<table>
<thead>
<tr>
<th>Title</th>
<th>Longitudinal analysis of Hamilton Depression Rating Scale (HDRS) scores (Statistical Inference of Records and Related Statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Matsumoto, Masato</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 1439: 185-201</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2005-07</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/47526">http://hdl.handle.net/2433/47526</a></td>
</tr>
<tr>
<td>Textversion</td>
<td>Departmental Bulletin Paper</td>
</tr>
</tbody>
</table>

Kyoto University
Longitudinal analysis of Hamilton Depression Rating Scale (HDRS) scores

北里大学 薬学研究科 松本正人（Masato Matsumoto）
CENTER FOR CLINICAL PHARMACY AND
CLINICAL SCIENCES KITASATO UNIVERSITY GRADUATE SCHOOL

ABSTRACT

Antidepressants are generally evaluated on the basis of the Hamilton Depression Scale Scores of the same patients measured repeatedly over time. The usual analysis of the scores measured at the end of the treatment period alone is, however, inadequate. To clarify the characteristic features of the test drugs, it is necessary to analyze the longitudinal patterns.

In this paper, we have analyzed actual clinical trial data in terms of longitudinal change of the score of individual subjects classified into three patterns (1. No variation, 2. Linear improvement, and 3. Early improvement). The clinical validity and usefulness of the analytical method presented are also examined.

Keywords: antidepressant, evaluation of drug effect, repeated measurements

INTRODUCTION

For the treatment of depression, TCAs (tricyclic antidepressants) have been widely used so far. In 1999, SSRI (Selective Serotonin Reuptake Inhibitor) was put on Japanese market. After that, other SSRIs and SNRI (Serotonin Noradrenaline Reuptake Inhibitor) were put on the market. From the many antidepressants, a proper antidepressant is chosen for each patient. For the proper choice, it is meaningful to characterize the antidepressants. In actual, a lot of meta-analyses (Examples are [1-8] .) and the comparison examinations (Examples are [9-12].) have been already performed.

The effects of antidepressants are generally evaluated using Hamilton Depression Rating Scale (HDRS) introduced by Max Hamilton in 1960 [13-15]. HDRS consists of 17 items and the total score of the 17 items is used for the measure of severity of depression. The maximum and minimum of the total score 48 points and 0
point, respectively. In clinical trials, HDRS scores are repeatedly measured on each patient. A decrease in the total indicates the improvement in the symptoms.

The efficacy of antidepressant is evaluated based on the mean of the decrease of HDRS scores at the final measurement point. In the current evaluation, however, the longitudinal pattern of HDRS scores of each patient is not considered. From clinical viewpoints, the evaluation is not appropriate. Longitudinal patterns of HDRS scores after the administration of an antidepressant can be grouped into the three patterns shown in Figure 1. In Pattern-1, pretreatment scores are maintained. This pattern corresponds to non-responders. Pattern-2 and Pattern-3 correspond to responders. In Pattern-2, HDRS scores decrease almost linearly. In Pattern-3, the scores decrease more rapidly. The patient population can be considered as a mixture of patients with the three patterns. We here suppose two drugs, Drug-1 and Drug-2, for which the mixing proportions of the three patterns are listed in Table 1.

<table>
<thead>
<tr>
<th>Drug</th>
<th>Pattern-1</th>
<th>Pattern-2</th>
<th>Pattern-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drug-1</td>
<td>20%</td>
<td>80%</td>
<td>0%</td>
</tr>
<tr>
<td>Drug-2</td>
<td>20%</td>
<td>40%</td>
<td>40%</td>
</tr>
</tbody>
</table>

If the evaluation is made based only on the mean of the decrease at the final measurement point, the proportion of responders is 80% in either drug. However, 40% of the patients in Drug-2 show Pattern-3 and respond more rapidly. It is clear that Drug-2 is clinically more preferable. Such an evaluation can not be made if the longitudinal patterns of HDRS scores are not considered. The efficacy of antidepressants should be evaluated based on the longitudinal patterns of HDRS scores.

We apply mixture models to actual clinical data of HDRS scores. We assume the following three patterns for the longitudinal patterns of HDRS scores, 1. No improvement pattern, 2. Linear improvement pattern, and 3. Early improvement pattern.

