General Approaches to Analysis of Course

Applying Growth Mixture Modeling to Randomized Trials of Depression Medication

BENGT MUTHÉN, HENDRICKS C. BROWN, AIMEE M. HUNTER,

IAN A. COOK, AND ANDREW F. LEUCHTER

Introduction

This chapter discusses the assessment of treatment effects in longitudinal randomized trials using growth mixture modeling (GMM) (Muthén & Shedden, 1999; Muthén & Muthén, 2000; Muthén et al., 2002; Muthén & Asparouhov, 2008). GMM is a generalization of conventional repeated measurement mixed-effects (multilevel) modeling. It captures unobserved subject heterogeneity in trajectories not only by random effects but also by latent classes corresponding to qualitatively different types of trajectories. It can be seen as a combination of conventional mixed-effects modeling and cluster analysis, also allowing prediction of class membership and estimation of each individual's most likely class membership. GMM has particularly strong potential for analyses of randomized trials because it responds to the need to investigate for whom a treatment is effective by allowing for different treatment effects in different trajectory classes.

The chapter is motivated by a University of California–Los Angeles study of depression medication (Leuchter, Cook, Witte, Morgan, & Abrams, 2002). Data on 94 subjects are drawn from a combination of three studies carried out with the same design, using three different types of medications: fluoxetine ($n = 14$), venlafaxine IR ($n = 17$), and venlafaxine XR ($n = 18$). Subjects were measured at baseline and again after a 1-week placebo lead-in phase. In the subsequent double-blind phase of the study, the subjects were randomized into medication ($n = 49$) and placebo ($n = 45$) groups. After randomization, subjects were measured at nine occasions: at 48 hours and at weeks 1–8. The current analyses consider the Hamilton Depression Rating Scale.

Several predictors of course of the Hamilton scale trajectory are available, including gender, treatment history, and a baseline measure of central cordance hypothesized to influence tendency to respond to treatment.

The results of studies of this kind are often characterized in terms of an end point analysis where the outcome at the end of the study, here at 8 weeks, is considered for the placebo group and for the medication group. A subject may be classified as a responder by showing a week 8 depression score below 10 or when dropping below 50% of the initial score. The treatment effect may be assessed by comparing the medication and placebo groups with respect to the ratio of responders to nonresponders.

As an alternative to end point analysis, conventional repeated measurement mixed-effects (multilevel) modeling can be used. Instead of focusing on only the last time point, this uses the outcome at all time points, the two pretreatment occasions and the nine posttreatment occasions. The trajectory shape over time is of key interest and is estimated by a model that draws on the information from all time points. The idea of considering trajectory shape in research on depression medication has been proposed by Quitkin et al. (1984), although not using a formal statistical growth model.

Rates of response to treatment with antidepressant drugs are estimated to be 50%–60% in typical patient populations. Of particular interest in this chapter is how to assess treatment effects in the presence of a placebo response. A *placebo response* is an improvement in depression ratings seen in the placebo group that is unrelated to medication. The improvement is often seen as an early steep drop in depression, often followed by a later upswing. An example is seen in Figure 7–2. A placebo response confounds the estimation of the true effect of medication and is an important phenomenon given its high prevalence of 25%–60% (Quitkin, 1984). Because the placebo response is pervasive, the statistical modeling must take it into account when estimating medication effects. This can be done by acknowledging the qualitative heterogeneity in trajectory shapes for responders and nonresponders.

It is important to distinguish among responder and nonresponder trajectory shapes in both the placebo and medication groups. Conventional repeated measures modeling may lead to distorted assessment of medication effects when individuals follow several different trajectory shapes. GMM avoids this problem while maintaining the repeated measures modeling advantages. The chapter begins by considering GMM with two classes, a nonresponder class and a responder class. The *responder class* is defined as those individuals who respond in the placebo group and who would have responded to placebo among those in the medication group. Responder class membership is observed for subjects in the placebo group but is unobserved in the medication group. Because of randomization, it can be assumed that this class of subjects is present in both the placebo and medication groups

and in equal numbers. GMM can identify the placebo responder class in the medication group. Having identified the placebo responder and placebo non-responder classes in both the placebo and medication groups, medication effects can more clearly be identified. In one approach, the medication effect is formulated in terms of an effect of medication on the trajectory slopes after the treatment phase has begun. This medication effect is allowed to be different for the nonresponder and responder trajectory classes. Another approach formulates the medication effect as increasing the probability of membership in advantageous trajectory classes and decreasing the probability of membership in disadvantageous trajectory classes.

Growth Mixture Modeling

This section gives a brief description of the GMM in the context of the current study. A two-piece, random effect GMM is applied to the Hamilton Depression Rating Scale outcomes at the 11 time points y_1–y_{11}. The first piece refers to the two time points y_1 and y_2 before randomization, and the second piece refers to the nine postrandomization time points y_3–y_{11}. Given only two time points, the first piece is by necessity taken as a linear model with a random intercept, defined at baseline, and a fixed effect slope. An exploration of each individual's trajectory suggests a quadratic trajectory shape for the second piece. The growth model for the second piece is centered at week 8, defining the random intercept as the systematic variation at that time point. All random effect means are specified as varying across latent trajectory classes. The medication effect is captured by a regression of the linear and quadratic slopes in the second piece on a medication dummy variable. These medication effects are allowed to vary across the latent trajectory classes. The model is shown in diagrammatic form at the top of Figure 7–1.[1]

The statistical specification is as follows. Consider the depression outcome y_{it} for individual i, let c denote the latent trajectory class variable, let η denote random effects, let a_t denote time, and let $\in_t$ denote residuals containing measurement error and time-specific variation. For the first, prerandomization piece, conditional on trajectory class k ($k = 1, 2 \ldots K$),

