Published in final edited form as:

J Am Stat Assoc. 2013 September 27; 108(503): . doi:10.1080/01621459.2013.797355.

Estimation and inference concerning ordered means in analysis of covariance models with interactions

Jason L. Morrisette* and Michael P. McDermott†

Jason L. Morrisette: jason_morrisette@urmc.rochester.edu; Michael P. McDermott: mikem@bst.rochester.edu

*PhD Student, University of Rochester Medical Center, Rochester, NY 14642

†Professor of Biostatistics and Neurology, University of Rochester Medical Center, Rochester, NY 14642

Abstract

When interactions are identified in analysis of covariance models it becomes important to identify values of the covariates for which there are significant differences or, more generally, significant contrasts among the group mean responses. Inferential procedures that incorporate a priori order restrictions among the group mean responses would be expected to be superior to those that ignore this information. In this paper we focus on analysis of covariance models with pre-specified order restrictions on the mean response across the levels of a grouping variable when the grouping variable may interact with model covariates. In order for the restrictions to hold in the presence of interactions, it is necessary to impose the requirement that the restrictions hold over all levels of interacting categorical covariates and across pre-specified ranges of interacting continuous covariates. The parameter estimation procedure involves solving a quadratic programming minimization problem with a carefully specified constraint matrix. Simultaneous confidence intervals for treatment group contrasts and tests for equality of the ordered group mean responses are determined by exploiting previously unconnected literature. The proposed methods are motivated by a clinical trial of the dopamine agonist pramipexole for the treatment of early-stage Parkinson's disease.

Keywords

Order constrained inference; quadratic programming; likelihood ratio test; Johnson-Neyman procedure; Scheffé method

1 INTRODUCTION

Analysis of covariance is a commonly used technique for incorporating covariate information when comparing the means of several normally distributed populations. In the setting of randomized controlled trials, the use of covariate information can substantially increase the precision with which treatment effects are estimated. In other settings, relevant covariate information can be used for bias reduction. We consider the problem of inference on the means of k groups in the setting of an analysis of covariance model that may include interactions between the grouping variable of interest and possibly multiple covariates. In addition, the mean responses are known a priori to satisfy an order restriction across all levels of categorical covariates and across pre-specified ranges of continuous covariates. Betcher and Peddada (2009) previously considered inference on analysis of covariance models with order restrictions on the group mean responses; however, they did not consider interactions involving the grouping variable.

The problem to be addressed in this paper is well-motivated by data from a multi-center, randomized, double-blind, placebo-controlled trial of pramipexole, a dopamine agonist, for the treatment of early-stage Parkinson's disease (Parkinson Study Group, 1997). In this trial, 264 participants from 20 centers were randomly assigned to receive pramipexole at a dosage of either 1.5 mg/day ($n = 54$), 3.0 mg/day ($n = 50$), 4.5 mg/day ($n = 54$), or 6.0 mg/day ($n = 55$), or matching placebo ($n = 51$). Participants were followed for 10 weeks on study medication. The primary outcome variable for efficacy was the ten-week change from baseline in the total score on the Unified Parkinson's Disease Rating Scale (UPDRS), an instrument commonly used to assess the severity of the clinical features of Parkinson's disease (e.g., tremor, rigidity, bradykinesia, postural instability/gait difficulty) (Lang and Fahn, 1989). The distribution of UPDRS scores at baseline ranged from 5 – 78, with a mean of 29.3 and a standard deviation of 14.2; higher scores indicate higher severity of Parkinson's disease. The goal of the trial was to assess the safety, tolerability, and efficacy of pramipexole at multiple dosages.

The primary analysis used an analysis of covariance model that treated dosage as a continuous variable (with the placebo group assigned a dosage of zero) and included center and baseline UPDRS score as covariates; no treatment by covariate interactions were included in the model (Parkinson Study Group, 1997). A test for significance of the coefficient for dosage yielded a p -value < 0.001 . A secondary analysis used a similar model but treated dosage as a categorical variable. The estimated treatment effects (relative to placebo) from this model were significant ($p < 0.005$) and remarkably consistent across the dosages, ranging from -5.1 to -5.9 points (with negative values indicating improvement). An additional finding in an exploratory analysis was an interaction between treatment group and baseline UPDRS score in this latter model ($p = 0.002$). As illustrated in Figure 1, the nature of the interaction suggests that the effect of pramipexole may be greater in patients with more severe Parkinson's disease features (higher baseline UPDRS scores).

The fitted model illustrated in Figure 1 raises interesting questions regarding the dose-response relationship. It seems clear from the results of the secondary analysis given above that this relationship is not linear, a conclusion verified by a lack of fit test ($p = 0.019$). Instead of assuming a particular functional form for this relationship, it may be reasonable to assume that the relationship satisfies an order restriction such as simple order, where the mean responses are non-increasing with dosage, or so-called simple tree order, where the mean responses for each active treatment group are at least as small as that for the placebo group. Moreover, it may be reasonable to assume that these order restrictions hold for each value of the baseline UPDRS score in a specified range. If the lines estimated in Figure 1 are in fact not parallel, it is impossible in general to impose an order restriction across the entire range of baseline UPDRS scores (the fitted lines will intersect eventually), but one may assume the order restriction to hold across a relevant subset of this range. The fitted model in Figure 1 is consistent with the 1.5 mg/day dosage being better than the 3.0 mg/day dosage for baseline UPDRS scores above 30, and with the 4.5 mg/day dosage being worse than placebo for baseline UPDRS scores lower than 10, neither of which may be plausible. To the extent that the order restrictions are valid, an analysis of covariance model that incorporates these restrictions may provide an improved description of the dose-response relationship.

Another important problem is introduced by the finding of an interaction between baseline UPDRS score and treatment group in the exploratory analysis. This interaction indicates that the differences among the treatment group means depend on the baseline UPDRS score. It is then of interest to identify the values of baseline UPDRS score, if any, for which there is a significant difference in the mean response across the treatment groups. In the case of comparing only two treatment groups, this procedure is attributed to Johnson and Neyman (1936). The method of Scheffé (1953) can be used to construct simultaneous confidence

intervals for any contrast among the treatment group mean responses. If a priori order restrictions are included when fitting the model it would be useful to also incorporate this information when identifying the covariate regions for which the group differences (or contrasts) are significant.