In applying mixture models, it is common to assume that longitudinal patterns can be described by low-degree polynomials of elapsed time after the beginning of treatment [16-19]. However, the low-degree polynomial models are not necessarily appropriate for describing the longitudinal patterns of HDRS scores. In Chapter 3, We propose a model using a monotone decreasing function to describe the early
improvement pattern. Furthermore, we investigate variance-covariance structures within a subject. In Chapter 4, we conduct simulation studies to evaluate the performance of the proposed model in Chapter 3. In the chapter of discussion, we arrange the result in this study. We derive the conclusion by present. And we refer the problem of the proposed method and the view of the future.

MOTIVATING EXAMPLE

The present data are HDRS scores of 84 patients in a randomized, double-blind, comparative study of antidepressants. The criteria for selecting the subjects are that the total score for HDRS items 1-17 was 16 or higher and the depressive mood score of was 2 or higher, before the start of the treatment. The antidepressants were given for 4 weeks, following a fixed-flexible regimen. The main item of evaluation was the final general improvement rating (FGIR) evaluated by the physicians, taking into account the changes in the HDRS scores and the clinical symptoms. FGIR was classified into eight categories, i.e., significant improvement, moderate improvement, mild improvement, no change, slight worsening, worsening, serious worsening and impossible to evaluate. The HDRS scores were evaluated at five measurement points, i.e., before the treatment and 1, 2, 3 and 4 weeks after the beginning of the treatment. The individual and mean profiles of HDRS scores are shown in Figure 2 and Figure 3, respectively. FGIR classification for the present data is shown in Table 2.
Figure 1.

Pattern 1

Pattern 2

Pattern 3
Table 2. FGIR classification for the present data

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25</td>
<td>21</td>
<td>18</td>
<td>11</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>(29.8%)</td>
<td>(54.8%)</td>
<td>(28.2%)</td>
<td>(89.3%)</td>
<td>(85.2%)</td>
<td>(100)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Significant improvement, 2. Moderate improvement, 3. Mild improvement, 4. No change,
5. Slight worsening, 6. Worsening, 7. Serious worsening, 8. impossible to evaluate

The cumulative percentages are shown in the parentheses.

Figure 2. Individual profiles of HDRS scores

Figure 3. Mean profile and Standard Error of HDRS scores
MIXTURE DISTRIBUTION MODEL FOR LONGITUDINAL DATA

Mixture distribution models are often applied to the analysis of longitudinal patterns of repeated measurements [16-19].

In this chapter, mixture distribution models are applied to HDRS score data obtained in an actual clinical trial of antidepressants.

(1) The model

As stated in the first chapter, longitudinal patterns of HDRS scores after the administration of antidepressants are grouped into three patterns. We define the three patterns as follows.

1. No improvement pattern: the scores show no improvement maintaining the pretreatment scores.
2. Linear improvement pattern: the scores show almost linear improvement.
3. Early improvement pattern: the scores show rapid improvement.

These three patterns correspond to the three patterns shown in Figure 1. All of the subjects are assumed to belong to one of the three patterns.

Let \( y_{ij} \) denote the HDRS score of the patient \( i \) (\( i = 1, \cdots, n \)) at the measurement point \( t_j \) (\( j = 1, \cdots, 5 \)). In the mixture distribution model, the probability density function of the observation vector \( \mathbf{y}_i = (y_{i1}, \cdots, y_{i5}) \) is given by

\[
g (\mathbf{y}_i \mid \mathbf{p}, \mathbf{\theta}) = \sum_{m=1}^{3} p_m \cdot f_m (\mathbf{y}_i \mid \theta_m),
\]

where \( \mathbf{y}_i = (y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5}) \) is the measurement vector for the patient \( i \), \( \mathbf{p} = (p_1, p_2, p_3) \) \((p_1 + p_2 + p_3 = 1)\) is the vector of the mixing proportions of the three patterns, \( f_m (\bullet) \) is the density function for the \( m \)-th pattern \((m = 1, 2, 3)\), \( \theta_m \) is the vector of the parameters that define the density function \( f_m (\bullet) \) \((m = 1, 2, 3)\), \( \mathbf{\theta} = (\theta_1, \theta_2, \theta_3) \) denotes the vector of all the parameters in \( \theta_1, \theta_2 \) and \( \theta_3 \).

For the three longitudinal patterns stated above, We assume the following model.