---

1. In Figure 7–1 the observed outcomes are shown in boxes and the random effects in circles. Here, i, s, and q denote intercept, linear slope, and quadratic slope, respectively. In the following formulas, these random effects are referred to as η_0, η_1, and η_2. The treatment dummy variable is denoted x.
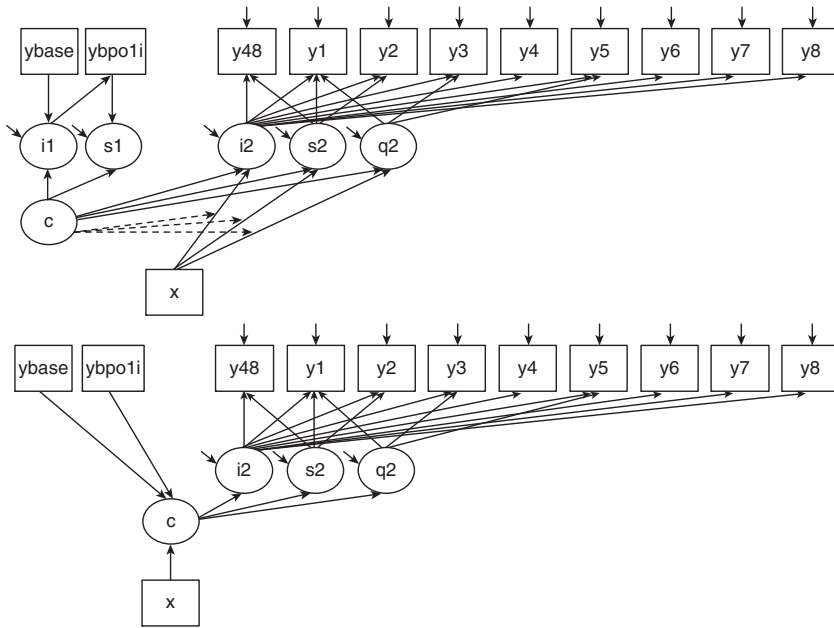
Figure 7.1 Two alternative GMM approaches.

$$Y_{it}^{pre}\big|_{c_i=k} = \eta_{0i}^{pre} = \eta_{1i}^{pre}a_t + \epsilon_{it}^{pre}, \tag{1}$$

with $\alpha_1 = 0$ to center at baseline, and random effects

$$\eta_{10i}^{pre}\big|_{c_i=k} = \alpha_{10k} + \zeta_{10i}, \tag{2}$$

$$\eta_{11i}^{pre}\big|_{c_i=k} = \alpha_{11k} + \zeta_{11i}, \tag{3}$$

with only two prerandomization time points, the model is simplified to assume a nonrandom slope, $V(\zeta_{11}) = 0$, for identification purposes. For the second, postrandomization piece,

$$y_{it}\big|_{c_i=k} = \eta_{0i} + \eta_{1i}a_t + \eta_{2i}a_t^2 + \epsilon_{it}, \tag{4}$$

with $\alpha_{11} = 0$, defining η_{0i} as the week 8 depression status. The remaining α_t values are set according to the distance in timing of measurements. Assume for simplicity a single drug and denote the medication status for individual i by the dummy variable x_i ($x = 0$ for the placebo group and $x = 1$ for the medication group).[2] The random effects are allowed to be influenced by

2. In the application three dummy variables are used to represent the three different medications.

group and a covariate, w, their distributions varying as a function of trajectory class (k),

$$\eta_{0i}\big|_{c_i=k} = \alpha_{0k} + \gamma_{01k}x_i + \gamma_{02k}w_i + \zeta_{0i}, \tag{5}$$

$$\eta_{1i}\big|_{c_i=k} = \alpha_{1k} + \gamma_{11k}x_i + \gamma_{12k}w_i + \zeta_{1i}, \tag{6}$$

$$\eta_{2i}\big|_{c_i=k} = \alpha_{2k} + \gamma_{21k}x_i + \gamma_{22k}w_i + \zeta_{2i}, \tag{7}$$

The residuals ζ_i in the first and second pieces have a 4×4 covariance matrix Ψ_k, here taken to be constant across classes k. For both pieces the residuals $\in_{it}$ have a $T \times T$ covariance matrix Θ_k, here taken to be constant across classes. For simplicity, Ψ_k and Θ_k are assumed to not vary across treatment groups. As seen in equations 5–7, the placebo group ($x_i = 0$) consists of subjects from the two different trajectory classes that vary in the means of the growth factors, which in the absence of covariate w are represented by α_{0k}, α_{1k}, and α_{2k}. This gives the average depression development in the absence of medication. Because of randomization, the placebo and medication groups are assumed to be statistically equivalent at the first two time points. This implies that x is assumed to have no effect on η_{10i} or η_{11i} in the first piece of the development. Medication effects are described in the second piece by γ_{01k}, γ_{11k}, and γ_{21k} as a change in average growth rate that can be different for the classes.

This model allows the assessment of medication effects in the presence of a placebo response. A key parameter is the medication-added mean of the intercept random effect centered at week 8. This is the γ_{01k} parameter of equation 5. This indicates how much lower or higher the average score is at week 8 for the medication group relative to the placebo group in the trajectory class considered. In this way, the medication effect is specific to classes of individuals who would or would not have responded to placebo. The modeling will be extended to allow for the three drugs of this study to have different γ parameters in equations 5–7.

Class membership can be influenced by baseline covariates as expressed by a logistic regression (e.g., with two classes),

$$log[P(c_i = 1|x_i)/P(c_i = 2\big|x_i)] = \alpha_c + \gamma_c w_i, \tag{8}$$

where $c = 1$ may refer to the nonresponder class and $c = 2$, the responder class. It may be noted that this model assumes that medication status does not influence class membership. Class membership is conceptualized as a quality characterizing an individual before entering the trial.

A variation of the modeling will focus on postrandomization time points. Here, an alternative conceptualization of class membership is used. Class membership is thought of as being influenced by medication so that the class probabilities are different for the placebo group and the three medication groups. Here, the medication effect is quantified in terms of differences across groups in class probabilities. This model is shown in diagrammatic form at the bottom of Figure 7–1. It is seen that the GMM involves only the postrandomization outcomes, which is logical given that treatment influences the latent class variable, which in turn influences the posttreatment outcomes. In addition to the treatment variable, pretreatment outcomes may be used as predictors of latent class, as indicated in the figure. The treatment and pretreatment outcomes may interact in their influence on latent class membership.