In Section 2, we present preliminaries including a reparameterized version of the analysis of covariance model with interactions followed by a brief overview of order restrictions. A procedure to obtain maximum likelihood estimates of the model parameters subject to order restrictions is described in Section 3. Methods for more formal inferences (simultaneous interval estimation and hypothesis testing) regarding the model parameters are described in Section 4. These include a procedure for identifying regions of the covariates for which the ordered group means significantly differ and tests for significance of the grouping variable subject to the order restriction. We return to the pramipexole clinical trial example in Section 5 and illustrate the utility of the proposed methods using these data. Further discussion and conclusions are summarized in Section 6.

2 PRELIMINARIES

For convenience, we will express the analysis of covariance model as

$$\mathbf{Y} = \mathbf{X}_1\mu + \mathbf{X}_2\beta + \mathbf{X}_3\gamma + \epsilon, \quad (1)$$

where $\mathbf{Y}$ is an $N \times 1$ vector of continuous response variables and $\mathbf{X}_1 = (\mathbf{X}_{11}, \dots, \mathbf{X}_{1k})$ is an $N \times k$ matrix of indicator variables designating the k levels of the grouping variable of interest. The columns of the $N \times r$ matrix $\mathbf{X}_2$ represent the collection of continuous and categorical covariates that do not interact with the grouping variable of interest. Each of the categorical covariates is represented here by a set of indicator variables numbering one fewer than the number of levels of the covariate. The collection of continuous and categorical covariates that do interact with the grouping variable of interest can be similarly represented by an $N \times t$ matrix $\mathbf{Z}$. The matrix $\mathbf{X}_3$ is formed by $(\mathbf{X}_{11} \cdot \mathbf{Z}, \mathbf{X}_{12} \cdot \mathbf{Z}, \dots, \mathbf{X}_{1k} \cdot \mathbf{Z})$, where the j^{th} column of $\mathbf{X}_{1i} \cdot \mathbf{Z}$ is the Hadamard product of the indicator variable $\mathbf{X}_{1i}$ with the j^{th} column of $\mathbf{Z}$ (element-wise multiplication), $i = 1, \dots, k$. Also, $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$ and μ , β , and γ are the corresponding vectors of parameters. In this parameterization, if a single continuous covariate interacted with the grouping variable of interest whereby the relationship between the covariate and the response differed for each group, the coefficients μ and γ would represent the different intercepts and slopes for the k groups, respectively. The model from equation (1) can be represented more compactly as $\mathbf{Y} = \mathbf{X}\theta$, where $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$ and $\theta = (\mu, \beta, \gamma)$. In the model without interactions – corresponding to $t = 0$ – the linear relationships between the response variable and the covariates are the same across all levels of the grouping variable. When the parameter space for θ is unrestricted, the solution to

$$\arg \min_{\theta \in \Theta} \|\mathbf{Y} - \mathbf{X}\theta\|^2$$

has a known simple form.

Estimation of the model parameters will be considered for any order restriction on the mean responses defined by a finite number of homogeneous linear inequalities, which describes a polyhedral cone. For simplicity, consider the case of the model from equation (1) that includes only the grouping variable of interest. The polyhedral cone is defined as $C(\mathbf{A}) = \{\mu \in \mathbb{R}^k | \mathbf{A}\mu \geq \mathbf{0}\}$ for some $k \times m$ matrix $\mathbf{A}$. The matrix $\mathbf{A}$ should have no redundant columns, where a redundant column is any column which, once removed, yields the same polyhedral cone. Two commonly applied order restrictions are simple order and simple tree order,

which are the spaces $C(\mathbf{A}) = \{\boldsymbol{\mu} | \mu_1 = \mu_2 = \dots = \mu_k\} = \{\boldsymbol{\mu} | \mathbf{A} \boldsymbol{\mu} = \mathbf{0}\}$ and $C(\mathbf{A}) = \{\boldsymbol{\mu} | \mu_1 = \mu_i, i = 2, \dots, k\} = \{\boldsymbol{\mu} | \mathbf{A} \boldsymbol{\mu} = \mathbf{0}\}$, where the $k \times (k-1)$ matrices $\mathbf{A}$ associated with these cones are

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ -1 & 1 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & -1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 1 & \dots & 1 \\ -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{pmatrix},$$

respectively. The simple tree ordering is particularly applicable in placebo-controlled studies in which it is assumed treated populations will have at least as large a mean response as the placebo group, but no ordering assumptions are made among the mean responses in the active treatment groups.

3 MAXIMUM LIKELIHOOD ESTIMATION OF MODEL PARAMETERS

In the general analysis of covariance model in Section 2, recall that $\mathbf{X}_3 = (\mathbf{X}_{11} \cdot \mathbf{Z}, \mathbf{X}_{12} \cdot \mathbf{Z}, \dots, \mathbf{X}_{1k} \cdot \mathbf{Z})$ where $\mathbf{Z}$ is the $N \times t$ matrix containing the covariates that interact with the

grouping variable of interest. We can similarly decompose $\boldsymbol{\gamma} = (\gamma'_1, \dots, \gamma'_k)'$, where γ_j is the $t \times 1$ vector of parameters corresponding to the terms in $\mathbf{X}_{1j} \cdot \mathbf{Z}, j = 1, \dots, k$. Now let

$$\alpha_i(\mathbf{x}) = \mathbf{x}' \boldsymbol{\theta} = \mathbf{x}'_1 \boldsymbol{\mu} + \mathbf{x}'_2 \boldsymbol{\beta} + \mathbf{x}'_3 \boldsymbol{\gamma} = \mu_i + \mathbf{x}'_2 \boldsymbol{\beta} + \mathbf{z}' \boldsymbol{\gamma}_i$$

denote the expected response for the i^{th} level of the grouping variable at the values

$\mathbf{x} = (\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3)'$ of the grouping variable and the covariates, where $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, and $\mathbf{z}$ are $k \times 1, r \times 1, kt \times 1$, and $t \times 1$ vectors, respectively. The vector $\mathbf{x}$ is a single realization of each of the variables corresponding to the columns of the $N \times (k + r + kt)$ matrix $\mathbf{X}$. For an observation from group i , the vector $\mathbf{x}_1$ will have a 1 in the i^{th} position and 0 elsewhere, and the vector $\mathbf{x}_3$ will consist of the $t \times 1$ vector $\mathbf{z}$ in the $[(i-1)t+1]^{th}$ through $[it]^{th}$ positions and 0 elsewhere, $i = 1, \dots, k$.