1. No improvement pattern

\[ y_{ij} = (\alpha_i + b_{ij}) + \epsilon_{3ij} \]
2. Linear improvement pattern
\[ y_{ij} = (\alpha_{2} + b_{2i}) + \beta_{2} \cdot t_{j} + \epsilon_{2ij} \]

3. Early improvement pattern
\[ y_{ij} = \exp\left(-\left(t_{j}/\alpha_{3}\right)^{\beta_{3}}\right) \cdot (\gamma_{3} + b_{3i}) + \epsilon_{3ij} \]

In this model, it is assumed that the pretreatment scores \( a_{1}, a_{2}, \) and \( a_{3} \) are common to all the patients, \( b_{1i}, b_{2i} \) and \( b_{3i} \) are the patient-specific variations of the pretreatment scores normally distributed as \( b_{1i}\sim\text{N}(0, s_{b12}), b_{2i}\sim\text{N}(0, s_{b22}) \) and \( b_{3i}\sim\text{N}(0, s_{b32}) \), respectively and \( \epsilon_{1ij}, \epsilon_{2ij} \) and \( \epsilon_{3ij} \) is the error term normally distributed with mean 0 and variance-covariance matrix \( \Sigma_{s1}, \Sigma_{e1} \) and \( \Sigma_{e3} \), respectively. The function of early improvement pattern comes from the following:
\[
1 - \text{(the Weibull distribution function)} = 1 - (1 - \exp(-t/\alpha_{3})) = \exp(-t/\alpha_{3})^{\beta_{3}}.
\]

This function is parsimonious and useful for describing monotone decreasing function. In addition, this function can be used for describing the feature of HDRS pattern that the variance for the early improvement pattern becomes smaller as the clinical trial advances. The details are given later.

(2) The variance-covariance within a patient

For the variance-covariance matrices of the error terms, the following three structures are employed: simple variance (SV), first-order autoregressive (AR(1)), and toeplitz (TOEP), which are commonly used in the analysis of clinical longitudinal data [20].

When SV is assumed, the variance-covariance matrix of the marginal distribution becomes a compound symmetry type in the no improvement pattern and linear improvement pattern as follows:
\[
[1]
\begin{bmatrix}
1 \\
1 \\
1 \\
1 \\
1
\end{bmatrix}
[\sigma_{bm}^{2} \{ 1 1 1 1 \} + \sigma_{\epsilon_{0}0^{m}}^{2}0000 \sigma_{m}^{2}0000 \sigma_{m}^{2}0000 \sigma_{m}^{2}0000 \sigma_{b_{m}}000 \ovalbox{\tt\small REJECT} 0_{2}^{m}], m=1, 2.
\]

In the early improvement pattern, the variance-covariance matrix is given by
\[
\begin{bmatrix}
\exp\left(-\frac{0}{\alpha_3}\right)^{\beta_3} \\
\exp\left(-\frac{1}{\alpha_3}\right)^{\beta_3} \\
\exp\left(-\frac{2}{\alpha_3}\right)^{\beta_3} \\
\exp\left(-\frac{3}{\alpha_3}\right)^{\beta_3} \\
\exp\left(-\frac{4}{\alpha_3}\right)^{\beta_3}
\end{bmatrix}
\begin{bmatrix}
\exp\left(-\frac{0}{\alpha_3}\right)^{\beta_3} \\
\exp\left(-\frac{1}{\alpha_3}\right)^{\beta_3} \\
\exp\left(-\frac{2}{\alpha_3}\right)^{\beta_3} \\
\exp\left(-\frac{3}{\alpha_3}\right)^{\beta_3} \\
\exp\left(-\frac{4}{\alpha_3}\right)^{\beta_3}
\end{bmatrix}^T
\begin{bmatrix}
\sigma_{\epsilon_3}^2 & 0 & 0 & 0 & 0 \\
0 & \sigma_{\epsilon_3}^2 & 0 & 0 & 0 \\
0 & 0 & \sigma_{\epsilon_3}^2 & 0 & 0 \\
0 & 0 & 0 & \sigma_{\epsilon_3}^2 & 0 \\
0 & 0 & 0 & 0 & \sigma_{\epsilon_3}^2
\end{bmatrix}.
\]

From this structure, it is found that the variance becomes smaller as the clinical trial advances and that the covariance becomes smaller as the interval between the measurement points becomes longer.