Estimation and Model Choice

The GMM can be fitted into the general latent variable framework of the Mplus program (Muthén & Muthén, 1998–2008). Estimation is carried out using maximum likelihood via an expectation-maximization (EM) algorithm. Missing data under the missing at random (MAR) assumption are allowed for the outcomes. Given an estimated model, estimated posterior probabilities for each individual and each class are produced. Individuals can be classified into the class with the highest probability. The classification quality is summarized in an entropy value with range 0–1, where 1 corresponds to the case where all individuals have probability 1 for one class and 0 for the others. For model fitting strategies, see Muthén et al. (2002), Muthén (2004), and Muthén and Asparouhov (2008). A common approach to decide on the number of classes is to use the Bayesian information criterion (BIC), which puts a premium on models with large log-likelihood values and a small number of parameters. The lower the BIC, the better the model. Analyses of depression trial data have an extra difficulty due to the typically small sample sizes. Little is known about the performance of BIC for samples as small as in the current study. Bootstrapped likelihood ratio testing can be performed in Mplus (Muthén & Asparouhov, 2008), but the power of such testing may not be sufficient at these sample sizes. Plots showing the agreement between the class-specific estimated means and the trajectories for individuals most likely belonging to a class can be useful in visually inspecting models but are of only limited value in choosing between models.

A complication of maximum-likelihood GMM is the presence of local maxima. These are more prevalent with smaller samples such as the current ones for the placebo group, the medication group, as well as for the combined sample. To be confident that a global maximum has been found, many

random starting values need to be used and the best log-likelihood value needs to be replicated several times. In the present analyses, between 500 and 4,000 random starts were used depending on the complexity of the model.

Growth Mixture Analyses

In this section the depression data are analyzed in three steps using GMM. First, the placebo group is analyzed alone. Second, the medication group is analyzed alone. Third, the placebo and medication groups are analyzed jointly according to the GMM just presented in order to assess the medication effects.

Analysis of the Placebo Group

A two-class GMM analysis of the 45 subjects in the placebo group resulted in the model-estimated mean curves shown in Figure 7–2. As expected, a responder class (class 1) shows a postrandomization drop in the depression score with a low of 7.9 at week 5 and with an upswing to 10.8 at week 8. An estimated 32% of the subjects belong to the responder class. In contrast, the nonresponder class has a relatively stable level for weeks 1–8, ending with a depression score of 15.6 at week 8. The sample standard deviation at week 8 is 7.6. It may be noted that the baseline score is only slightly higher for the nonresponder class, 22.7 vs. 21.9. The standard deviation at baseline is 3.6.[3]

The observed trajectories of individuals classified into the two classes are plotted in Figure 7–3a and b as broken lines, whereas the solid curves show the model-estimated means. The figure indicates that the estimated mean curves represent the individual development rather well, although there is a good amount of individual variation around the mean curves.

It should be noted that the classification of subjects based on the trajectory shape approach of GMM will not agree with that using end point analysis. As an example, the nonresponder class of Figure 7–3b shows two subjects with scores less than 5 at week 8. The individual with the lowest score at week 8, however, has a trajectory that agrees well with the nonresponder mean curve for most of the trial, deviating from it only during the last 2 weeks. The week 8 score has a higher standard deviation than at earlier time points, thereby weighting this time point somewhat less. Also, the data coverage due to

3. The maximum log-likelihood value for the two-class GMM of Figure 7–2 is 1,055.974, which is replicated across many random starts, with 28 parameters and a BIC value of 2,219. The classification based on the posterior class probabilities is not clear-cut in that the classification entropy value is only 0.66.
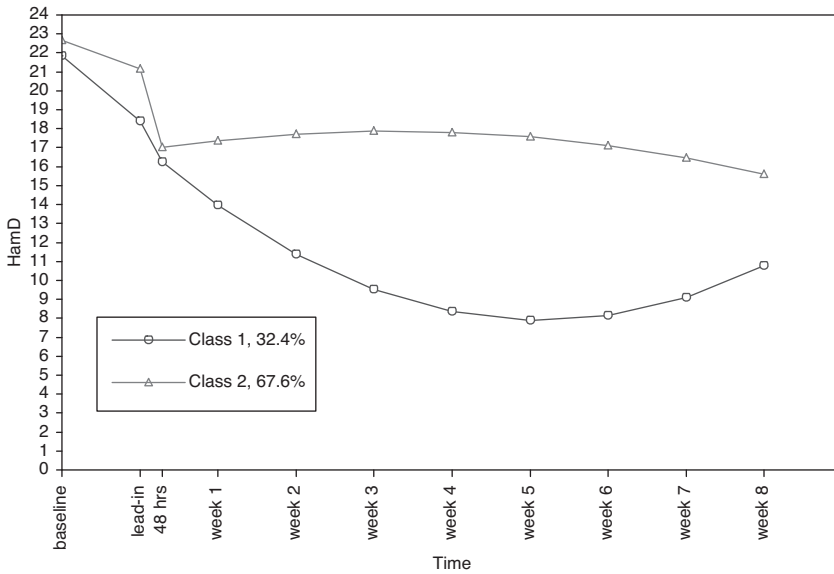
Figure 7.2 Two-class GMM for placebo group.

missing observations is considerably lower for weeks 5–7 than other weeks, reducing the weight of these time points. The individual with the second lowest score at week 8 deviates from the mean curve for week 5 but has missing data for weeks 6 and 7. This person is also ambiguously classified in terms of his or her posterior probability of class membership.

To further explore the data, a three-class GMM was also fitted to the 45 placebo subjects. Figure 7–4a shows the mean curves for this solution. This solution no longer shows a clear-cut responder class. Class 2 (49%) declines early, but the mean score does not go below 14. Class 1 (22%) ends with a mean score of 10.7 but does not show the expected responder trajectory shape of an early decline.[4] A further analysis was made to investigate if the lack of a clear responder class in the three-class solution is due to the sample size of $n = 45$ being too small to support three classes. In this analysis, the $n = 45$ placebo group subjects were augmented by the medication group subjects but using only the two prerandomization time points from the medication group. Because of randomization, subjects are statistically equivalent before randomization, so this approach is valid. The first, prerandomization piece of the GMM has nine parameters, leaving only 25 parameters to be estimated in the second, postrandomization piece by