For the moment, we will assume that there are only continuous covariates in the model. Suppose a priori we impose a set of linear inequality constraints of the form $\mathbf{A}(\mathbf{x}) \leq \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^{k+r+kt}$ and for some $k \times m$ matrix $\mathbf{A}$, on the vector of means $\boldsymbol{\alpha}(\mathbf{x}) = (\alpha_1(\mathbf{x}), \dots, \alpha_k(\mathbf{x}))$. Rewriting this constraint in matrix form, we have

$$\mathbf{A}' \boldsymbol{\mu} + \mathbf{A}' \mathbf{1}_k \mathbf{x}'_2 \boldsymbol{\beta} + \mathbf{A}' \tilde{\mathbf{X}} \boldsymbol{\gamma} \leq \mathbf{0}, \forall \mathbf{x} \in \mathbb{R}^{k+r+kt}, \quad (2)$$

where $\mathbf{1}_k$ is a vector of 1's and the $k \times kt$ matrix $\tilde{\mathbf{X}}$ is defined as

$$\tilde{\mathbf{X}} = \begin{pmatrix} \mathbf{z} & \mathbf{0}_t & \dots & \mathbf{0}_t \\ \mathbf{0}_t & \mathbf{z} & \dots & \mathbf{0}_t \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_t & \mathbf{0}_t & \dots & \mathbf{z} \end{pmatrix},$$

with $\mathbf{0}_t$ denoting a $t \times 1$ vector of 0's. In general, the only way these constraints can hold for all $\mathbf{x} \in \mathbb{R}^{k+r+kt}$ is if $\mu = \mathbf{0}$ and $\beta = \mathbf{0}$, which reduces the model to a standard analysis of variance model.

In practice, one is typically interested in estimating the mean responses in the analysis of covariance model over pre-specified intervals of the continuous covariates rather than the entire space. For example, inference beyond the approximate range of the observed covariate values is often not desired. Alternatively, specific study populations may be of interest. These intervals are at the investigator's discretion to choose. Let $C_j = [C_{j1}, C_{j2}]$ denote the pre-specified interval of interest for the j^{th} continuous covariate, $j = 1, \dots, r+t$. The combination of these $r+t$ intervals creates a $(r+t)$ -dimensional hyper-rectangle, denoted by $\mathcal{H}$, so that estimation of the mean responses in the analysis of covariance model is equivalent to estimating k different hyper-planes within this hyper-rectangle. Since these hyper-planes represent only linear relationships, the constraint $\mathbf{A}'(\mathbf{x}) \leq \mathbf{0}$ for all $\mathbf{x} \in \mathcal{H}$ is satisfied if and only if $\mathbf{A}'\alpha(\mathcal{V}) \leq \mathbf{0}$ for all vertices $\mathcal{V}$ of the hyper-rectangle $\mathcal{H}$. Let

$\mathcal{V}_I = (\mathcal{V}'_{I1}, \mathcal{V}'_{I2})'$, $I = 1, \dots, 2^{r+t}$, denote the $(r+t) \times 1$ vectors representing the vertices of $\mathcal{H}$, where $\mathcal{V}_I$ is decomposed into the $r \times 1$ vector $\mathcal{V}'_{I1}$ corresponding to the non-interacting covariates in $\mathbf{X}_2$ and the $t \times 1$ vector $\mathcal{V}'_{I2}$ corresponding to the interacting covariates in $\mathbf{Z}$. The set of constraints in equation (2) above can then be written as

$\mathbf{A}'\mu + \mathbf{A}'\mathbf{1}_k\mathcal{V}'_{I1}\beta + \mathbf{A}'\tilde{\mathcal{V}}_{I2}\gamma \leq \mathbf{0}$, $I = 1, \dots, 2^{r+t}$, or simply $\tilde{\mathbf{A}}'\theta \leq \mathbf{0}$, where

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A}' & \mathbf{A}'\mathbf{1}_k\mathcal{V}'_{11} & \mathbf{A}'\tilde{\mathcal{V}}_{12} \\ \vdots & \vdots & \vdots \\ \mathbf{A}' & \mathbf{A}'\mathbf{1}_k\mathcal{V}'_{2^{r+t}1} & \mathbf{A}'\tilde{\mathcal{V}}_{2^{r+t}2} \end{pmatrix} \text{ and } \tilde{\mathcal{V}}_{I2} = \begin{pmatrix} \mathcal{V}'_{I2} & \mathbf{0}_t & \cdots & \mathbf{0}_t \\ \mathbf{0}_t & \mathcal{V}'_{I2} & \cdots & \mathbf{0}_t \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_t & \mathbf{0}_t & \cdots & \mathcal{V}'_{I2} \end{pmatrix}. \quad (3)$$

Therefore the maximum likelihood estimator (MLE) for the parameter vector θ under the order restriction is the solution to the quadratic programming problem

$$\text{minimize } \|\mathbf{Y} - \mathbf{X}\theta\|^2 \text{ subject to the constraint } \tilde{\mathbf{A}}'\theta \leq \mathbf{0}. \quad (4)$$

There are various algorithms available to solve this quadratic programming problem; see for example Silvapulle and Sen (2005) and Meyer (2010).

