
1.0 & .74 & .70 & .69 & .67 & .66 & .66 & \\
1.0 & .75 & .72 & .70 & .69 & .69 & \\
1.0 & .76 & .73 & .72 & .71 & \\
1.0 & .78 & .75 & .74 & \\
1.0 & .79 & .77 & \\
1.0 & .80 & \\
\end{bmatrix}
\]

\[
\text{(sym)}
\]

149
Table 3.24
Covariance Parameter Estimates for the Cell Mean Model
Mixed Model With Random Effects Plus an Autoregressive Error

Model with $\nu = 1$, $-2\lambda = 746.0$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{11}$</td>
<td>900</td>
<td>931.7</td>
<td>92.88</td>
<td>10.0</td>
</tr>
<tr>
<td>$\delta_{12}$</td>
<td>1</td>
<td>-2.0</td>
<td>9.52</td>
<td>-0.2</td>
</tr>
<tr>
<td>$\delta_{22}$</td>
<td>10</td>
<td>9.4</td>
<td>1.71</td>
<td>5.5</td>
</tr>
<tr>
<td>$\sigma^2_\epsilon$</td>
<td>500</td>
<td>457.9</td>
<td>17.12</td>
<td>26.7</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.25</td>
<td>0.15</td>
<td>0.029</td>
<td>5.1</td>
</tr>
</tbody>
</table>

Model with $\nu = 2$, $-2\lambda = 745.3$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{11}$</td>
<td>900</td>
<td>941.7</td>
<td>91.84</td>
<td>10.3</td>
</tr>
<tr>
<td>$\delta_{12}$</td>
<td>1</td>
<td>-3.6</td>
<td>9.27</td>
<td>-0.4</td>
</tr>
<tr>
<td>$\delta_{22}$</td>
<td>10</td>
<td>9.7</td>
<td>1.64</td>
<td>5.9</td>
</tr>
<tr>
<td>$\sigma^2_\epsilon$</td>
<td>500</td>
<td>451.9</td>
<td>15.29</td>
<td>29.6</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.25</td>
<td>0.14</td>
<td>0.023</td>
<td>5.9</td>
</tr>
</tbody>
</table>

Model with $\nu = \bar{\nu}$, $-2\lambda = 746.4$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{11}$</td>
<td>880</td>
<td>832.5</td>
<td>230.86</td>
<td>3.6</td>
</tr>
<tr>
<td>$\delta_{12}$</td>
<td>4.8</td>
<td>7.8</td>
<td>17.24</td>
<td>0.5</td>
</tr>
<tr>
<td>$\delta_{22}$</td>
<td>8</td>
<td>7.2</td>
<td>3.52</td>
<td>2.0</td>
</tr>
<tr>
<td>$\sigma^2_\epsilon$</td>
<td>500</td>
<td>524.9</td>
<td>166.19</td>
<td>3.2</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.25</td>
<td>0.25</td>
<td>0.232</td>
<td>1.2</td>
</tr>
<tr>
<td>$\bar{\nu}$</td>
<td>0.10</td>
<td>0.44</td>
<td>0.369</td>
<td>1.2</td>
</tr>
</tbody>
</table>
Table 3.25
Covariance Parameter Estimates for the Cubic Model
Mixed Model With Random Effects Plus an Autoregressive Error

<table>
<thead>
<tr>
<th>Model with $\nu = 1, -2\lambda = 776.9$</th>
<th>Parameter</th>
<th>Starting Estimate</th>
<th>Parameter Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{11}$</td>
<td></td>
<td>900</td>
<td>932.1</td>
<td>92.89</td>
<td>10.0</td>
</tr>
<tr>
<td>$\delta_{12}$</td>
<td></td>
<td>1</td>
<td>-1.9</td>
<td>9.50</td>
<td>-0.2</td>
</tr>
<tr>
<td>$\delta_{22}$</td>
<td></td>
<td>10</td>
<td>9.3</td>
<td>1.70</td>
<td>5.5</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td></td>
<td>500</td>
<td>461.8</td>
<td>17.13</td>
<td>27.0</td>
</tr>
<tr>
<td>$\rho$</td>
<td></td>
<td>0.25</td>
<td>0.14</td>
<td>0.029</td>
<td>4.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model with $\nu = 2, -2\lambda = 776.0$</th>
<th>Parameter</th>
<th>Starting Estimate</th>
<th>Parameter Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{11}$</td>
<td></td>
<td>900</td>
<td>941.7</td>
<td>91.84</td>
<td>10.3</td>
</tr>
<tr>
<td>$\delta_{12}$</td>
<td></td>
<td>1</td>
<td>-3.3</td>
<td>9.27</td>
<td>-0.4</td>
</tr>
<tr>
<td>$\delta_{22}$</td>
<td></td>
<td>10</td>
<td>9.6</td>
<td>1.64</td>
<td>5.8</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td></td>
<td>500</td>
<td>456.7</td>
<td>15.42</td>
<td>29.6</td>
</tr>
<tr>
<td>$\rho$</td>
<td></td>
<td>0.25</td>
<td>0.13</td>
<td>0.023</td>
<td>5.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model with $\nu = \tilde{\nu}, -2\lambda = 793.5$</th>
<th>Parameter</th>
<th>Starting Estimate</th>
<th>Parameter Estimates</th>
<th>Standard $^\dagger$</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{11}$</td>
<td></td>
<td>880</td>
<td>864.9</td>
<td>192.20</td>
<td>4.4</td>
</tr>
<tr>
<td>$\delta_{12}$</td>
<td></td>
<td>4.8</td>
<td>5.5</td>
<td>16.38</td>
<td>0.4</td>
</tr>
<tr>
<td>$\delta_{22}$</td>
<td></td>
<td>8</td>
<td>7.6</td>
<td>3.33</td>
<td>2.2</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td></td>
<td>500</td>
<td>504.6</td>
<td>126.38</td>
<td>4.1</td>
</tr>
<tr>
<td>$\rho$</td>
<td></td>
<td>0.25</td>
<td>0.21</td>
<td>0.186</td>
<td>1.2</td>
</tr>
<tr>
<td>$\tilde{\nu}$</td>
<td></td>
<td>0.10</td>
<td>0.50</td>
<td>0.378</td>
<td>1.2</td>
</tr>
</tbody>
</table>

$^\dagger$ Restricted ML using quasi-scoring was used to obtain these parameter estimates, but this algorithm did not provide standard errors. Approximate standard errors were obtained by using these estimates as starting values in a Newton-Raphson algorithm and setting the number of iterations equal to one.
Table 3.26
Covariance Parameter Estimates for Models with Compound Symmetry
Plus an Autoregressive Error – Cell Mean Model