4. The log-likelihood value for the model in Figure 7–4a is 1,048,403, replicated across several random starts, with 34 parameters and a BIC value of 2,226. Although the BIC value is slightly worse than for the two-class solution, the classification is better, as shown by the entropy value of 0.85.
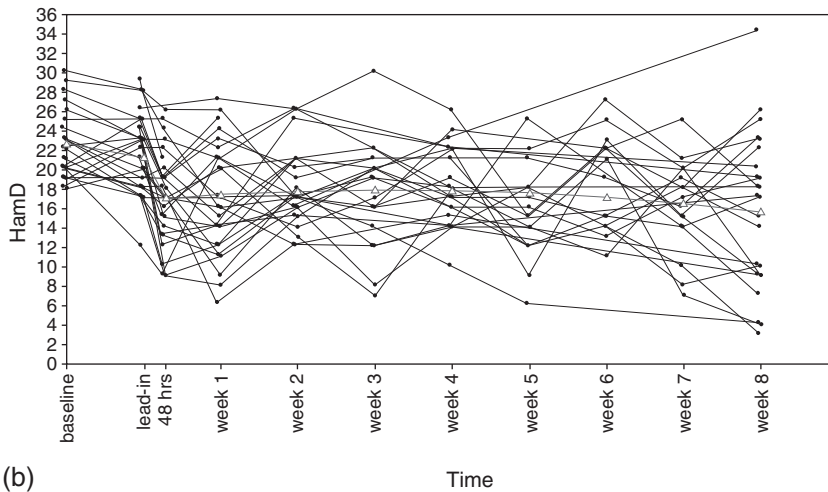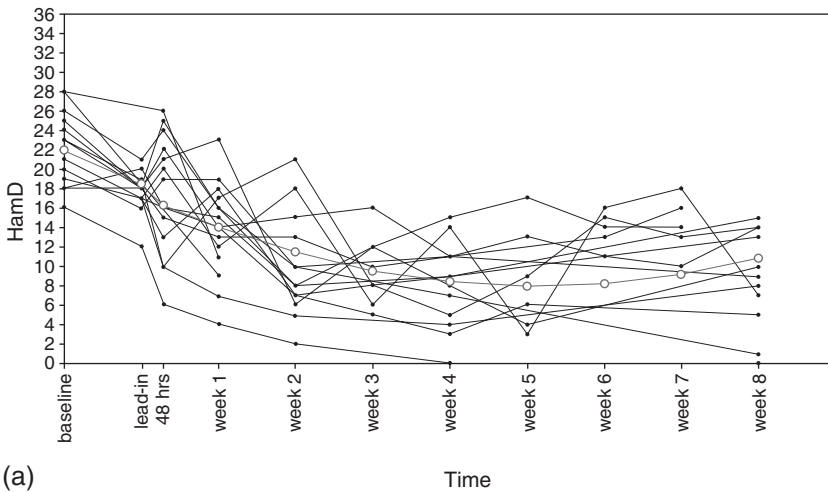
(a)   Time

(b)   Time

Figure 7.3 Individual trajectories for placebo subjects classified into (a) the responder class and (b) the non-responder class.

the n = 45 placebo subjects alone. Figure 7–4b shows that a responder class (class 2) is now found, with 21% of the subjects estimated to be in this class. High (class 3) and low (class 1) nonresponder classes are found, with 18% and 60% estimated to be in these classes, respectively. Compared to Figure 7–3, the observed individual trajectories within class are somewhat less hetero-geneous (trajectories not shown).[5]

5. The log-likelihood value for the model in Figure 7–4b is 1,270.030, replicated across several random starts, with 34 parameters and a BIC value of 2,695. The entropy value is 0.62. Because a different sample size is used, these values are not comparable to the earlier ones.
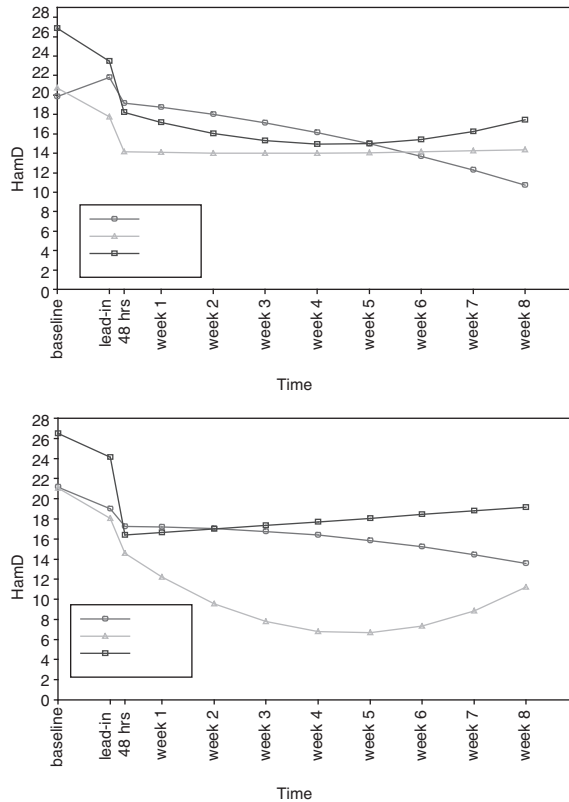
Figure 7.4 (a) Three-class GMM for placebo group. (b) Three-class GMM for placebo group and pre-randomization medication group individuals.

Analysis of the Medication Group

Two major types of GMMs were applied to the medication group. The first type analyzes all time points and either makes no distinction among the three drugs (fluoxetine, venlafaxine IR, venlafaxine XR) or allows drug differences for the class-specific random effect means of the second piece of the GMM. It would not make sense to also let class membership vary as a function of drug since class membership is conceptualized as a quality characterizing an individual before entering the trial. Class membership influences prerandomization outcomes, which cannot be influenced by drugs.

To investigate class membership, the second type of GMM analyzes the nine postrandomization time points both to focus on the period where the medications have an effect and to let the class membership correspond to only postrandomization variables. Here, not only are differences across the three drugs allowed for the random effect means for each of the classes but the drug type is also allowed to influence class probabilities.

Analysis of All Time Points

A two-class GMM analysis of the 49 subjects in the medication group resulted in the model-estimated mean curves shown in Figure 7–5. As expected, one of the classes is a large responder class (class 1, 85%). The other class (class 2, 15%) improves initially but then worsens.[6]

A three-class GMM analysis of the 49 subjects in the medication group resulted in the model-estimated mean curves shown in Figure 7–6. The three mean curves show the expected responder class (class 3, 68%) and the class (class 2, 15%) found in the two-class solution showing an initial improvement but later worsening. In addition, a nonresponse class (class 1, 17%) emerges, which has no medication effect throughout.[7]

Allowing for drug differences for the class-specific random effect means of the second piece of the GMM did not give a trustworthy solution in that the best log-likelihood value was not replicated. This may be due to the fact that this model has more parameters than subjects (59 vs. 49).