When the original constraint matrix $\mathbf{A}$ satisfies $\sum_{i=1}^k A_{ij} = 0$, $j = 1, \dots, m$, the final constraint matrix $\tilde{\mathbf{A}}$ can be simplified further because $\mathbf{A}'\mathbf{1}_k = \mathbf{0}_k$ and so

$\mathbf{A}'\mu + \mathbf{A}'\mathbf{1}_k\mathcal{V}'_{I1}\beta + \mathbf{A}'\tilde{\mathcal{V}}_{I2}\gamma \leq \mathbf{0}$ reduces to $\mathbf{A}'\tilde{\mathcal{V}}_{I2}\gamma \leq \mathbf{0}$. The constraint matrix then becomes

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A}' & \mathbf{0}_{m \times r} & \mathbf{A}'\tilde{\mathcal{V}}_{12} \\ \vdots & \vdots & \vdots \\ \mathbf{A}' & \mathbf{0}_{m \times r} & \mathbf{A}'\tilde{\mathcal{V}}_{2^{r+t}2} \end{pmatrix}$$

with $\tilde{\mathcal{V}}_{I2}$ defined as above. This means that if $\mathbf{A}$ satisfies this property, then the constraint does not involve any of the covariates that do not interact with the grouping variable. As a result, although there are 2^{r+t} vertices, only 2^t of the vertices produce non-redundant rows of $\tilde{\mathbf{A}}$. Therefore, when these redundant rows are removed, $\tilde{\mathbf{A}}$ would be of order $m2^t \times (k + r + kt)$.

Two special cases of this model, corresponding to $t = 0$ and $r = 0$, are also note-worthy. The parameter constraint matrix $\tilde{\mathbf{A}}$ for these cases takes the forms

$$\begin{pmatrix} \mathbf{A}' & \mathbf{A}' \tilde{\mathcal{V}}_{12} \\ \vdots & \vdots \\ \mathbf{A}' & \mathbf{A}' \tilde{\mathcal{V}}_{2^{t-2}} \end{pmatrix} \text{ and } \begin{pmatrix} \mathbf{A}' & \mathbf{A}' \mathbf{1}_k \mathcal{V}'_{11} \\ \vdots & \vdots \\ \mathbf{A}' & \mathbf{A}' \mathbf{1}_k \mathcal{V}'_{2^{r-1}} \end{pmatrix},$$

respectively, with $\mathcal{V}_{12}$ again defined as above. Note that even though the latter model contains no interactions, additional constraints beyond $\mathbf{A} \boldsymbol{\mu} = \mathbf{0}$ are needed to ensure that the order restriction on the mean responses holds for all values of the covariates.

Extending the above methods to include categorical covariates with a fixed and finite number of categories is straightforward. Since there are only a finite number of categories, the constraints can be repeated for each level of the categorical covariate, ensuring that the constraints are valid over all levels of the covariate.

4 INFERENCE FOR THE MODEL PARAMETERS

In the presence of interactions, it is often of interest to identify the regions of the covariates over which the mean response is significantly different between any two levels of the grouping variable. One possible use of such regions is as decision making tools in clinical practice. When comparing a treatment to a placebo the region could be used to determine the subset of patients for whom it is worthwhile to prescribe the treatment based on measurable baseline characteristics. One may also be interested in more general contrasts among the group means, such as a comparison of the average response among multiple dosages of a drug with that for placebo. Simultaneous confidence regions for these contrasts can be created using the method of Scheffé (1953), and regions of significance could be determined using the approach of Johnson and Neyman (1936). In our case, however, the parameters of the model lie in the polyhedral cone defined by the order restriction. Let $\hat{\boldsymbol{\theta}} = (\mathbf{X} \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y}$ be the unrestricted MLE with covariance matrix $\mathbf{D} = (\mathbf{X} \mathbf{X})^{-1}$. An unbiased estimator for σ^2

is $S^2 = \frac{1}{v} (\mathbf{Y} - \mathbf{X} \hat{\boldsymbol{\theta}})' (\mathbf{Y} - \mathbf{X} \hat{\boldsymbol{\theta}})$, where $v = N - k - r - kt$. We are interested in contrasts of the form $\mathbf{a}'(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{H}$, where $\mathbf{a} \in \mathbb{R}^k$ is a vector of contrast coefficients satisfying $\mathbf{A} \mathbf{a} = \mathbf{0}$. Furthermore, we restrict attention to the case where $\mathbf{A}$ satisfies the property that

$\sum_{i=1}^k A_{ij} = 0, j = 1, \dots, m$, since this will be of interest in most practical applications.

Following similar logic to that in Section 3, it is more convenient to write the contrasts $\mathbf{a}'(\mathbf{x})$

in terms of the model parameters rather than the expected group means $\mathbf{a}'(\mathbf{x})$. We can rewrite $\mathbf{a}'(\mathbf{x})$ as $\mathbf{c}'(\mathbf{x})$ for some $\mathbf{c}(\mathbf{x})$ satisfying $\tilde{\mathbf{A}} \mathbf{c}(\mathbf{x}) = \mathbf{0}$. As an example, the comparison between the j^{th} and $(j+1)^{\text{th}}$ levels of the grouping variable would involve the contrast vector $\mathbf{a} = (0'_{j-1}, 1, -1, 0'_{k-j-1})'$, which can be written as a contrast vector on using

$$\mathbf{c}(\mathbf{x}) = (\underbrace{0, \dots, 0}_{j-1}, 1, -1, \underbrace{0, \dots, 0}_{k-j-1}, \underbrace{0, \dots, 0}_r, \underbrace{0, \dots, 0}_{t(j-1)}, \mathbf{z}', -\mathbf{z}', \underbrace{0, \dots, 0}_{t(k-j-1)})'.$$

From the definition of $\tilde{\mathbf{A}}$ in equation (3), if $\mathbf{a}$ satisfies $\mathbf{A} \mathbf{a} = \mathbf{0}$, then it follows that $\mathbf{c}(\mathbf{x})$ will satisfy $\tilde{\mathbf{A}} \mathbf{c}(\mathbf{x}) = \mathbf{0}$.