Model with $\nu = 1$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Starting Estimate</th>
<th>Parameter Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^2$</td>
<td>400</td>
<td>355.9</td>
<td>18.85</td>
<td>18.9</td>
</tr>
<tr>
<td>$\omega^2$</td>
<td>900</td>
<td>-479.8</td>
<td>4262.78</td>
<td>-0.1</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>300</td>
<td>1757.9</td>
<td>4253.84</td>
<td>0.4</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.1</td>
<td>0.03</td>
<td>0.076</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Model with $\nu = 2$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Starting Estimate</th>
<th>Parameter Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^2$</td>
<td>400</td>
<td>402.4</td>
<td>13.18</td>
<td>30.5</td>
</tr>
<tr>
<td>$\omega^2$</td>
<td>900</td>
<td>895.2</td>
<td>108.87</td>
<td>8.2</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>300</td>
<td>337.1</td>
<td>73.17</td>
<td>3.6</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.1</td>
<td>0.03</td>
<td>0.011</td>
<td>3.1</td>
</tr>
</tbody>
</table>

Model with $\nu = 3$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Starting Estimate</th>
<th>Parameter Estimates</th>
<th>Standard Error</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^2$</td>
<td>400</td>
<td>424.7</td>
<td>12.73</td>
<td>33.9</td>
</tr>
<tr>
<td>$\omega^2$</td>
<td>100</td>
<td>956.6</td>
<td>90.57</td>
<td>10.6</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>300</td>
<td>251.3</td>
<td>38.88</td>
<td>6.5</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.1</td>
<td>0.01</td>
<td>0.002</td>
<td>4.2</td>
</tr>
</tbody>
</table>
Table 3.27
Likelihood Ratio Tests
Cell Mean Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Covariance Structure</th>
<th>$-2\lambda$</th>
<th>Comparison</th>
<th>Model</th>
<th>$\chi^2$</th>
<th>d.f.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unstructured</td>
<td>670.1</td>
<td>18 + 45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Mixed Model</td>
<td>772.4</td>
<td>18 + 4</td>
<td>1</td>
<td>102.3</td>
<td>41</td>
</tr>
<tr>
<td>3</td>
<td>Compound Symmetry</td>
<td>938.5</td>
<td>18 + 2</td>
<td>2</td>
<td>166.1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>AR-1, $\nu = 1$</td>
<td>1148.8</td>
<td>18 + 2</td>
<td>1</td>
<td>478.7</td>
<td>43</td>
</tr>
<tr>
<td>5</td>
<td>Mix./Mod. + AR, $\nu = 1$</td>
<td>746.0</td>
<td>18 + 5</td>
<td>2</td>
<td>26.4</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Mix./Mod. + AR, $\nu = 2$</td>
<td>745.3</td>
<td>18 + 5</td>
<td>2</td>
<td>27.1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>C. S. + AR, $\nu = 1$</td>
<td>NC</td>
<td>18 + 4</td>
<td>3</td>
<td>NC</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>C. S. + AR, $\nu = 2$</td>
<td>778.3</td>
<td>18 + 4</td>
<td>3</td>
<td>160.2</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>C. S. + AR, $\nu = 3$</td>
<td>786.9</td>
<td>18 + 4</td>
<td>3</td>
<td>151.6</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>AR-1, $\bar{\nu} = 0.29$</td>
<td>779.1</td>
<td>18 + 3</td>
<td>4</td>
<td>402.4</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>AR-1, $\nu = 2$</td>
<td>2029.9</td>
<td>18 + 2</td>
<td>10</td>
<td>1250.8</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: All models had same model for fixed effects

28,000 has been subtracted from $-2\lambda$

$p$ is the number of parameters used in the expected value model

$q$ is the number of parameters used to describe the covariance matrix

NC means the estimation algorithm did not converge

All LR tests were significant at the $\alpha = 0.01$ level
### Table 3.28
Likelihood Ratio Tests
Cubic Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Covariance Structure</th>
<th>$-2\lambda$</th>
<th>Comparison</th>
<th>$\chi^2$</th>
<th>d.f.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unstructured</td>
<td>702.7</td>
<td>8 + 45</td>
<td>1</td>
<td>98.4</td>
</tr>
<tr>
<td>2</td>
<td>Mixed Model</td>
<td>801.1</td>
<td>8 + 4</td>
<td>1</td>
<td>160.9</td>
</tr>
<tr>
<td>3</td>
<td>Compound Symmetry</td>
<td>962.0</td>
<td>8 + 2</td>
<td>2</td>
<td>483.6</td>
</tr>
<tr>
<td>4</td>
<td>AR-1, $\nu = 1$</td>
<td>1186.3</td>
<td>8 + 2</td>
<td>1</td>
<td>25.1</td>
</tr>
<tr>
<td>5</td>
<td>Mix./Mod. + AR, $\nu = 1$</td>
<td>776.9</td>
<td>8 + 5</td>
<td>2</td>
<td>24.2</td>
</tr>
<tr>
<td>6</td>
<td>Mix./Mod. + AR, $\nu = 2$</td>
<td>776.0</td>
<td>8 + 5</td>
<td>2</td>
<td>25.1</td>
</tr>
<tr>
<td>7</td>
<td>C.S. + $\nu = 1$</td>
<td>804.9</td>
<td>8 + 4</td>
<td>3</td>
<td>157.1</td>
</tr>
<tr>
<td>8</td>
<td>C.S. + $\nu = 2$</td>
<td>808.3</td>
<td>8 + 4</td>
<td>3</td>
<td>153.7</td>
</tr>
<tr>
<td>9</td>
<td>C.S. + $\nu = 3$</td>
<td>815.6</td>
<td>8 + 4</td>
<td>3</td>
<td>146.4</td>
</tr>
<tr>
<td>10</td>
<td>AR-1, $\bar{\nu} = 0.28$</td>
<td>810.0</td>
<td>8 + 3</td>
<td>4</td>
<td>376.3</td>
</tr>
<tr>
<td>11</td>
<td>AR-1, $\nu = 2$</td>
<td>2051.9</td>
<td>8 + 2</td>
<td>10</td>
<td>1241.9</td>
</tr>
</tbody>
</table>

Note: All models had same model for fixed effects

$-2\lambda$ has been subtracted from $-2\lambda$

$p$ is the number of parameters used in the expected value model

$q$ is the number of parameters used to describe the covariance matrix

NC means the estimation algorithm did not converge

All LR tests were significant at the $\alpha = 0.01$ level
<table>
<thead>
<tr>
<th>Visit</th>
<th>Elapsed Time in Months</th>
<th>Diet Status</th>
<th>Random. Status</th>
<th>Post-randomization Visit Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>pre-diet</td>
<td>pre</td>
<td></td>
</tr>
<tr>
<td>4 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>pre-diet</td>
<td>pre</td>
<td></td>
</tr>
<tr>
<td>4 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>on-diet</td>
<td>pre</td>
<td></td>
</tr>
<tr>
<td>4 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>on-diet</td>
<td>pre</td>
<td></td>
</tr>
<tr>
<td>4 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>on-diet</td>
<td>pre</td>
<td></td>
</tr>
<tr>
<td>2 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4.5</td>
<td>on-diet</td>
<td>post</td>
<td>2-month</td>
</tr>
<tr>
<td>2 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>on-diet</td>
<td>post</td>
<td>First 6-month</td>
</tr>
<tr>
<td>2 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>on-diet</td>
<td>post</td>
<td>2-month</td>
</tr>
<tr>
<td>2 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>on-diet</td>
<td>post</td>
<td>2-month</td>
</tr>
<tr>
<td>2 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>13</td>
<td>on-diet</td>
<td>post</td>
<td>First Annual</td>
</tr>
</tbody>
</table>
Table 4.2