(3) Results

Table 3 shows the number of the parameters, maximum log likelihood and AIC [21-23] for the three variance-covariance structures, SV, AR(1), or TOEP. The AICs indicate that the variance-covariance structure AR(1) is the best among the three structures.

<table>
<thead>
<tr>
<th>The number of the parameters</th>
<th>SV (Parameter)</th>
<th>AR(1) (Parameter)</th>
<th>TOEP (Parameter)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum log likelihood</td>
<td>-1149.2</td>
<td>-1114.8</td>
<td>-1112.2</td>
</tr>
<tr>
<td>AIC</td>
<td>2326.3</td>
<td>2265.6</td>
<td>2276.4</td>
</tr>
</tbody>
</table>

The result when AR(1) is assumed is shown as follows. Table 4 lists the maximum likelihood estimates of the parameters and their standard errors. Figure 4 shows the estimated mean profiles of the HDRS scores.
Table 4. The maximum likelihood estimates (MLE) and their standard errors

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Parameter</th>
<th>MLE</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. No improvement pattern</td>
<td>$p_1$</td>
<td>0.298</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>$q_1$</td>
<td>22.4</td>
<td>1.762</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{b_1}^2$</td>
<td>8.43</td>
<td>2.639</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{e_1}^2$</td>
<td>40.4</td>
<td>2.349</td>
</tr>
<tr>
<td></td>
<td>$\rho_1$</td>
<td>0.875</td>
<td>0.008</td>
</tr>
<tr>
<td>2. Linear improvement pattern</td>
<td>$p_2$</td>
<td>0.581</td>
<td>0.122</td>
</tr>
<tr>
<td></td>
<td>$q_2$</td>
<td>23.4</td>
<td>0.818</td>
</tr>
<tr>
<td></td>
<td>$\beta_2$</td>
<td>-4.0</td>
<td>0.285</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{b_2}^2$</td>
<td>0.00</td>
<td>3.847</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{e_2}^2$</td>
<td>33.6</td>
<td>3.893</td>
</tr>
<tr>
<td></td>
<td>$\rho_2$</td>
<td>0.553</td>
<td>0.066</td>
</tr>
<tr>
<td>3. Early improvement pattern</td>
<td>$p_3$</td>
<td>0.121</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$q_3$</td>
<td>0.283</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>$\beta_3$</td>
<td>0.309</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>$\gamma_3$</td>
<td>23.5</td>
<td>3.554</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{b_3}^2$</td>
<td>31.9</td>
<td>9.686</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{e_3}^2$</td>
<td>11.5</td>
<td>0.477</td>
</tr>
<tr>
<td></td>
<td>$\rho_3$</td>
<td>0.859</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Figure 4. The estimated mean profiles of HDRS scores for the three patterns

![Graph showing HDRS scores over time for three patterns with percentages 29.8%, 58.1%, and 12.1%](image_url)
Given the estimates for all the parameters, the probabilities that the patient i with the data $y_i$ belongs to each of the three patterns can be estimated by Bayes theorem [24-26]. By assuming that each patient belongs to the pattern for which the probability is the largest, the patients can be classified into the three patterns. The proportions of the patients classified into the three patterns are 27.4% (23/84: No improvement), 60.7% (51/84: Linear improvement) and 11.9% (10/84: Early improvement). The relationship between the classification and FGIR measured in the clinical trial is shown in Table 5. Figure 5 shows the individual profiles of the patients classified into the three patterns.

Table 5. The relationship between the classification and FGIR

<table>
<thead>
<tr>
<th>Pattern</th>
<th>FGIR 1</th>
<th>FGIR 2</th>
<th>FGIR 3</th>
<th>FGIR 4</th>
<th>FGIR 5</th>
<th>FGIR 6</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>No improvement</td>
<td>0.0</td>
<td>0.0</td>
<td>0.10</td>
<td>0.08</td>
<td>0.02</td>
<td>0.03</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(43.5)</td>
<td>(34.8)</td>
<td>(8.7)</td>
<td>(13.0)</td>
<td>(100)</td>
</tr>
<tr>
<td>Linear improvement</td>
<td>0.15</td>
<td>0.21</td>
<td>0.08</td>
<td>0.03</td>
<td>0.03</td>
<td>0.01</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>(29.4)</td>
<td>(41.2)</td>
<td>(15.7)</td>
<td>(5.9)</td>
<td>(5.9)</td>
<td>(2.0)</td>
<td>(100)</td>
</tr>
<tr>
<td>Early improvement</td>
<td>0.10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(100.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(100)</td>
</tr>
<tr>
<td>Sum</td>
<td>0.25</td>
<td>0.21</td>
<td>0.18</td>
<td>0.11</td>
<td>0.05</td>
<td>0.04</td>
<td>0.84</td>
</tr>
</tbody>
</table>