Shapiro (2003) extended the method of Scheffé (1953) for constructing simultaneous confidence intervals for all linear combinations of the model parameters, with the coefficients restricted to lie in a polyhedral cone. When we are interested in contrasts among the mean responses across the levels of the grouping variable where the contrast vector $\mathbf{c}(\mathbf{x})$ satisfies $\tilde{\mathbf{A}}\mathbf{c}(\mathbf{x}) = \mathbf{0}$, simultaneously all confidence intervals of the form

$$\left(-\infty, \mathbf{c}'(\mathbf{x})\hat{\theta} + w_{\alpha}^{1/2} \sqrt{S^2(\mathbf{c}'(\mathbf{x}) (\mathbf{X}'\mathbf{X})^{-1} \mathbf{c}(\mathbf{x}))}^{1/2} \right]$$

cover the true value of $\mathbf{c}(\mathbf{x})$ with probability $1 - \alpha$. A slight modification of Shapiro (2003) shows that the critical value w_{α} is chosen to satisfy

$$\alpha = \sum_{i=0}^h p_i P\{(i+d)F_{i+d,\nu} \geq w_{\alpha}\}, \quad (5)$$

where F_{df_1, df_2} is an F random variable with numerator degrees of freedom df_1 and denominator degrees of freedom df_2 , $h = 2^t m$, and $d = 2^t(k-1-m)$ (see Appendix A for more details on the modification). The level probability p_i is the probability that the orthogonal projection of a random variable $\mathbf{L}_1 \sim \mathcal{N}_h(\mathbf{0}, \tilde{\mathbf{A}}\tilde{\mathbf{D}}\tilde{\mathbf{A}})$ onto the space $\{\mathcal{L} \in \mathbb{R}^h | \mathcal{L} \in \mathbf{0}\}$ falls in a linear subspace of dimension i , $i = 0, 1, \dots, h$ (see Robertson et al. (1988) and Silvapulle and Sen (2005) for more details). In Appendix B it is shown that this critical value, w_{α} , will always be smaller than that of Scheffé (1953) when $C(\tilde{\mathbf{A}}) \subseteq \mathbb{R}^{k+r+kt}$ and thus the confidence intervals incorporating the prior ordering information will be narrower.

In most cases the level probabilities are difficult to calculate so one useful solution to this problem is to simulate data from $\mathcal{N}_h(\mathbf{0}, \tilde{\mathbf{A}}\tilde{\mathbf{D}}\tilde{\mathbf{A}})$, calculate the orthogonal projections, and use the empirical probabilities as estimates for the true p_i . When simulation is necessary, the amount of time needed to solve the quadratic programming problem in equation (4) is an important consideration. Meyer (2010) provides a computationally efficient algorithm for finding the restricted MLE $\hat{\theta}^*$.

In an unrestricted parameter space, the usual F -test for equality of means will reject the null hypothesis if and only if at least one of the infinite number of possible contrasts is statistically significant, according to Scheffé's (1953) method for multiple comparisons. In the context of order constrained means, Mukerjee, Robertson, and Wright (1987, p. 903) note that "Dykstra, in a personal communication, showed that the [likelihood ratio test] can be thought of as being based on the maximum contrast with an infinite set of directions in H_1 ," where H_1 corresponds to the restricted parameter space. The method of Shapiro (2003) for constructing simultaneous confidence intervals has a similar dual relationship with the F test statistic for testing the significance of the grouping variable, i.e., the test of $H_0: \mu_1(\mathbf{x}) = \dots = \mu_k(\mathbf{x})$, $\mathbf{x} \in \mathcal{H}$, against the restricted alternative $H_1 \setminus H_0$, where $H_1: \mathbf{A}(\mathbf{x}) \neq \mathbf{0}$, $\mathbf{x} \in \mathcal{H}$. The test statistic takes the form $F = \|\mathbf{X}^* - \mathbf{X}_0\|^2 / S^2$ where $\mathbf{X}_0$ is the MLE under H_0 and $\mathbf{X}^*$ is the MLE under H_1 (Silvapulle and Sen, 2005, Sec 3.9). These estimators are the orthogonal projections of the unrestricted MLE onto the null and alternative spaces, respectively. Under the null hypothesis this test statistic has distribution function

$$P\{\bar{F} \leq c\} = \sum_{i=0}^h p_i P\{(i+d)F_{i+d,\nu} \leq c\}, \quad (6)$$

which takes the same form as in equation (5) above. The F statistic is closely related to the likelihood ratio test statistic which takes the form $\bar{E}_{01}^2 = \|\mathbf{X}\hat{\theta}^* - \mathbf{X}\hat{\theta}_0\|^2 / S_0^2$, where $S_0^2 = \|\mathbf{Y} - \mathbf{X}\hat{\theta}_0\|^2$. Under the null hypothesis this test statistic has distribution function

$$P\{\bar{E}_{01}^2 \leq c\} = \sum_{i=0}^h p_i P\left\{B_{\frac{i+d}{2}, \frac{\nu+\delta-i}{2}} \leq c\right\}, \quad (7)$$

where $B_{\cdot}$ is a beta random variable, $\nu = h$, and d, h , and p_i are defined as above (Robertson, Wright, and Dykstra, 1988). In large samples, these two test statistics are approximately equivalent (Silvapulle and Sen, 2005).

Under this framework, more general hypotheses may also be tested including, but not limited to, the test of $H_0: \mathbf{R} = \mathbf{0}$ against $H_1 \setminus H_0$ with $H_1: \tilde{\mathbf{A}} = \mathbf{0}$, where $\tilde{\mathbf{A}}$ is a sub-matrix of the full rank matrix $\mathbf{R}$, and the test of the same H_1 against $H_2 \setminus H_1$ with $H_2: \mathbb{R}^{k+r+kt}$. An example of the first type of test is when one is interested in testing the null hypothesis that all group means are the same (i.e., $\mu_1(\mathbf{x}) = \dots = \mu_k(\mathbf{x})$) against an alternative where only some of the group means are restricted; the second type of test can be used to test the null hypothesis that the group means follow a specified ordering against the unrestricted alternative. The test statistics and their corresponding null distributions take similar forms as in equations (6) and (7) above; however, $h = \text{rank}(\tilde{\mathbf{A}})$ and $d = \text{rank}(\mathbf{R})$ for the first type of test and $h = k + r + kt$, $d = 0$, and $\nu = 0$ for the second type of test (see Silvapulle and Sen (2005) for more details).