Model with 2 Averaged Baseline Values Treated as Additional Time Points
Unstructured Covariance Matrix with a Cell Mean Model

Pre-randomization\textsuperscript{†}
Visit Averages

\[
\begin{bmatrix}
X_1 & X_2 & 7 & 13 & 19 & 25 & 31 & 37 & 43 & 49 & 55 \\
949 & 754 & 691 & 697 & 572 & 577 & 665 & 739 & 690 & 766 & 716 \\
905 & 804 & 745 & 631 & 655 & 748 & 754 & 701 & 777 & 750 & 750 \\
1339 & 1020 & 868 & 839 & 941 & 944 & 851 & 952 & 940 & 940 \\
1493 & 1065 & 1021 & 1160 & 1096 & 1075 & 1048 & 1087 & 1087 \\
1374 & 1106 & 1156 & 1124 & 1114 & 1074 & 1094 & 1094 \\
1518 & 1172 & 1125 & 1199 & 1155 & 1201 & 1199 & 1155 & 1201 \\
1792 & 1134 & 1270 & 1339 & 1427 & 1651 & 1359 & 1401 & 1488 \\
1651 & 1359 & 1401 & 1488 & (sym) & 1856 & 1593 & 2100 & \end{bmatrix}
\]

\textsuperscript{†} Pre-randomization averages:
Pre-diet mean: $\bar{X}_1$ = mean of (visits 1 and 2)
On-diet mean: $\bar{X}_2$ = mean of (visits 3, 4, and 5)
Table 4.3

Model with 5 Baseline Values Treated as Additional Time Points
Unstructured Covariance Matrix with a Cell Mean Model

<table>
<thead>
<tr>
<th>Pre-randomization† Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  2  3  4  5  7  13 19 25 31 37 43 49 55</td>
</tr>
</tbody>
</table>

| 1026 | 804 | 698 | 711 | 773 | 663 | 655 | 558 | 577 | 627 | 698 | 651 | 739 | 679 |
| 1161 | 761 | 788 | 790 | 720 | 740 | 583 | 575 | 702 | 780 | 729 | 792 | 763 |
| 987  | 790 | 805 | 741 | 695 | 639 | 604 | 723 | 746 | 687 | 753 | 705 |
| 1049 | 848 | 813 | 779 | 621 | 673 | 756 | 739 | 723 | 781 | 790 |
| 1218 | 856 | 760 | 631 | 688 | 763 | 776 | 691 | 796 | 770 |
| 1400 | 1021| 868 | 840 | 941 | 943 | 851 | 953 | 944 |
| 1493 | 1063| 1022| 1160| 1095| 1074| 1049| 1082|
| 1371 | 1106| 1154| 1122| 1111| 1072| 1088|
| 1520 | 1172| 1124| 1200| 1157| 1192|
| 1791 | 1332| 1269| 1339| 1415|
| 1648 | 1356| 1346| 1396|
| (sym) |
| 1773 | 1420| 1479|
| 1855 | 1589|
| 2088 |

† Pre-randomization visits: 1 and 2 are pre-diet
3, 4, and 5 are on-diet
Table 4.4
Likelihood Ratio Tests
Model with 5 Baseline Values Treated as Additional Time Points
Cell Mean Model Plus Cubic Contrasts

<table>
<thead>
<tr>
<th>Model</th>
<th>Covariance Structure</th>
<th>Adjusted $-2\lambda$</th>
<th>$p + q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unstructured</td>
<td>0</td>
<td>18 + 105</td>
</tr>
<tr>
<td>2</td>
<td>Mixed Model</td>
<td>576.0</td>
<td>18 + 4</td>
</tr>
<tr>
<td>3</td>
<td>Compound Symmetry</td>
<td>1156.7</td>
<td>18 + 2</td>
</tr>
<tr>
<td>4</td>
<td>C.S. + AR, $c = 0.50$</td>
<td>398.0</td>
<td>18 + 4</td>
</tr>
<tr>
<td>5</td>
<td>C.S. + AR, $c = 0.80$</td>
<td>407.1</td>
<td>18 + 4</td>
</tr>
<tr>
<td>7</td>
<td>C.S. + AR, $c = 1.00$</td>
<td>417.4</td>
<td>18 + 4</td>
</tr>
<tr>
<td>8</td>
<td>AR-1, $\tilde{\nu} = 0.222$</td>
<td>406.5</td>
<td>18 + 3</td>
</tr>
<tr>
<td>9</td>
<td>AR-1, $\nu = 1.0$</td>
<td>2773.6</td>
<td>18 + 2</td>
</tr>
</tbody>
</table>

† For models of this type with other values of $c$, see Table 4.5.
‡ For models of this type with other values of $\nu$, see Table 4.6.

Note: All models had same model for fixed effects.
The $-2\lambda$ value of the unstructured model (46,126.16) has been subtracted from each model $-2\lambda$ for easier comparisons.
$p$ is the number of parameters used in the expected value model.
$q$ is the number of parameters used to describe the covariance matrix.
Table 4.5

Estimates for Mixed Model With Compound Symmetry Plus AR(1) Error
Model with 5 Baseline Values Treated as Additional Time Points
Cell Mean Model Plus Cubic Contrasts

<table>
<thead>
<tr>
<th></th>
<th>0.50†</th>
<th>0.60</th>
<th>0.70</th>
<th>0.80</th>
<th>0.90</th>
<th>1.0</th>
<th>1.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\tau}^2 )</td>
<td>243.21</td>
<td>261.71</td>
<td>276.61</td>
<td>288.96</td>
<td>299.44</td>
<td>308.53</td>
<td>316.47</td>
</tr>
<tr>
<td>( \hat{\sigma}^2 )</td>
<td>10.94</td>
<td>490.00</td>
<td>645.80</td>
<td>719.57</td>
<td>760.82</td>
<td>786.28</td>
<td>803.46</td>
</tr>
<tr>
<td>( \hat{\sigma}^2 )</td>
<td>1264.86</td>
<td>767.52</td>
<td>596.87</td>
<td>510.59</td>
<td>458.39</td>
<td>423.05</td>
<td>396.83</td>
</tr>
<tr>
<td>( \hat{\alpha} )</td>
<td>0.176</td>
<td>0.274</td>
<td>0.323</td>
<td>0.357</td>
<td>0.366</td>
<td>0.363</td>
<td>0.354</td>
</tr>
<tr>
<td>(-2\lambda)^†</td>
<td>397.98</td>
<td>398.96</td>
<td>403.16</td>
<td>407.09</td>
<td>411.42</td>
<td>417.41</td>
<td>423.54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1.20</th>
<th>1.25</th>
<th>1.30</th>
<th>1.35</th>
<th>1.50</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\tau}^2 )</td>
<td>323.39</td>
<td>326.42</td>
<td>329.10</td>
<td>331.32</td>
<td>333.47</td>
<td>334.88</td>
</tr>
<tr>
<td>( \hat{\omega}^2 )</td>
<td>816.14</td>
<td>821.78</td>
<td>827.43</td>
<td>833.57</td>
<td>859.16</td>
<td>901.90</td>
</tr>
<tr>
<td>( \hat{\sigma}^2 )</td>
<td>375.94</td>
<td>366.59</td>
<td>357.62</td>
<td>348.73</td>
<td>321.22</td>
<td>280.50</td>
</tr>
<tr>
<td>( \hat{\alpha} )</td>
<td>0.343</td>
<td>0.338</td>
<td>0.334</td>
<td>0.333</td>
<td>0.369</td>
<td>0.500</td>
</tr>
<tr>
<td>(-2\lambda)</td>
<td>430.09</td>
<td>433.47</td>
<td>436.90</td>
<td>440.34</td>
<td>450.48</td>
<td>474.35</td>
</tr>
</tbody>
</table>