Analysis of Postrandomization Time Points

As a first step, two- and three-class analyses of the nine postrandomization time points were performed, not allowing for differences across the three drugs. This gave solutions that were very similar to those of Figures 7–5 and 7–6. The similarity in mean trajectory shape held up also when allowing for class probabilities to vary as a function of drug. Figure 7–7 shows the estimated mean curves for this latter model. The estimated class probabilities for the three drugs show that in the responder class (class 2, 63%) 21% of the subjects are on fluoxetine, 29% are on venlafaxine IR, and 50% are on venlafaxine XR. For the nonresponder class that shows an initial improvement and a later worsening (class 3, 19%), 25% are on fluoxetine, 75% are on venlafaxine IR, and 0% are on venlafaxine XR. For the nonresponder class that shows no improvement at any point (class 1, 19%), 58% are on fluoxetine, 13% are on venlafaxine IR, and 29% are on venlafaxine XR. Judged across all three trajectory classes, this suggests that venlafaxine XR has the better outcome, followed by venlafaxine IR, with fluoxetine last. Note, however, that for these data subjects were not randomized to the different medications; therefore, comparisons among medications are confounded by subject differences.[8]

6. The log-likelihood value for the model in Figure 7–5 is –1,084.635, replicated across many random starts, with 28 parameters and a BIC value of 2,278. The entropy value is 0.90.
7. The log-likelihood value for the model in Figure 7–6 is –1,077.433, replicated across many random starts, with 34 parameters and a BIC value of 2,287. The BIC value is worse than for the two-class solution. The entropy value is 0.85.
8. The log-likelihood value for the model of Figure 7–7 is –873.831, replicated across many random starts, with 27 parameters and a BIC value of 1,853. The entropy value is 0.79.

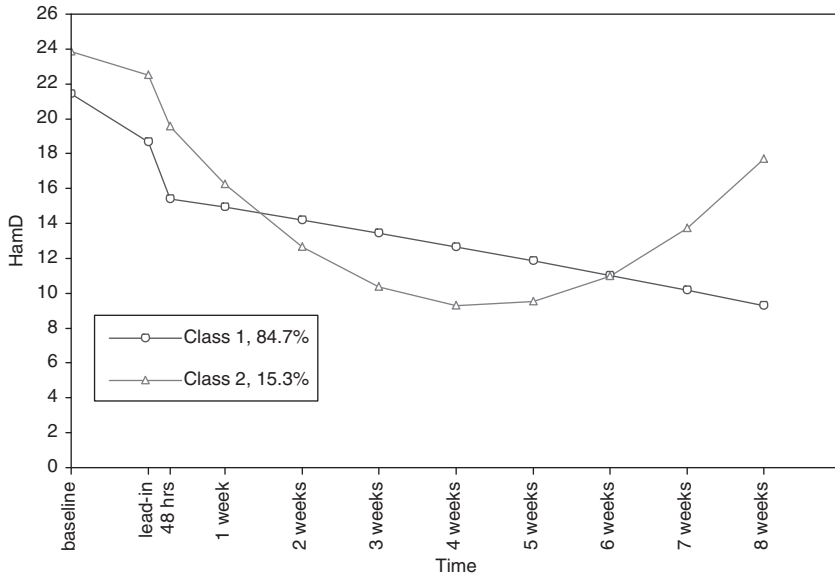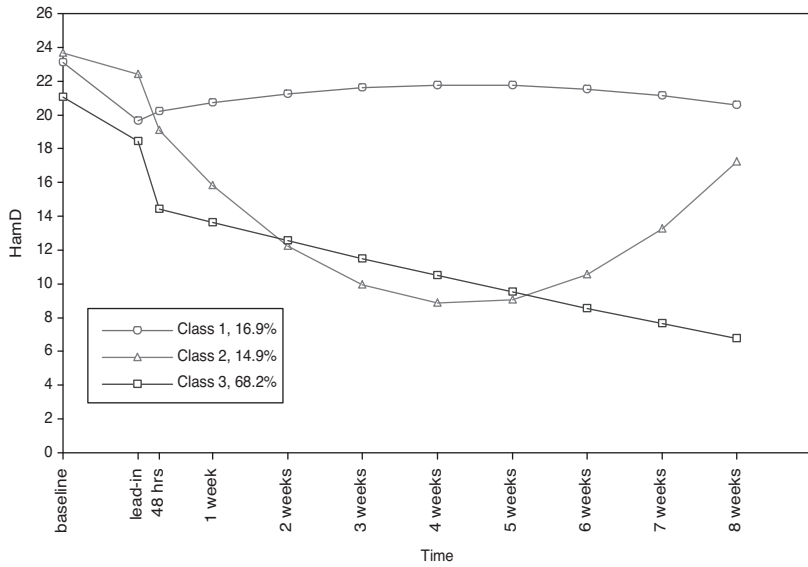*Figure 7.5* Two-class GMM for medication group.

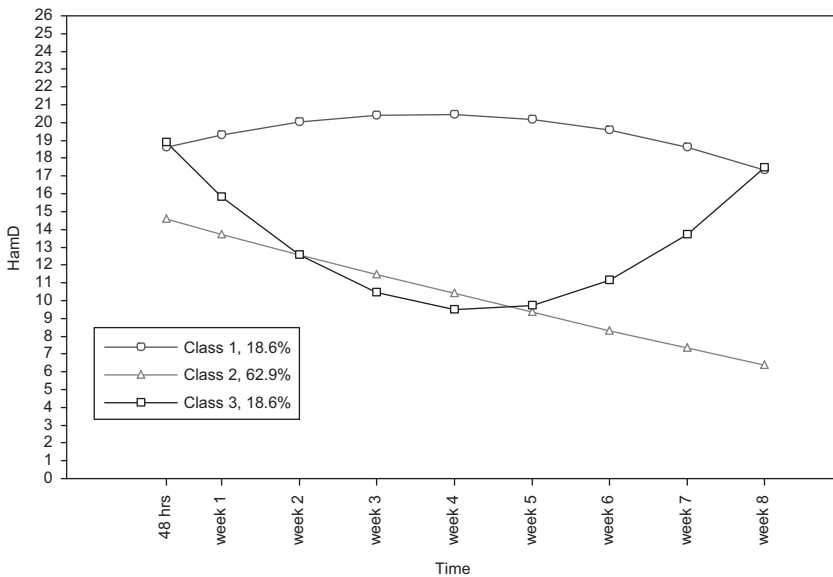*Figure 7.6* Three-class GMM for medication group.