5 PRAMIPEXOLE EXAMPLE

In order to demonstrate the proposed methods, we now return to the example introduced in Section 1 regarding the clinical trial of pramipexole. Following the notation used above, there are $r = 19$ non-interacting covariates (indicator variables for the 20 study centers), $t = 1$ interacting covariate (baseline UPDRS score), and $k = 5$ treatment groups. Under the above parameterization, $\mu_1, \dots, \mu_5, \mu_{1,1}, \dots, \mu_{1,19}$, and $\mu_{1,1}, \dots, \mu_{1,5}$ are the parameters corresponding to the treatment group indicators, study center indicators, and baseline-UPDRS-score $\times$ treatment-group product terms, respectively. Since baseline UPDRS score is an interacting continuous covariate it is necessary to choose (in advance) an interval over which the order restriction must hold. In patients with early-stage Parkinson's disease – the study population of this trial – UPDRS scores higher than 60 are quite rare. Furthermore, some symptoms would need to have manifested before intervention is considered, which would indicate a minimum larger than 0. Since the smallest observed baseline UPDRS score was 5, one may, prior to observing the follow-up data, choose the interval $[5, 60]$ as the interval of interest. In our example, this interval eliminates some outliers (the highest observed score was 78) but it still captures 95.8% of the observed baseline UPDRS scores. The simple tree order restriction for this data set would be represented as $\tilde{\mathbf{A}} = \mathbf{0}$, where

$$\tilde{\mathbf{A}} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & \mathbf{0}'_{19} & -5 & 5 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & \mathbf{0}'_{19} & -5 & 0 & 5 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & \mathbf{0}'_{19} & -5 & 0 & 0 & 5 & 0 \\ -1 & 0 & 0 & 0 & 1 & \mathbf{0}'_{19} & -5 & 0 & 0 & 0 & 5 \\ -1 & 1 & 0 & 0 & 0 & \mathbf{0}'_{19} & -60 & 60 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & \mathbf{0}'_{19} & -60 & 0 & 60 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & \mathbf{0}'_{19} & -60 & 0 & 0 & 60 & 0 \\ -1 & 0 & 0 & 0 & 1 & \mathbf{0}'_{19} & -60 & 0 & 0 & 0 & 60 \end{pmatrix}$$

and $\mathbf{0}_{19}$ is a 19×1 vector of zeros. The matrix $\tilde{\mathbf{A}}$ corresponding to the simple order restriction would be

$$\tilde{\mathbf{A}} = \left(\begin{array}{ccccc|ccccc} -1 & 1 & 0 & 0 & 0 & \mathbf{0}'_{19} & -5 & 5 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & \mathbf{0}'_{19} & 0 & -5 & 5 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & \mathbf{0}'_{19} & 0 & 0 & -5 & 5 & 0 \\ 0 & 0 & 0 & -1 & 1 & \mathbf{0}'_{19} & 0 & 0 & 0 & -5 & 5 \\ -1 & 1 & 0 & 0 & 0 & \mathbf{0}'_{19} & -60 & 60 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & \mathbf{0}'_{19} & 0 & -60 & 60 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & \mathbf{0}'_{19} & 0 & 0 & -60 & 60 & 0 \\ 0 & 0 & 0 & -1 & 1 & \mathbf{0}'_{19} & 0 & 0 & 0 & -60 & 60 \end{array} \right).$$

Figure 2 shows the model fits using these two order restrictions. It can easily be seen that the restrictions hold for all values of baseline UPDRS score in the range [5, 60]. To test the null hypothesis of no treatment effect, i.e., $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5$, against the simple tree order alternative, the likelihood ratio test statistic is $\bar{E}_{01}^2 = 0.13942$, $d = 0$, $h = 8$, and $n = 264$. Using Monte Carlo simulation to estimate the p_i , the p-value was calculated as $p_{sto} = 0.0000151$. The usual F-test for H_0 against an unrestricted alternative yields $F_{8,235} = 4.9824$ and a corresponding p-value of $p_u = 0.00001022$. Although the conclusions from these two tests are the same in this case, the test that incorporates the information on the order restriction is more powerful when the true μ satisfies this restriction.

Since one of the primary aims of the trial was to study the efficacy of various dosages of pramipexole, it would be useful to identify intervals of the baseline UPDRS score over which there is a significant treatment effect for each dosage compared to the placebo group. Using the Johnson-Neyman-type method for constrained confidence region construction described above, we obtain the critical value $w = 11.14052$ for the case of simple tree order. Figure 3 shows the confidence regions for the differences between each of the four treatment groups and the placebo group as a function of baseline UPDRS score. These indicate that the mean 10-week changes in UPDRS score for the 1.5 mg/day, 3.0 mg/day, 4.5 mg/day, and 6.0 mg/day groups are significantly different from the mean 10-week change for the placebo group ranging over the values [27.13, 51.79], [27.76, 29.15], [25.83, 60.00], and [27.41, 58.72] of baseline UPDRS score, respectively. These intervals vary from the unrestricted intervals only in the critical value used. In the unrestricted case the critical value is $w = 8F_{8,235;0.05} = 15.82353$ and the corresponding ranges of baseline UPDRS score over which the mean 10-week changes in UPDRS score for the 1.5 mg/day, 4.5 mg/day, and 6.0 mg/day groups are significantly different from the mean 10-week change for the placebo group are [32.36, 38.63], [27.77, 60.00], and [32.10, 44.02], respectively; the mean 10-week change in UPDRS score for the 3.0 mg/day group was not found to differ significantly from the placebo group. The unrestricted confidence intervals are wider, as expected, and consequently the regions of significance are smaller in the unrestricted case than in the restricted case.

6 CONCLUSION

In randomized trials, interactions between treatment and baseline covariates are frequently encountered as exemplified in the randomized trial of pramipexole in Parkinson's disease. This complicates the interpretation of treatment effects and raises a natural question: For what values of the covariates does the treatment have an effect? For the linear model that we

have considered but with no order restrictions imposed on the model parameters, this question can be addressed using the well-known Johnson-Neyman (1936) procedure. This procedure constructs confidence regions for treatment group differences in mean response using a Scheffé correction to enable examination of an infinite number of such differences across all possible values of the (possibly continuous) covariates. Our work extends this procedure to the case where the treatment group means are assumed to satisfy an order restriction. Maximum likelihood estimation of the model parameters subject to the order restriction is accomplished through the careful specification of the constraint matrix $\tilde{\mathbf{A}}$ and solving the resulting quadratic programming problem. We show that the method for simultaneous confidence interval construction has a dual relationship with the F -test, which is asymptotically equivalent to the likelihood ratio test, for overall significance of the treatment group variable under the order restriction. It should be noted that the methods introduced here can be easily extended to the case where the covariance of the error terms is known up to a scale factor, i.e., $\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{V})$ for a known matrix $\mathbf{V}$. Extensions to non-normal responses is a topic of ongoing research.