† For easier comparisons, 46,126.16 has been subtracted from each \(-2\lambda\) value.
‡ Models in which \(c \leq 0.49\) produced negative variance estimates. See below.

Note: the parameter \(c\) is a constant in the estimation algorithm for all models.

<table>
<thead>
<tr>
<th></th>
<th>0.45</th>
<th>0.49</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\tau}^2 )</td>
<td>232.7</td>
<td>241.1</td>
</tr>
<tr>
<td>( \hat{\omega}^2 )</td>
<td>-935.2</td>
<td>-95.8</td>
</tr>
<tr>
<td>( \hat{\sigma}^2 )</td>
<td>2221.8</td>
<td>1373.6</td>
</tr>
<tr>
<td>( \hat{\alpha} )</td>
<td>0.102</td>
<td>0.163</td>
</tr>
<tr>
<td>(-2\lambda)</td>
<td>397.20</td>
<td>397.80</td>
</tr>
</tbody>
</table>
Table 4.6

Estimates for Models with Stationary Variance and Exponentially Decreasing Correlation
Model with 5 Baseline Values Treated as Additional Time Points
Cell Mean Model Plus Cubic Contrasts

<table>
<thead>
<tr>
<th>( c = )</th>
<th>0.16</th>
<th>0.18</th>
<th>0.20</th>
<th>.22</th>
<th>( \tilde{\nu} = 0.222^\dagger )</th>
<th>0.24</th>
<th>0.26</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma^2 )</td>
<td>1555.0</td>
<td>1546.4</td>
<td>1537.0</td>
<td>1527.3</td>
<td>1526.3</td>
<td>1517.6</td>
<td>1508.2</td>
</tr>
<tr>
<td>( \hat{\rho} )</td>
<td>0.703</td>
<td>0.700</td>
<td>0.697</td>
<td>0.693</td>
<td>0.693</td>
<td>0.689</td>
<td>0.683</td>
</tr>
<tr>
<td>( -2\lambda^\dagger )</td>
<td>440.00</td>
<td>421.25</td>
<td>410.39</td>
<td>406.52</td>
<td>406.49</td>
<td>408.93</td>
<td>417.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( c = )</th>
<th>0.28</th>
<th>0.30</th>
<th>0.35</th>
<th>0.40</th>
<th>0.50</th>
<th>0.75</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma^2 )</td>
<td>1499.3</td>
<td>1491.2</td>
<td>1475.6</td>
<td>1468.0</td>
<td>1481.1</td>
<td>1647.5</td>
<td>1804.7</td>
</tr>
<tr>
<td>( \hat{\rho} )</td>
<td>0.678</td>
<td>0.671</td>
<td>0.654</td>
<td>0.634</td>
<td>0.592</td>
<td>0.459</td>
<td>0.223</td>
</tr>
<tr>
<td>( -2\lambda )</td>
<td>430.36</td>
<td>448.53</td>
<td>513.10</td>
<td>602.07</td>
<td>842.96</td>
<td>1728.6</td>
<td>2773.6</td>
</tr>
</tbody>
</table>

\( ^\dagger \) For easier comparisons, 46,126.16 has been subtracted from each \( -2\lambda \) value.
\( ^\ddagger \) The actual MLE for \( \nu \) was estimated in this model, with \( \tilde{\nu} = 0.2220 \). For all other models, \( \nu \) is set to a was a constant, \( c \), in the estimation algorithm.
Table 4.7
Likelihood Ratio Tests
Model with 2 Average Baseline Values Treated as Additional Time Points
Cell Mean Model Plus Cubic Contrasts

<table>
<thead>
<tr>
<th>Model</th>
<th>Covariance Structure</th>
<th>Adjusted $-2\lambda$</th>
<th>$p + q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unstructured</td>
<td>0</td>
<td>12 + 66</td>
</tr>
<tr>
<td>2</td>
<td>Mixed Model</td>
<td>394.5</td>
<td>12 + 4</td>
</tr>
<tr>
<td>3</td>
<td>Compound Symmetry</td>
<td>716.7</td>
<td>12 + 2</td>
</tr>
<tr>
<td>4</td>
<td>C.S. + AR, $c = 0.35$†</td>
<td>325.6</td>
<td>12 + 4</td>
</tr>
<tr>
<td>5</td>
<td>C.S. + AR, $c = 0.50$</td>
<td>329.4</td>
<td>12 + 4</td>
</tr>
<tr>
<td>7</td>
<td>C.S. + AR, $c = 1.00$</td>
<td>347.2</td>
<td>12 + 4</td>
</tr>
<tr>
<td>8</td>
<td>AR-1, $\varphi = 0.2727$‡</td>
<td>330.1</td>
<td>12 + 3</td>
</tr>
<tr>
<td>9</td>
<td>AR-1, $\nu = 1.0$</td>
<td>1275.5</td>
<td>12 + 2</td>
</tr>
</tbody>
</table>

† For models of this type with other values of $c$, see Table 4.8.
‡ For models of this type with other values of $\nu$, see Table 4.9.

Note: All models had same model for fixed effects.
The $-2\lambda$ value of the unstructured model (35512.60) has been subtracted from each model $-2\lambda$ for easier comparisons.
$p$ is the number of parameters used in the expected value model.
$q$ is the number of parameters used to describe the covariance matrix.
Table 4.8

Estimates for Mixed Model With Compound Symmetry Plus AR(1) Error
Model with 2 Averaged Baseline Values Treated as Additional Time Points
Cell Mean Model Plus Cubic Contrasts

\[ c = \begin{array}{cccccccc}
0.30^\dagger & 0.35 & 0.40 & 0.50 & 0.60 & 0.70 & 0.80 \\
\hline
\tilde{\tau}^2 & -57.80 & 9.65 & 60.55 & 132.79 & 181.74 & 217.68 & 245.54 \\
\tilde{\omega}^2 & 587.03 & 651.22 & 697.34 & 756.75 & 791.95 & 812.98 & 825.31 \\
\tilde{\sigma}^2 & 993.10 & 860.98 & 763.42 & 630.50 & 544.84 & 486.11 & 443.96 \\
\hat{\alpha} & 0.661 & 0.650 & 0.636 & 0.600 & 0.557 & 0.509 & 0.457 \\
-2\lambda^\dagger & 324.64 & 325.63 & 326.74 & 329.35 & 332.38 & 335.76 & 339.40 \\
\hline
\end{array} \]

\[ c = \begin{array}{cccc}
0.90 & 1.00 & 1.50 & 2.00 \\
\hline
\tilde{\tau}^2 & 268.07 & 286.96 & 349.09 & 314.35 \\
\tilde{\omega}^2 & 831.83 & 834.01 & 815.74 & 946.79 \\
\tilde{\sigma}^2 & 412.80 & 389.52 & 335.80 & 251.55 \\
\hat{\alpha} & 0.405 & 0.354 & 0.153 & 0.487 \\
-2\lambda & 343.20 & 347.20 & 367.08 & 388.04 \\
\end{array} \]

\hspace{0.5cm} ^\dagger For easier comparisons, 35512.60 has been subtracted from each \(-2\lambda\) value.