4. No change, 5. Slight worsening, 6. Worsening

The percentages to the total in each pattern are shown in the parentheses.
Figure 5. Individual profiles of the patients classified into the three patterns

1. No improvement pattern: n=23

2. Linear improvement pattern: n=51

3. Early improvement pattern: n=10
From the results assuming the variance-covariance structure AR(1), the following points can be found.

i. The estimated pretreatment scores are 22.4, 23.4 and 23.5 for the no improvement pattern, linear improvement pattern and early improvement pattern, respectively. There seems to be no great differences among the three patterns.

ii. The estimated mixing proportions are 29.8%, 58.1%, and 12.1% for the no improvement pattern, linear improvement pattern, and early improvement pattern, respectively. About 70% of the patients belong to either the linear improvement pattern or early improvement pattern.

iii. The estimated scores at 1 week after the beginning of the treatment are 19.4 and 5.4 in the linear improvement pattern and early improvement pattern, respectively. The estimated scores at 2 weeks are 15.4 and 3.8 in the linear improvement pattern and early improvement pattern, respectively. The results suggest that the HDRS scores had been improved clinically well enough at 1 week in the early improvement pattern.

iv. The estimated scores at the final measurement point (4 weeks after the beginning of the treatment) are 22.4, 7.4 and 2.4 in the no improvement pattern, linear improvement pattern and early improvement pattern, respectively. The HDRS scores were improved in both the linear and early improvement patterns.

v. The estimated probabilities that each patient belong to each of the three patients range from 0.504 to 1.000 with mean 0.892. 25, 50 and 75 percentiles are 0.836, 0.966 and 0.998, respectively.

vi. The proportions of the patients classified into the three patterns are 27.4% (23/84), 60.7% (51/84), and 11.9% (10/84) for the no improvement pattern, linear improvement pattern, and early improvement pattern, respectively. These are almost the same as the estimates of the mixing proportion.

vii. The relationship between the results of the classification and FGIR indicates that all the patients classified into the early improvement pattern showed the significant improvement in FGIR and that about 85 % of the patients classified into the linear improvement pattern showed the mild or better improvement in FGIR. On the other hand, the patients classified into the no improvement pattern...
did not show the moderate or better improvement in FGIR.

SIMULATION STUDY: DETECTION OF TRUE VARIANCE-COVARIANCE STRUCTURE

When repeated measurements of HDRS scores are analyzed using the mixture distribution model consisting of the three patterns (no improvement, linear improvement, and early improvement) presented in Chapter 3, it is important to examine the influence of the assumption of the within-subject covariance structure on the parameter estimates. In this chapter, the following two simulation studies are conducted to examine this issue.

In the simulation study, we suppose the situation in which the model with the true within-subject covariance structure is included in the applied models. Under this situation, it is examined whether the selected model can detect the true structure of the within-subject covariance. In addition, we examine the influence of the mis-specified within-subject covariance structure on the accuracy of the parameter estimates.

In this simulation study, the following three structures, SV, AR(1), and TOEP are assumed for the within-subject covariance structure. Table 6 shows the true values of the parameters. These values are determined by referring to the results in Chapter 3.

Under the true structure, 100 data sets are simulated. Each data set consists of the data of 100 subjects. For each data set, the three mixture distribution models with the within-subject covariance structure SV, AR(1), and TOEP are applied and the goodness of each model is evaluated based on the AIC[21-23].

Table 7 shows the proportions that each mixture distribution model is selected based on AIC. The proportion that the true within-subject covariance structure model is selected is about 95% for each of the three within-subject covariance structure. This result suggests that the proposed approach can select the true within-subject covariance structure under the situation in which the model with the true within-subject covariance structure is included in the applied models.