The pramipexole clinical trial example provides an excellent illustration of the usefulness of these methods. Trials examining dose-response relationships may be particularly well-suited to the application of an order restriction, which can be viewed as a compromise between specifying a particular functional form for the dose-response relationship and not incorporating any prior knowledge concerning this relationship. Of course, the methods can be applied in settings other than clinical trials. An important principle is that the order restriction needs to be specified prior to examination of the data. Also, one should be confident that the order restriction is valid prior to application of the methods.

In cases where the covariates that interact with the grouping variable are continuous, pre-specification of restricted ranges of the covariates is necessary to ensure that the order restriction can be satisfied at all possible values of the covariates in these ranges. It is also necessary to assume that the relationships between these continuous covariates and the dependent variable $\mathbf{Y}$ are linear for the same reason. If these relationships are non-linear, the order restriction is not guaranteed to be satisfied for all possible values of the covariates within pre-specified ranges even if the order restriction holds at the endpoints of these ranges. Additional work is needed to extend the methods presented here to accommodate non-linear relationships.

Acknowledgments

This work partially supported by Award Number T32AI083206 from the National Institute of Allergy and Infectious Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Allergy and Infectious Diseases or the National Institutes of Health.

The authors thank the Parkinson Study Group and the Steering Committee, investigators, and coordinators of the pramipexole trial (Parkinson Study Group, 1997) for providing access to the trial data, as well as the editors and two anonymous reviewers for their helpful comments and suggestions.

References

- Betcher J, Peddada SD. Statistical inference under order restrictions in analysis of covariance using a modified restricted maximum likelihood estimator. *Sankhya Ser B*. 2009; 71:79–96. [PubMed: 20957069]
- Johnson PO, Neyman J. Tests of certain linear hypothesis and their application to some educational problems. *Statistical Research Memoirs*. 1936; 1:57–93.
- Lang, AE.; Fahn, S. Assessment of Parkinson's disease. In: Munsat, TL., editor. *Quantification of Neurologic Deficit*. Boston, MA: Butterworth-Heinemann; 1989. p. 285-309.

- Meyer, MC. Tech rep. Colorado State University; 2010. An Algorithm for Quadratic Programming with Applications in Statistics.
- Mukerjee H, Robertson T, Wright FT. Comparison of Several Treatments With a Control Using Multiple Contrasts. *Journal of the American Statistical Association*. 1987; 82:902–910.
- Parkinson Study Group. Safety and Efficacy of Pramipexole in Early Parkinson Disease: A Randomized Dose-Ranging Study. *JAMA*. 1997; 278:125–130. [PubMed: 9214527]
- Robertson, T.; Wright, FT.; Dykstra, RL. *Order Restricted Statistical Inference*. New York: John Wiley & Sons; 1988.
- Scheffé H. A method for judging all contrasts in the analysis of variance. *Biometrika*. 1953; 40:87–104.
- Shapiro A. Scheffe's method for constructing simultaneous confidence intervals subject to cone constraints. *Statistics & Probability Letters*. 2003; 64:403–406.
- Silvapulle, MJ.; Sen, PK. *Constrained Statistical Inference*. New York: John Wiley & Sons; 2005.

A APPENDIX

The method of Shapiro (2003) considers simultaneous confidence intervals for all linear combinations of the form $\mathbf{a}'\boldsymbol{\mu}$, where $\boldsymbol{\mu} \sim \mathcal{N}_k(\boldsymbol{\mu}, \mathbf{V})$, $\mathbf{a}$ lies in a polyhedral cone denoted by C , $\mathbf{V}$ is known, and $\mathbf{a}'\boldsymbol{\mu}$ is estimated with ν degrees of freedom. The critical value is chosen to satisfy

$$\alpha = \sum_{i=1}^q p_i P\{iF_{i,\nu} \geq w_\alpha\}.$$

Suppose for simplicity that in our problem we have $r = t = 0$. We then have $\boldsymbol{\mu} \sim \mathcal{N}_k(\boldsymbol{\mu}, \mathbf{D})$ and we are interested in all contrasts among the μ_j such that the vector of contrast coefficients lies in a polyhedral cone $C(\mathbf{A}) = \{\mathbf{c} | \mathbf{A}\mathbf{c} = \mathbf{0}\}$, where it is assumed that the $k \times m$ matrix $\mathbf{A}$ has no redundant columns and satisfies the property that $\sum_{i=1}^k A_{ij} = 0, j = 1, \dots, m$ (i.e., the columns of $\mathbf{A}$ are contrast vectors).

Since $\mathbf{A}$ has no redundant columns and they each sum to 0, it is easy to show that $\mathbf{A}$ must be of full rank. Furthermore, $\text{rank}(\mathbf{A}) = m - k + 1$ since there are only $k - 1$ linearly independent contrast vectors of length k . It then follows that there must be exactly $k - 1 - m$ contrast vectors of length k that are linearly independent of the m contrast vectors in $\mathbf{A}$.

Suppose the columns of $\mathbf{B}$ contain these vectors. Let $\mathbf{R}' = \begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix}$ and make the transformation

$$\mathbf{R}\boldsymbol{\mu} = \boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \mathbf{A}'\boldsymbol{\mu} \\ \mathbf{B}'\boldsymbol{\mu} \end{pmatrix}.$$