\hspace{0.5cm} ^\ddagger Models in which \(c \leq 0.35\) produced negative variance estimates.

Note: the parameter \(c\) is a constant in the estimation algorithm for all models.
Table 4.9

Estimates for Models with Stationary Variance and Exponentially Decreasing Correlation
Model with 2 Average Baseline Values Treated as Additional Time Points Cell Mean Model Plus Cubic Contrasts

<table>
<thead>
<tr>
<th></th>
<th>$c = 0.20$</th>
<th>$c = 0.25$</th>
<th>$c = 0.26$</th>
<th>$c = 0.27$</th>
<th>$\bar{\nu} = 0.2727^\dagger$</th>
<th>$c = 0.28$</th>
<th>$c = 0.30$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>1535.3</td>
<td>1524.1</td>
<td>1521.8</td>
<td>1519.4</td>
<td>1518.8</td>
<td>1517.1</td>
<td>1512.5</td>
</tr>
<tr>
<td>$\hat{\rho}$</td>
<td>0.723</td>
<td>0.725</td>
<td>0.725</td>
<td>0.725</td>
<td>0.725</td>
<td>0.724</td>
<td>0.724</td>
</tr>
<tr>
<td>$-2\lambda^\dagger$</td>
<td>347.13</td>
<td>331.68</td>
<td>330.62</td>
<td>330.17</td>
<td>330.15</td>
<td>330.30</td>
<td>332.23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$c = 0.40$</th>
<th>$c = 0.50$</th>
<th>$c = 0.75$</th>
<th>$c = 1.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>1493.3</td>
<td>1485.1</td>
<td>1527.2</td>
<td>1624.3</td>
</tr>
<tr>
<td>$\hat{\rho}$</td>
<td>0.717</td>
<td>0.706</td>
<td>0.668</td>
<td>0.603</td>
</tr>
<tr>
<td>$-2\lambda$</td>
<td>370.72</td>
<td>448.80</td>
<td>780.44</td>
<td>1275.5</td>
</tr>
</tbody>
</table>

† For easier comparisons, 35512.60 has been subtracted from each $-2\lambda$ value.
‡ The actual MLE for $\nu$ was estimated in this model, with $\bar{\nu} = 0.2727$. For all other models, $\nu$ is set to a was a constant, $c$, in the estimation algorithm.
Table 4.10

Model with 2 Averaged Baseline Values Treated as Additional Time Points
Comparison of $\hat{\sigma}$ for Diggle's Model

<table>
<thead>
<tr>
<th>Pre-randomization†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visit Averages</td>
</tr>
<tr>
<td>$\bar{x}_1$</td>
</tr>
<tr>
<td>1522</td>
</tr>
<tr>
<td>1522</td>
</tr>
<tr>
<td>1522</td>
</tr>
<tr>
<td>1522</td>
</tr>
<tr>
<td>$c = 0.35$</td>
</tr>
<tr>
<td>1522</td>
</tr>
<tr>
<td>1522</td>
</tr>
<tr>
<td>(sym)</td>
</tr>
<tr>
<td>1522</td>
</tr>
</tbody>
</table>

| 1520   | 1253   | 1206  | 1073  | 1009  | 968   | 938  | 915  | 896  | 881  | 868  |
| 1520   | 1253   | 1087  | 1018  | 974   | 942   | 918  | 899  | 883  | 870  |       |
| 1520   | 1103   | 1027  | 980   | 947   | 921   | 902  | 886  | 872  |       |       |
| 1520   | 1103   | 1027  | 980   | 947   | 921   | 902  | 886  |       |       |       |
| $c = 0.50$ |       |       |       |       |       |       |       |       |       |       |
| 1520   | 1103   | 1027  | 980   | 947   | 921   | 902  |       |       |       |       |
| 1520   | 1103   | 1027  | 980   | 947   |       |       |       |       |       |       |
| (sym)  |       |       |       |       |       |       |       |       |       |       |
| 1520   |       |       |       |       |       |       |       |       |       |       |
| 1520   |       |       |       |       |       |       |       |       |       |       |

† Pre-randomization averages:
Pre-diet mean: $\bar{x}_1 =$ mean of (visits 1 and 2)
On-diet mean: $\bar{x}_2 =$ mean of (visits 3, 4, and 5)
Table 4.11

Model with 2 Averaged Baseline Values Treated as Additional Time Points
\( \bar{Z} \) for the General Mixed Model

<table>
<thead>
<tr>
<th>( \bar{X}_1 )</th>
<th>( \bar{X}_2 )</th>
<th>7</th>
<th>13</th>
<th>19</th>
<th>25</th>
<th>31</th>
<th>37</th>
<th>43</th>
<th>49</th>
<th>55</th>
</tr>
</thead>
<tbody>
<tr>
<td>1175</td>
<td>748</td>
<td>751</td>
<td>767</td>
<td>784</td>
<td>801</td>
<td>818</td>
<td>835</td>
<td>852</td>
<td>868</td>
<td>885</td>
</tr>
<tr>
<td>1181</td>
<td>754</td>
<td>773</td>
<td>792</td>
<td>811</td>
<td>830</td>
<td>848</td>
<td>867</td>
<td>886</td>
<td>905</td>
<td></td>
</tr>
<tr>
<td>1187</td>
<td>778</td>
<td>799</td>
<td>819</td>
<td>840</td>
<td>860</td>
<td>881</td>
<td>902</td>
<td>922</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1241</td>
<td>843</td>
<td>876</td>
<td>908</td>
<td>941</td>
<td>973</td>
<td>1006</td>
<td>1038</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1318</td>
<td>932</td>
<td>976</td>
<td>1021</td>
<td>1065</td>
<td>1110</td>
<td>1154</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1418</td>
<td>1044</td>
<td>1101</td>
<td>1157</td>
<td>1213</td>
<td>1270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1543</td>
<td>1181</td>
<td>1249</td>
<td>1317</td>
<td>1386</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1691</td>
<td>1341</td>
<td>1421</td>
<td>1501</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1863</td>
<td>1525</td>
<td>1617</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(sym)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ Z' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.08 & 0.25 & 0.4 & 1.4 & 2.4 & 3.4 & 4.4 & 5.4 & 6.4 & 7.4 & 8.4 \end{bmatrix} \]

\[ \dagger \] Pre-randomization averages:
- Pre-diet mean: \( \bar{X}_1 = \text{mean of (visits 1 and 2)} \)
- On-diet mean: \( \bar{X}_2 = \text{mean of (visit 3, 4, and 5)} \)
Table 4.12