Figure 7.7 Three-class GMM for medication group post randomization.

As a second step, a three-class model was analyzed by a GMM, where not only class membership probability was allowed to vary across the three drugs but also the class-varying random effect means. This analysis showed no significant drug differences in class membership probabilities. As shown in Figure 7–8, the classes are essentially of different nature for the three drugs.[9]

Analysis of Medication Effects, Taking Placebo Response Into Account

The separate analyses of the 45 subjects in the placebo group and the 49 subjects in the medication group provide the basis for the joint analysis of all 94 subjects. Two types of GMMs will be applied. The first is directly in line with the model shown earlier under Growth Mixture Modeling, where medication effects are conceptualized as postrandomization changes in the slope means. The second type uses only the postrandomization time points and class membership is thought of as being influenced by

9. The log-likelihood value for the model of Figure 7–8 is –859.577, replicated in only a few random starts, with 45 parameters and a BIC value of 1, 894. The entropy value is 0.81. It is difficult to choose between the model of Figure 7–7 and the model of Figure 7–8 based on statistical indices. The Figure 7–7 model has the better BIC value, but the improvement in the log-likelihood of the Figure 7–8 model is substantial.
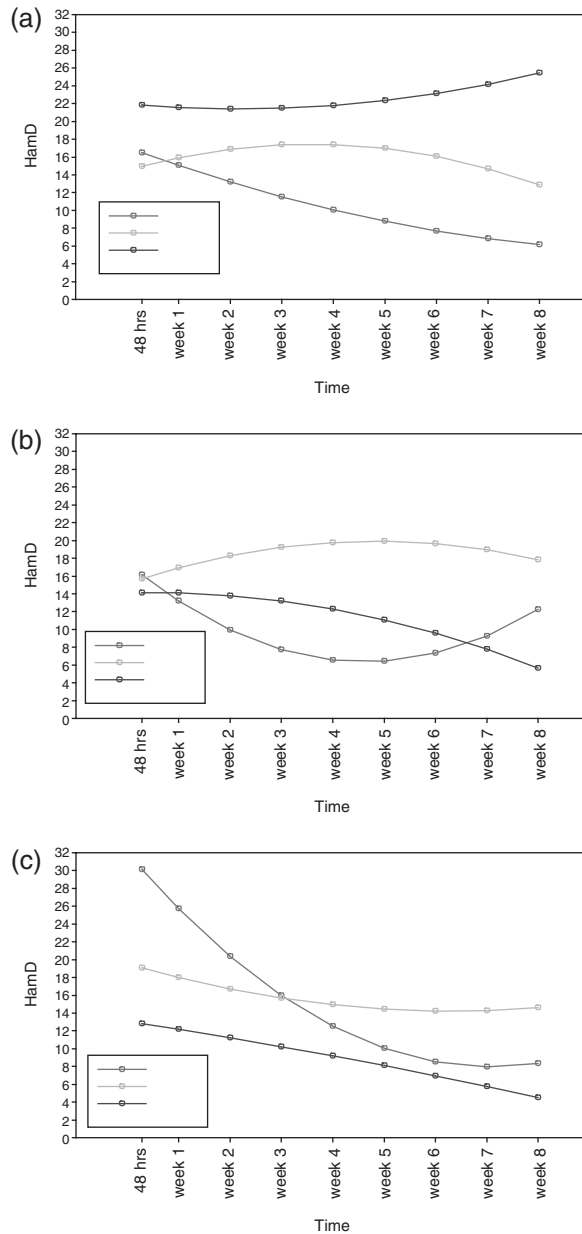
Figure 7.8 Three-class GMM for (a) fluoxetine subjects, (b) venlafaxine IR subjects, and (c) venlafaxine XR subjects.

medication, in line with the Figure 7–7 model. Here, the class probabilities are different for the placebo group and the three medication groups so that medication effect is quantified in terms of differences across groups in class probabilities.

Analysis of All Time Points

For the analysis based on the earlier model (see Growth Mixture Modeling), a three-class GMM will be used, given that three classes were found to be interpretable for both the placebo and the medication groups. Figure 7–9 shows the estimated mean curves for the three-class solution for the placebo group, the fluoxetine group, the venlafaxine IR group, and the venlafaxine XR group. It is interesting to note that for the placebo group the Figure 7–9a mean curves are similar in shape to those of Figure 7–4b, although the responder class (class 3) is now estimated to be 34%. Note that for this model the class percentages are specified to be the same in the medication groups as in the placebo group. The estimated mean curves for the three medication groups shown in Figure 7–9b–d are similar in shape to those of the medication group analysis shown in Figure 7–8a–c. These agreements with the separate-group analyses strengthen the plausibility of the modeling.

This model allows the assessment of medication effects in the presence of a placebo response. A key parameter is the medication-added mean of the intercept random effect centered at week 8. This is the γ_{01k} parameter of equation 5. For a given trajectory class, this indicates how much lower or higher the average score is at week 8 for the medication group in question relative to the placebo group. In this way, the medication effect is specific to classes of individuals who would or would not have responded to placebo.

The γ_{01k} estimates of the Figure 7–9 model are as follows. The fluoxetine effect for the high nonresponder class 1 at week 8 as estimated by the GMM is significantly positive (higher depression score than for the placebo group), 7.4, indicating a failure of this medication for this class of subjects. In the low nonresponder class 2 the fluoxetine effect is small but positive, though insignificant. In the responder class, the fluoxetine effect is significantly negative (lower depression score than for the placebo group), –6.3. The venlafaxine IR effect is insignificant for all three classes. The venlafaxine XR effect is significantly negative, –11.7, for class 1, which after an initial slight worsening turns into a responder class for venlafaxine XR. For the nonresponder class 2 the venlafaxine XR effect is insignificant, while for the responder class it is significantly negative, –7.8. In line with the medication group analysis shown in Figure 7–7, the joint analysis of placebo and medication subjects indicates that venlafaxine XR has the most desirable outcome relative to the placebo group. None of the drugs is significantly effective for the low nonresponder class 2.[10]