The newly transformed parameter vector $\boldsymbol{\lambda}$ is a $k - 1$ dimensional vector. We now have $\boldsymbol{\lambda} \sim \mathcal{N}_{k-1}(\mathbf{R}\boldsymbol{\mu}, \mathbf{R}\mathbf{D}\mathbf{R}')$ and we are interested in all linear combinations of the form $\mathbf{a}'\boldsymbol{\lambda}$ such that the first m coefficients of $\mathbf{a}$ are restricted to lie in the negative orthant and the remaining $k - 1 - m$ coefficients are unrestricted. We can now directly apply the method of Shapiro (2003) to find w satisfying

$$\alpha = \sum_{i=1}^{k-1} p_i P\{iF_{i,\nu} \geq w_\alpha\}.$$

Since $k-1-m$ parameters are unrestricted, we can simplify this expression because $p_i = 0$ for all $i = 1, 2, \dots, k-1-m-1$. Therefore, by shifting the indices we find w satisfying

$$\begin{aligned} \alpha &= \sum_{i=1}^{k-1} p_i P\{iF_{i,\nu} \geq w_\alpha\} = \sum_{i=k-1-m}^{k-1} p_i P\{iF_{i,\nu} \geq w_\alpha\} \\ &= \sum_{i=0}^m p_{i+k-1-m} P\{(i+k-1-m)F_{i+k-1-m,\nu} \geq w_\alpha\} \\ &= \sum_{i=0}^h p_{i+d} P\{(i+d)F_{i+d,\nu} \geq w_\alpha\}, \end{aligned}$$

where $h = \text{rank}(\mathbf{A}) = m$ and $d = k-1-m$. We are finally left with

$$\sum_{i=0}^h p_i P\{(i+d)F_{i+d,\nu} \geq w_\alpha\}$$

since we defined the level probabilities to run from $i = 0$ to h rather than from $i = d$ to $h+d$.

When $r = 0$ or $t = 0$ we have 2^t unique vertices of the hyper-rectangle at which these constraints must hold. Following the same process as above it is easy to show that $h = \text{rank}(\tilde{\mathbf{A}}) = 2^t m$ and $d = 2^t(k-1-m)$.

B APPENDIX

Theorem 1

Let F_α^* denote the critical value obtained by the method of Scheffé (1953) for contrast vectors (among the group means) in the space $\mathbb{R}^P$, satisfying

$$\alpha = P\{pF_{p,\nu} > F_\alpha^*\} = P\{(h+d)F_{h+d,\nu} > F_\alpha^*\}$$

and let w denote the critical value obtained by the method of Shapiro (2003) for contrast vectors in the cone $C \subseteq \mathbb{R}^P$ satisfying

$$\alpha = \sum_{i=0}^h p_i P\{(i+d)F_{i+d,\nu} > w_\alpha\}.$$

If $h > 0$, then $F_\alpha^* > w_\alpha$ for all $\alpha \in (0, 1)$.

Proof

A chi-squared random variable with i degrees of freedom, denoted χ_i^2 , is stochastically dominated by a chi-squared random variable with j degrees of freedom, $j > i$, so that $P\{\chi_i^2 > c\} < P\{\chi_j^2 > c\}$. Suppose χ_ν^2 is independent of both χ_i^2 and χ_j^2 for some $\nu > 0$. Since

$\chi^2_\nu > 0$, we find $P\{\frac{i\chi^2_{i^2}/i}{\chi^2_\nu/\nu} > c\} < P\{\frac{j\chi^2_{j^2}/j}{\chi^2_\nu/\nu} > c\}$; therefore, $P\{iF_{i^2} > c\} < P\{jF_{j^2} > c\}$ for all $i < j$, where F_{d_1, d_2} denotes a random variable with an F distribution with d_1 and d_2 degrees of freedom. Now

$$\begin{aligned} P\{(h+d)F_{h+d, \nu} > F_\alpha^*\} &= \alpha = \sum_{i=0}^h p_i P\{(i+d)F_{i+d, \nu} > w_\alpha\} = \sum_{i=d}^{h+d} p_{i-d} P\{iF_{i, \nu} > w_\alpha\} \\ &= p_h P\{(h+d)F_{h+d, \nu} > w_\alpha\} + \sum_{i=d}^{h+d-1} p_{i-d} P\{iF_{i, \nu} > w_\alpha\}, \end{aligned}$$

where the p_i are the level probabilities and $\sum_{i=0}^h p_i = 1$. From above, $P\{iF_{i^2} > c\} < P\{(h+d)F_{h+d, \nu} > c\}$ for $i = d, d+1, \dots, h+d-1$. Since $C \subset \mathbb{R}^p$, $\sum_{i=0}^{h-1} p_i > 0$ and so

$$\begin{aligned} P\{(h+d)F_{h+d, \nu} > F_\alpha^*\} &= p_h P\{(h+d)F_{h+d, \nu} > w_\alpha\} + \sum_{i=d}^{h+d-1} p_{i-d} P\{iF_{i, \nu} > w_\alpha\} \\ &< p_h P\{(h+d)F_{h+d, \nu} > w_\alpha\} + \sum_{i=d}^{h+d-1} p_{i-d} P\{(h+d)F_{h+d, \nu} > w_\alpha\} \\ &= P\{(h-d)F_{h+d, \nu} > w_\alpha\}, \end{aligned}$$

from which we can conclude that $F_\alpha^* > w_\alpha$.

The consequence of Theorem 1 is that the confidence regions incorporating the order restriction will always be smaller than the confidence regions not incorporating the restriction. In fact, the sizes of the confidence regions will be equal only if $C = \mathbb{R}^p$ since in that case $\sum_{i=0}^{h-1} p_i = 0$.

Figure 1.

Analysis of covariance model fit to the pramipexole trial data with no order restrictions imposed on the group mean responses. There was a significant interaction between treatment group and baseline UPDRS score ($p = 0.002$) with a pattern suggesting that the effect of pramipexole may be greater in patients with higher baseline UPDRS scores (more severe Parkinson's disease features). In patients with higher baseline UPDRS scores all four treatment groups had a larger improvement than the placebo group.

Figure 2.

Analysis of covariance model fits to the pramipexole trial data under the simple tree order and simple order restrictions. The mean responses across the treatment groups now satisfy the corresponding order restriction for every value of baseline UPDRS score in the specified range of 5 to 60 points.

Figure 3.

Johnson-Neyman-type confidence regions identifying intervals of baseline UPDRS score corresponding to significant treatment effects for the four pramipexole dosages, subject to the simple tree order restriction. A significant treatment effect is identified for an interval for which the upper confidence band for the treatment effect falls below the value of zero. These intervals, indicated by the dashed vertical lines, are [27.13, 51.79] for 1.5 mg/day, [27.76, 29.15] for 3.0 mg/day, [25.83, 60.00] for 4.5 mg/day, and [27.41, 58.72] for 6.0 mg/day.