Likelihood Ratio Tests
Model with Cholesterol Measure at Visit 1 Treated as a Covariate,
Baseline Measures from Visits 2–5 are Additional Time Points
Cell Mean Model Plus Cubic Contrasts

<table>
<thead>
<tr>
<th>Model</th>
<th>Covariance Structure</th>
<th>Adjusted $-2\lambda$</th>
<th>$p + q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unstructured</td>
<td>0</td>
<td>17 + 91</td>
</tr>
<tr>
<td>2</td>
<td>Mixed Model</td>
<td>535.4</td>
<td>17 + 4</td>
</tr>
<tr>
<td>3</td>
<td>Compound Symmetry</td>
<td>1094.1</td>
<td>17 + 2</td>
</tr>
<tr>
<td>4</td>
<td>AR-1, $\phi = 0.2302^\dagger$</td>
<td>528.7</td>
<td>17 + 3</td>
</tr>
<tr>
<td>5</td>
<td>AR-1, $\nu = 1.0$</td>
<td>2329.4</td>
<td>17 + 2</td>
</tr>
<tr>
<td>6</td>
<td>C.S. + AR, $c =$ constant$^\ddagger$</td>
<td>NC</td>
<td>17 + 4</td>
</tr>
</tbody>
</table>

$^\dagger$ For models of this type with other values of $\nu$, see Table 4.13.
$^\ddagger$ The estimation algorithm for Diggle's model never converged.

Note: All models had same model for fixed effects.
The $-2\lambda$ value of the unstructured model (42315.1) has been subtracted from each model $-2\lambda$ for easier comparisons.
$p$ is the number of parameters used in the expected value model.
$q$ is the number of parameters used to describe the covariance matrix.
Table 4.13

Estimates for Models with Stationary Variance and Exponentially Decreasing Correlation
Model with Cholesterol Measure at Visit 1 Treated as a Covariate, Baseline Measures from Visits 2–5 are Additional Time Points Cell Mean Model Plus Cubic Contrasts

<table>
<thead>
<tr>
<th>c</th>
<th>0.15</th>
<th>0.20</th>
<th>0.21</th>
<th>0.22</th>
<th>ϑ = 0.2302†</th>
<th>0.25</th>
<th>0.30</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ²</td>
<td>1113.0</td>
<td>1105.0</td>
<td>1103.2</td>
<td>1101.3</td>
<td>1099.5</td>
<td>1096.2</td>
<td>1090.1</td>
</tr>
<tr>
<td>ρ</td>
<td>0.582</td>
<td>0.579</td>
<td>0.578</td>
<td>0.576</td>
<td>0.574</td>
<td>0.570</td>
<td>0.556</td>
</tr>
<tr>
<td>−2λ†</td>
<td>547.22</td>
<td>534.50</td>
<td>531.28</td>
<td>529.38</td>
<td>528.74</td>
<td>531.02</td>
<td>555.30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>c</th>
<th>0.40</th>
<th>0.50</th>
<th>0.75</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ²</td>
<td>1093.0</td>
<td>1119.3</td>
<td>1251.3</td>
<td>1291.8</td>
</tr>
<tr>
<td>ρ</td>
<td>0.523</td>
<td>0.484</td>
<td>0.348</td>
<td>0.089</td>
</tr>
<tr>
<td>−2λ</td>
<td>670.86</td>
<td>860.05</td>
<td>1568.4</td>
<td>2329.4</td>
</tr>
</tbody>
</table>

† For easier comparisons, 42315.1 has been subtracted from each −2λ value.
‡ The actual MLE for ν was estimated in this model, with ϑ = 0.2302. For all other models, ν is set to a was a constant, c, in the estimation algorithm.
Table 4.14
Covariance Matrices
Model with Cholesterol Measure at Visit 1 Treated as a Covariate,
Baseline Measures from Visits 2–5 are Additional Time Points
Cell Mean Model Plus Cubic Contrasts

<table>
<thead>
<tr>
<th>Pre-randomization Visits 2–4</th>
<th>Unstructured</th>
</tr>
</thead>
<tbody>
<tr>
<td>536 212 229 187</td>
<td>195 222 132 111 204 231 213 213 224</td>
</tr>
<tr>
<td>512 306 278 292</td>
<td>253 245 300 273 247 250 243</td>
</tr>
<tr>
<td>556 312 354 328</td>
<td>236 276 324 256 272 269 324</td>
</tr>
<tr>
<td>638 354 266 203</td>
<td>248 288 250 197 239 255</td>
</tr>
<tr>
<td>976 602 523 479</td>
<td>543 494 438 474 513</td>
</tr>
<tr>
<td>1092 715 664 771</td>
<td>655 661 573 656</td>
</tr>
<tr>
<td>1112 820 832 749</td>
<td>779 668 731</td>
</tr>
<tr>
<td>1225 836 731 847</td>
<td>739 815</td>
</tr>
<tr>
<td>1419 908 879 882</td>
<td>1003</td>
</tr>
<tr>
<td>1174 915 837 933</td>
<td>1053</td>
</tr>
<tr>
<td>1370 948 1084</td>
<td>1633</td>
</tr>
</tbody>
</table>

General Mixed Model

| 730 290 291 292 | 293 305 316 328 339 351 362 374 386 |
| 732 292 293 294 | 307 320 333 345 358 371 384 396 |
| 734 295 296 310 | 324 337 351 365 379 393 407 |
| 736 297 312 322 | 342 357 373 388 403 418 |
| 739 315 331 347 | 363 380 396 412 429 |
| 786 377 408 439 | 470 501 532 563 |
| 863 468 514 560 | 605 651 697 |
| 969 589 650 710 | 771 831 |
| 1105 740 615 815 | 890 965 |
| 1270 920 1010 1100 | 1234 |
| 1465 1129 1234 | 1689 1368 |
| 1943 | |

Murray’s Model, $\tilde{\nu} = 0.2302$

| 1099 806 764 737 | 717 608 560 529 505 486 470 457 445 |
| 1099 806 764 | 737 613 563 531 507 488 472 458 446 |
| 1099 806 | 764 619 567 533 509 489 473 459 447 |
| 1099 | 806 625 570 535 510 490 474 460 448 |
| 1099 | 631 573 538 512 492 475 461 449 |
| 1099 | 631 |
| 1099 | |
Table 4.15
Likelihood Ratio Tests
Model with Mean Baseline Measure from Visits 1 & 2 Treated as a Covariate,
Mean Baseline Measure from Visits 3-5 is an Additional Time Point
Cell Mean Model Plus Cubic Contrasts

<table>
<thead>
<tr>
<th>Model</th>
<th>Covariance Structure</th>
<th>Adjusted $-2\lambda$</th>
<th>$p + q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unstructured</td>
<td>0</td>
<td>11 + 55</td>
</tr>
<tr>
<td>2</td>
<td>Mixed Model</td>
<td>344.7</td>
<td>11 + 4</td>
</tr>
<tr>
<td>3</td>
<td>Compound Symmetry</td>
<td>612.9</td>
<td>11 + 2</td>
</tr>
<tr>
<td>4</td>
<td>AR-1, $\nu = 0.2562^\dagger$</td>
<td>388.0</td>
<td>11 + 3</td>
</tr>
<tr>
<td>5</td>
<td>AR-1, $\nu = 1.0$</td>
<td>1150.5</td>
<td>11 + 2</td>
</tr>
<tr>
<td>6</td>
<td>C.S. + AR, $c = \text{constant}^\ddagger$</td>
<td>NC</td>
<td>11 + 4</td>
</tr>
</tbody>
</table>

$^\dagger$ For models of this type with other values of $\nu$, see Table 4.16.
$^\ddagger$ The estimation algorithm for Diggle's model never converged.