10. The log-likelihood value for the model shown in Figure 7–9 is –2,142.423, replicated across a few random starts, with 61 parameters and a BIC value of 4,562. The entropy value is 0.76.
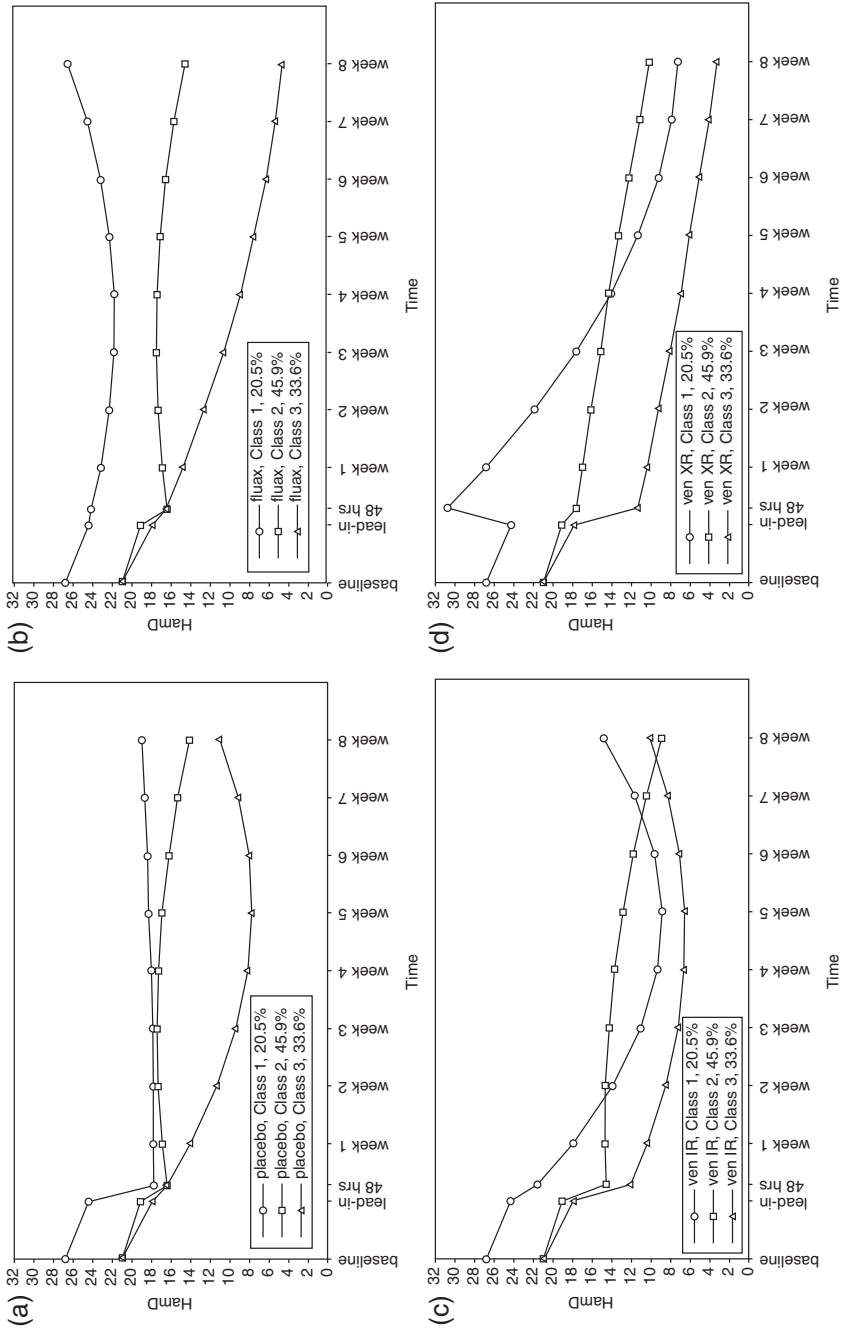
Figure 7.9 Three-class GMM of both groups: (a) Placebo subjects, (b) fluoxetine subjects, (c) venlafaxine IR subjects, and (d) venlafaxine XR subjects.

Analysis of Postrandomization Time Points

As a final analysis, the placebo and medication groups were analyzed together for the postrandomization time points. Figure 7–10 displays the estimated three-class solution, which again shows a responder class, a nonresponder class which initially improves but then worsens (similar to the placebo response class found in the placebo group), and a high nonresponder class.[11] As a first step, it is of interest to compare the joint placebo–medication group analysis of Figure 7–10 to the separate placebo group analysis of Figure 7–4b and the separate medication group analysis of Figure 7–6.

Comparing the joint analysis in Figure 7–10 to that of the placebo group analysis of Figure 7–4b indicates the improved outcome when medication group individuals are added to the analysis. In the placebo group analysis of Figure 7–4b 78% are in the two highest, clearly nonresponding trajectory classes, whereas in the joint analysis of Figure 7–10 only 36% are in the highest, clearly nonresponding class. In this sense, medication seems to have a positive effect in reducing depression. Furthermore, in the placebo analysis, 21% are in the placebo-responding class which ultimately worsens, whereas in the joint analysis 21% are in this type of class and 43% are in a clearly responding class.

Comparing the joint analysis in Figure 7–10 to that of the medication group analysis of Figure 7–6 indicates the worsened outcome when placebo group individuals are added to the analysis. In the medication group analysis of Figure 7–6 only 17% are in the nonresponding class compared to 36% in the joint analysis of Figure 7–10. Figure 7–6 shows 15% in the initially improving but ultimately worsening class compared to 21% in Figure 7–10. Figure 7–6 shows 68% in the responding class compared to 43% in Figure 7–10. All three of these comparisons indicate that medication has a positive effect in reducing depression.