The description of the result is omitted, and the following is confirmed. The accuracy of the estimates is especially worsened for the following cases: SV is assumed when the true structure is AR, and SV or AR(1) is assumed when the true structure is TOEP.
Table 6. The true values of the parameters for Simulation study

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Proportion</th>
<th>Mean structure</th>
<th>Variance of intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>No improvement</td>
<td>p1=30%</td>
<td>y_{ij}=23</td>
<td></td>
</tr>
<tr>
<td>Linear improvement</td>
<td>p2=50%</td>
<td>y_{ij}=23-4\cdot t</td>
<td>10</td>
</tr>
<tr>
<td>Early improvement</td>
<td>p3=20%</td>
<td>y_{ij}=\exp(-{(y/1.0)^{0.55}})\cdot25</td>
<td>40</td>
</tr>
</tbody>
</table>

Y_{ij}: the observation of the patient i at the measurement point j, t: the measurement point

<table>
<thead>
<tr>
<th>True structure</th>
<th>Pattern</th>
<th>Variance covariance matrix within subject: $\sigma^2_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>No improvement</td>
<td>15 (i=j)</td>
</tr>
<tr>
<td></td>
<td>Linear improvement</td>
<td>20 (i=j)</td>
</tr>
<tr>
<td></td>
<td>Early improvement</td>
<td>15 (i=j)</td>
</tr>
<tr>
<td>AR (1)</td>
<td>No improvement</td>
<td>15 (i=j)</td>
</tr>
<tr>
<td></td>
<td>Linear improvement</td>
<td>20 (i=j)</td>
</tr>
<tr>
<td></td>
<td>Early improvement</td>
<td>15 (i=j)</td>
</tr>
<tr>
<td>TOEP</td>
<td>No improvement</td>
<td>i=j</td>
</tr>
<tr>
<td></td>
<td>Linear improvement</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Early improvement</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
</tr>
</tbody>
</table>

Table 7. The proportions that each mixture distribution model is selected by AIC

<table>
<thead>
<tr>
<th>True structure</th>
<th>Assumed structure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SV</td>
</tr>
<tr>
<td>SV</td>
<td>0.95</td>
</tr>
<tr>
<td>AR(1)</td>
<td>0.00</td>
</tr>
<tr>
<td>TOEP</td>
<td>0.00</td>
</tr>
</tbody>
</table>
It is problem from clinical viewpoints that the efficacy of antidepressant is evaluated based on the mean of the decrease of HDRS scores at the final measurement point, because the longitudinal pattern of HDRS scores of each patient is not considered. In the present study, we have evaluated the results on the basis of longitudinal patterns. By evaluating the average changes in each pattern and the mixing proportions, we could quantitatively evaluate the early onset of a characteristic feature of the drug. The analyses at each time point cause the statistical problem of multiplicity, and the results are difficult to understand. Because the objective of the analyses at each time point is to evaluate on the longitudinal patterns, the evaluation is possible by this method.

The results of this study and the clinical evaluation (FGIR) of the subjects have a certain level of agreement. Therefore, we can conclude that this method is appropriate from the clinical point of view. The results suggest that FGIR is an evaluation in which longitudinal patterns are taken into account. By classifying the subjects into one of the three patterns, we can examine the differences in background factors among subjects having these different patterns.

By applying this method, it is possible to execute comparison between drugs by the mixture proportions of the drug. The null hypothesis of comparison between Drug-1 and Drug-2 in this case is as follows.

\[
H_0: p_{m,\text{Drug-1}} = p_{m,\text{Drug-2}} : m=1,2,3,
\]

where \( p_m \) is mixing proportion of \( m \)-th pattern.

This part will need examining in the future.

One problem with this method is how to decide on the number of patterns to be used. This has not been solved in the present study. In this study, the analysis is done assuming 3 patterns, taking into account the observed data and an easiness of clinical explanation. When the number of patterns is decided, we should decide it in consideration of a feature of the drug and a clinical meaning.

In the analysis of this study, we assumed that all the subjects belonged to one of the three variation patterns. It is quite possible that data of some subjects may be intermediate between two patterns in fact. Areas to be studied in the future include the problem of how to handle patients who are difficult to belong to any one pattern.
The variance-covariance structure within a patient is actually unknown. It is necessary to investigate the influence of misspecification the variance-covariance structure within a patient. In actual analyses, it is quite difficult to specify the correct variance-covariance structure within a subject. It will be desirable to use a robust estimation method against the mis-specification of variance-covariance structure within a subject.

REFERENCES


15. Williams JBW. Arch Gen Psychiatry 1988;45: 742-747