Note: All models had same model for fixed effects.
The $-2\lambda$ value of the unstructured model (31732.19) has been
subtracted from each model $-2\lambda$ for easier comparisons.

$p$ is the number of parameters used in the expected value model.
$q$ is the number of parameters used to describe the covariance matrix.
Table 4.16
Likelihood Ratio Tests
Model with Mean Baseline Measure from Visits 1 & 2 Treated as a Covariate,
Mean Baseline Measure from Visits 3-5 is an Additional Time Point
Cell Mean Model Plus Cubic Contrasts

<table>
<thead>
<tr>
<th>c</th>
<th>0.20</th>
<th>0.24</th>
<th>0.25</th>
<th>$\nu = 0.2562^{\dagger}$</th>
<th>0.26</th>
<th>0.30</th>
<th>0.40</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>1060.0</td>
<td>1054.2</td>
<td>1052.7</td>
<td>1051.8</td>
<td>1051.3</td>
<td>1046.0</td>
<td>1037.4</td>
</tr>
<tr>
<td>$\hat{\rho}$</td>
<td>0.5961</td>
<td>0.5990</td>
<td>0.5994</td>
<td>0.5996</td>
<td>0.5997</td>
<td>0.6000</td>
<td>0.5959</td>
</tr>
<tr>
<td>$-2\lambda^{\dagger}$</td>
<td>395.74</td>
<td>388.24</td>
<td>388.10</td>
<td>388.01</td>
<td>388.04</td>
<td>392.16</td>
<td>428.84</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>c</th>
<th>0.50</th>
<th>0.75</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>1037.6</td>
<td>1074.5</td>
<td>1130.2</td>
</tr>
<tr>
<td>$\hat{\rho}$</td>
<td>0.588</td>
<td>0.552</td>
<td>0.483</td>
</tr>
<tr>
<td>$-2\lambda$</td>
<td>496.66</td>
<td>771.81</td>
<td>1150.51</td>
</tr>
</tbody>
</table>

$^{\dagger}$ For easier comparisons, 31732.19 has been subtracted from each $-2\lambda$ value.
$^{\dagger}$ The actual MLE for $\nu$ was estimated in this model, with $\nu = 0.2302$. For all other models, $\nu$ is set to a was a constant, c, in the estimation algorithm.
Table 4.17  
Covariance Matrices  
Model with Mean Baseline Measure from Visits 1 & 2 Treated as a Covariate,  
Mean Baseline Measure from Visits 3-5 is an Additional Time Point  
Cell Mean Model Plus Cubic Contrasts  

<table>
<thead>
<tr>
<th>Pre-randomization,†</th>
<th>On-diet Average</th>
<th>Unstructured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visit</td>
<td>X</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>302</td>
</tr>
<tr>
<td></td>
<td></td>
<td>900</td>
</tr>
<tr>
<td></td>
<td></td>
<td>993</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1079</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1205</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1334</td>
</tr>
<tr>
<td></td>
<td>(sym)</td>
<td></td>
</tr>
</tbody>
</table>

† Pre-randomization, on-diet mean: $\bar{X}$ = mean of (visits 3, 4, and 5)

Mixed Model

|                      |                  | 786| 359 | 368 | 378 | 387 | 397 | 407 | 416 | 426 | 435 |
|                      |                  | 790| 372 | 383 | 395 | 406 | 417 | 429 | 440 | 452 |        |
|                      |                  | 824| 417 | 440 | 463 | 485 | 508 | 531 | 553 |    |        |
|                      |                  | 880| 485 | 519 | 553 | 587 | 621 | 655 |    |    |        |
|                      |                  | 960| 576 | 621 | 666 | 711 | 757 |    |    |    |        |
|                      |                  | 1061|689  |745  |802  |858 |    |    |    |    |        |
|                      |                  | 1186|825  |892  |960 |    |    |    |    |    |        |
|                      | (sym)            |    |     |     |     |     |     |     |     |     |     |

Murray’s Model, $\hat{\nu} = 0.2562$

|                      |                  | 1052|764  |618  |564 |529 |503 |483 |466 |451 |438 |
|                      |                  | 1052|631  |571  |534 |507 |486 |468 |453 |440 |    |
|                      |                  | 1052|631  |571  |534 |507 |486 |468 |453 |453 |    |
|                      |                  |    |     |     |     |     |     |     |     |     |     |
| (sym)                |                  |    |     |     |     |     |     |     |     |     |     |

|                      |                  |    |     |     |     |     |     |     |     |     |     |

171
Figure 3.1
Simple Group Means

Total Cholesterol

Year in Study

A
B

LRC Study
Minnesota, N=393
Figure 3.2
Correlation Plot
Simple Correlation Plot

Correlations

Lag: Difference in Years

LRC Study
Minnesota Clinic, N=393
Figure 3.3
Covariance Plot
Simple Covariance Matrix

LRC Study
Minnesota Clinic, N=393
Figure 3.4
Model-Based Estimates of Group Means
Cell Mean Model, Unstructured Covariance

Total Cholesterol

Year in Study

LRC Study
Minnesota, N=393
Figure 3.5
Model-Based Estimates of Group Means
Cubic Model, Unstructured Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.6
Model-Based Estimates of Group Means
Quadratic Model, Unstructured Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.7
Model-Based Estimates of Group Means
Linear Model, Unstructured Covariance

Total Cholesterol

Year in Study

LRC Study
Minnesota Clinic, N=393
Figure 3.8
Correlation Plot
Cubic Model with Unstructured Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.9
Covariance Plot
Cubic Model with Unstructured Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.10
Correlation Plot
Cell Mean Model, Unstructured Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.11
Covariance Plot
Cell Mean Model, Unstructured Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.12
Covariance Plot
Ordinary Least Squares Method

LRC Study
Minnesota, N=393
Figure 3.13
Covariance Plot
Cell Mean Model, Mixed Model Covariance

LRC Study
Minneapolis Clinic, N=393
Figure 3.14
Correlation Plot
Cell Mean Model, Mixed Model Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.16
Correlation Plot
Cell Mean Model with AR(1) Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.17
Covariance Plot
Cubic Model with Mixed Model Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.18
Correlation Plot
Cubic Model with Mixed Model Covariance

Correlations

Lag: Difference in Years

LRC Study
Minnesota Clinic, N=393
LRC Study
Minnesota Clinic, N=393
Figure 3.20
Correlation Plot
Cubic Model with AR(1) Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.21
Model-Based Estimates of Group Means
Cubic Model, Mixed Model Covariance

Total Cholesterol

Year in Study

LRC Study
Minnesota Clinic, N=393
Figure 3.22
Model-Based Estimates of Group Means
Cubic Model, Mixed Model + AR Error Cov.