As a second step, it is of interest to study the medication effects for each medication separately. The joint analysis model allows this because the class probabilities differ between the placebo group and each of the three medication groups, as expressed by equation 8. The results are shown in Figure 7–11. For the placebo group, the responder class (class 3) is estimated to be 26%, the initially improving nonresponder class (class 1) to be 22%, and the high nonresponder class (class 2) to be 52%. In comparison, for the fluoxetine group the responder class is estimated to be 48% (better than placebo), the initially improving nonresponder class to be 0% (better than placebo), and the high nonresponder class to be 52% (same as placebo). For the

11. The log-likelihood value for the model shown in Figure 7–10 is –1,744.999, replicated across many random starts, with 29 parameters and a BIC value of 3,621. The entropy value is 0.69.
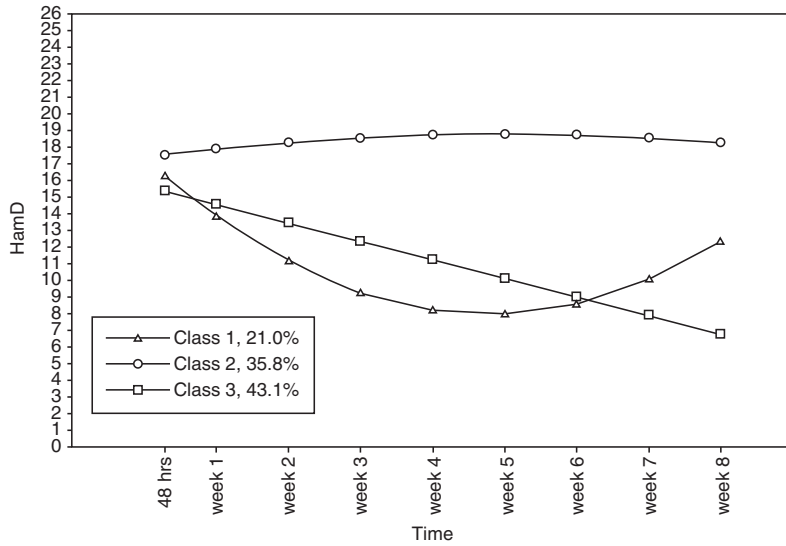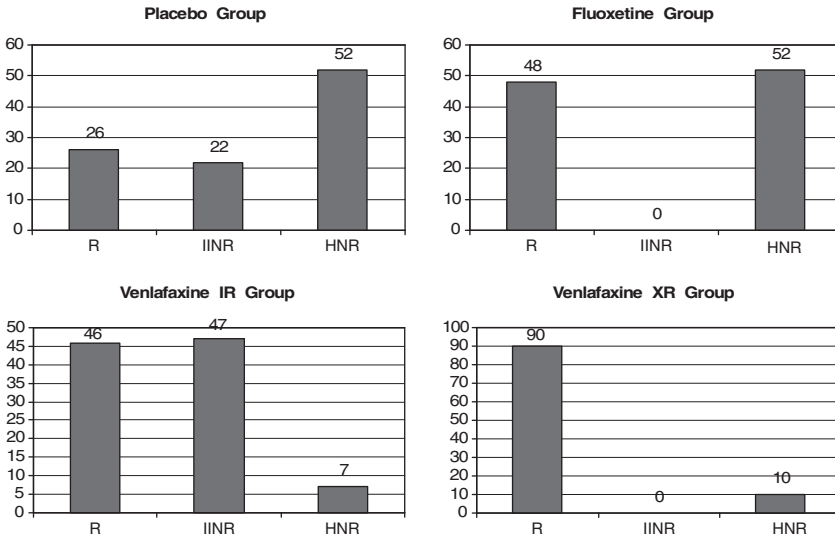
Figure 7.10 Three-class GMM analysis of both groups using post-randomization time points.

R = Responder Class

IINR = Initially Improving Non-Responder Class

HNR = High Non-Responder Class

Figure 7.11 Medication effects in each of 3 trajectory classes.

venlafaxine IR group, the responder class is estimated to be 46% (better than placebo), the initially improving nonresponder class t be 47% (worse than placebo), and the high nonresponder class to be 7% (better than placebo). For the venlafaxine XR group, the responder class is estimated to be 90% (better than placebo), the initially improving nonresponder class to be 0% (better than placebo), and the high nonresponder class to be 10% (better than placebo).

Conclusions

The growth mixture analysis presented here demonstrates that, unlike conventional repeated measures analysis, it is possible to estimate medication effects in the presence of placebo effects. The analysis is flexible in that the medication effect is allowed to differ across trajectory classes. This approach should therefore have wide applicability in clinical trials. It was shown that medication effects could be expressed as causal effects. The analysis also produces a classification of individuals into trajectory classes.

Medication effects were expressed in two alternative ways, as changes in growth slopes and as changes in class probabilities. Related to the latter approach, a possible generalization of the model is to include two latent class variables, one before and one after randomization, and to let the medication influence the postrandomization latent class variable as well as transitions between the two latent class variables. Another generalization is proposed in Muthén and Brown (2009) considering four classes of subjects: (1) subjects who would respond to both placebo and medication, (2) subjects who would respond to placebo but not medication, (3) subjects who would respond to medication but not placebo, and (4) subjects who would respond to neither placebo nor medication. Class 3 is of particular interest from a pharmaceutical point of view.

Prediction of class membership can be incorporated as part of the model but was not explored here. Such analyses suggest interesting opportunities for designs of trials. If at baseline an individual is predicted to belong to a nonresponder class, a different treatment can be chosen.

References

Leuchter, A. F., Cook, I. A., Witte, E. A., Morgan, M., & Abrams, M. (2002). Changes in brain function of depressed subjects during treatment with placebo. *American Journal of Psychiatry, 159*, 122–129.

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park, CA: Sage Publications.

Muthén, B., & Asparouhov, T. (2008). Growth mixture modeling: Analysis with non-Gaussian random effects. In G. Fitzmaurice, M. Davidian, G. Verbeke, & G. Molenberghs (Eds.), *Longitudinal data analysis* (pp. 143–165). Boca Raton, FL: Chapman & Hall/CRC Press.

Muthén, B. & Brown, H. (2009). Estimating drug effects in the presence of placebo response: Causal inference using growth mixture modeling. *Statistics in Medicine, 28,* 3363–3385.

Muthén, B., Brown, C. H., Masyn, K., Jo, B., Khoo, S. T., Yang, C. C., et al. (2002). General growth mixture modeling for randomized preventive interventions. *Biostatistics, 3,* 459–475.

Muthén, B., & Muthén, L. (2000). Integrating person-centered and variable-centered analysis: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research, 24,* 882–891.

Muthén, B., & Muthén, L. (1998–2008). *Mplus user's guide* (5th ed.) Los Angeles: Muthén & Muthén.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics, 55,* 463–469.

Quitkin, F. M., Rabkin, J. G., Ross, D., & Stewart, J. W. (1984). Identification of true drug response to antidepressants. Use of pattern analysis. *Archives of General Psychiatry, 41,* 782–786.