Total Cholesterol

Year in Study

LRC Study
Minnesota Clinic, N=393
Figure 3.23
Covariance Plot: Cell Mean Model with Compound Symmetry + AR Error Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.24
Correlation Plot: Cell Mean Model with Compound Symmetry + AR Error Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.25
Covariance Plot: Cubic Model with Compound Symmetry + AR Error Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.26
Correlation Plot: Cubic Model with Compound Symmetry + AR Error Covariance

Correlations

Lag: Difference in Years

0 0.2 0.4 0.6 0.8 1
1 2 3 4 5 6 7 8

LRC Study
Minnesota Clinic, N=393
Figure 3.27
Covariance Plot: Cell Mean Model with Mixed Model + AR Error Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.28
Correlation Plot: Cell Mean Model with Mixed Model Plus AR Error Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.29
Covariance Plot: Cubic Model with Mixed Model + AR Error Covariance

LRC Study
Minnesota Clinic, N=393
Figure 3.30
Correlation Plot: Cubic Model with Mixed Model Plus AR Error Covariance

Correlations

Lag: Difference in Years

LRC Study
Minnesota Clinic, N=393
Figure 4.1
Model-Based Estimates of Group Means
5 Baseline Values are Added Time Points

Total Cholesterol

Year in Study

Expected Value Model: Cell Mean
Covariance Model: Unstructured
Figure 4.2
Model-Based Estimates of Group Means
2 Baseline Means are Added Time Points

Expected Value Model: Cell Mean
Covariance Model: Unstructured
Figure 4.3
Covariance Plot: Cell Mean Model
5 Baseline Values are Added Time Points

Expected Value Model: Cell Mean
Covariance Model: Unstructured
Figure 4.4
Covariance Plot: Cell Mean Model
2 Baseline Means are Added Time Points

Expected Value Model: Cell Mean
Covariance Model: Unstructured
Figure 4.5
Covariance Plot: Unstructured Covariance
5 Baseline Values are Added Time Points

Cell Mean Model and Cubic Contrasts
Figure 4.6
Covariance Plot: General Mixed Model
5 Baseline Values are Added Time Points

Cell Mean Model and Cubic Contrasts

Lag Time in Years

Covariances

0 1 2 3 4 5 6 7 8 9 10

500 1000 1500
Figure 4.7
Covariance Plot: Diggle's Model
5 Baseline Values are Added Time Points

Covariances

Lag Time in Years

Cell Mean Model and Cubic Contrasts
Figure 4.8

Covariance Plot: Murray's Model

5 Baseline Values are Added Time Points

Covariances

Lag Time in Years

Cell Mean Model and Cubic Contrasts
Figure 4.9

Model-Based Means, Unstructured Cov.

5 Baseline Values are Added Time Points

Expected Value: Cell Mean & Cubic Trends
Covariance Model: Unstructured
Figure 4.10
Model-Based Means, Murray’s Model Cov.
5 Baseline Values are Added Time Points

Expected Value: Cell Mean & Cubic Trends
Covariance Model: Murray’s Model w MLE
Figure 4.11

Covariance Plot: 2 Mean Baseline Values Are Added Time Points, Unstructured Cov.

Cell Mean Model and Cubic Contrasts
Unstructured Covariance
Figure 4.12
Covariance Plot: 2 Mean Baseline Values Are Added Time Points, Mixed Model Cov.

Covariances

Cell Mean Model and Cubic Contrasts
General Mixed Model Covariance
Figure 4.13
Covariance Plot: 2 Mean Baseline Values Are Added Time Points, Diggle’s Model

Cell Mean Model and Cubic Contrasts
Diggle’s Model Covariance
Figure 4.14
Covariance Plot: 2 Mean Baseline Values
Are Added Time Points, Murray's Model

Covariances

Lag Time in Years

Cell Mean Model and Cubic Contrasts
Murray's Model Covariance
Figure 4.15
Model-Based Estimates of Group Means
Covariate is CHL at Visit 1

Total Cholesterol

Year in Study

Expected Value Model: Cell Mean
Covariance Model: Unstructured
Figure 4.16
Covariance Plot: Cell Mean Model
Covariate is CHL at Visit 1

Expected Value Model: Cell Mean
Covariance Model: Unstructured
Figure 4.17
Model-Based Estimates of Group Means
Covariate is Mean CHL of Visits 1 & 2

Total Cholesterol

Year in Study

Expected Value Model: Cell Mean
Covariance Model: Unstructured
Figure 4.18
Covariance Plot: Cell Mean Model
Covariate is Mean CHL of Visits 1 & 2

Expected Value Model: Cell Mean
Covariance Model: Unstructured
Figure 4.19
Covariance Plot: Unstructured Covariance
Covariate is CHL at Visit 1

Covariances

Lag Time in Years

Cell Mean Model and Cubic Contrasts
Baseline Visits 2–5 are New Time Points
Figure 4.20
Covariance Plot: General Mixed Model
Covariate is CHL at Visit 1

Cell Mean Model and Cubic Contrasts
Baseline Visits 2-5 are New Time Points
Figure 4.21
Covariance Plot: Murray’s Model
Covariate is CHL at Visit 1

Cell Mean Model and Cubic Contrasts
Baseline Visits 2-5 are New Time Points
Figure 4.22
Covariance Plot: Unstructured Covariance
Covariate is Mean CHL of Visits 1 & 2

Cell Mean Model and Cubic Contrasts
Mean of Visits 3-5 is New Time Point
Figure 4.23
Covariance Plot: General Mixed Model
Covariate is Mean CHL of Visits 1 & 2

Cell Mean Model and Cubic Contrasts
Mean of Visits 3-5 is New Time Point
Figure 4.24
Covariance Plot: Murray’s Model
Covariate is Mean CHL of Visits 1 & 2

Cell Mean Model and Cubic Contrasts
Mean of Visits 3-5 is New Time Point
Figure 4.25
Model-Based Estimates of Group Means
Models: Covariate V1 vs 5 Added Values

Total Cholesterol

TRT, 5 added values     TRT, visit 1 cov
Plac, 5 added values    Plac, visit 1 cov

Expected Value Model: Cell Mean & Cubic
Covariance Model: Murray's Model
Figure 4.26
Model-Based Estimates of Group Means
Models: Mean Covariate vs 2 Added Means

Total Cholesterol

Year in Study

- TRT, 2 added means
- TRT, mean covariate
- Plac, 2 added means
- Plac, mean covariate

Expected Value Model: Cell Mean & Cubic
Covariance Model: Murray's Model
REFERENCES


