The correlation structure of longitudinal measurements of vision in patients with macular degeneration.

RESEARCH ARTICLE

Pharmaceutical Statistics

Received 8 October 2009, Revised 5 January 2010, Accepted 11 January 2010 Published online in Wiley Interscience

(www.interscience.wiley.com) DOI: 10.1002/pst.419

The correlation structure of longitudinal measurements of vision in patients with macular degeneration

Tomasz Burzykowski[a,b]* and M. Buyse[a,b]

***Background*: In age-related macular degeneration (ARMD) trials, the FDA-approved endpoint is the loss (or gain) of at least three lines of vision as compared to baseline. The use of such a response endpoint entails a potentially severe loss of information. A more efficient strategy could be obtained by using longitudinal measures of the change in visual acuity. In this paper we investigate, by using data from two randomized clinical trials, the mean and variance–covariance structures of the longitudinal measurements of the change in visual acuity.**

***Methods*: Individual patient data were collected in 234 patients in a randomized trial comparing interferon-α with placebo and in 1181 patients in a randomized trial comparing three active doses of pegaptanib with sham. A linear model for longitudinal data was used to analyze the repeated measurements of the change in visual acuity.**

***Results*: For both trials, the data were adequately summarized by a model that assumed a quadratic trend for the mean change in visual acuity over time, a power variance function, and an antedependence correlation structure. The power variance function was remarkably similar for the two datasets and involved the square root of the measurement time.**

***Conclusions*: The similarity of the estimated variance functions and correlation structures for both datasets indicates that these aspects may be a genuine feature of the measurements of changes in visual acuity in patients with ARMD. The feature can be used in the planning and analysis of trials that use visual acuity as the clinical endpoint of interest. Copyright © 2010 John Wiley & Sons, Ltd.**

Keywords: visual acuity; longitudinal data; variance function; correlation structure

BACKGROUND

In age-related macular degeneration (ARMD) trials, the FDA-approved endpoint is a binary response defined as the loss (or gain) of at least three lines of vision as compared to baseline one year after starting therapy. Secondary endpoints in ARMD trials include further binary responses defined as the loss (or gain) of more than a pre-defined number of lines of vision as compared to baseline at different time points.

The use of such response endpoints entails a potentially severe loss of information, for two main reasons. First, it does not allow to differentiate between patients who lose different numbers of lines of vision. Second, it does not take into account the information available at other time points. Additional problems of binary endpoints at fixed time points include misclassification of the outcome, potential for a ceiling or floor effect, and the need to impute missing data [1].

A more efficient strategy could be obtained by using longitudinal measures of the change in visual acuity, defined as the change in the number of letters correctly read as compared to baseline. It is worth noting that the binary response is actually closely linked to the change in visual acuity, because each line on the vision chart consists of five letters. Thus, the number of incorrectly read letters gives approximate information about the number of lines lost.

An important difficulty in using longitudinal measurements of the change in visual acuity is the need to use models that correctly reflect the variability and dependence between the measurements. In this paper we show, by using data from several randomized clinical trials, that the mean and variance–covariance structures of the longitudinal measurements of the change in visual acuity can be described by relatively simple models that seem to fit the data well for two different experimental drugs. Use of these models should facilitate the planning and analysis of trials that use such measurements.

MATERIAL AND METHODS

Trials

The first of the two randomized trials included in our analyses compared interferon-α with placebo [2]. The second trial compared different three doses (0.3, 1, and 3 mg) of intravitreous injections of pegaptanib, an anti-vascular endothelial growth factor therapy, or sham injections (with a syringe applied on the surface of the eye to simulate the pressure of

[a]*International Drug Development Institute (IDDI), Louvain-la-Neuve, Belgium*

[b]*Interuniversity Institute for Biostatistics and Statistical Bioinformatics (I-BioStat), Hasselt University, Diepenbeek, Belgium*

**Correspondence to: Tomasz Burzykowski, Hasselt University, Agoralaan, Building D, 3590 Diepenbeek, Belgium.*
E-mail: tomasz.burzykowski@uhasselt.be

an injection), every 6 weeks over a period of one year [3]. Tables I and II present summary characteristics of the measurements of the change in visual acuity versus baseline obtained in the two trials.

Methods

The longitudinal measurements of the change in visual acuity versus baseline were analyzed by using a general linear model for longitudinal data [4]. For each treatment arm, the mean value of the change in visual acuity was modelled as a quadratic function of time, with treatment-specific coefficients. Three forms of the variance–covariance structure of the visual acuity measurements were assumed. Model 1 used a general structure with unconstrained correlations and variances. Model 2 specified an antedependence correlation structure, with unconstrained variances. Model 3 assumed an antedependence correlation structure and variances expressed as a power function of time. Finally, Model 4 was similar to Model 3, but the correlation coefficients, which defined the antedependence structure, were modeled as a linear function of time on the scale defined by Fisher's Z-transform.

The general variance–covariance structure assumes that variances of the longitudinal measurements of change in visual acuity differ freely between time points. Thus, for subsequent K measurements, different K variances are assumed. Additionally, the correlation matrix for the measurements is expressed as follows:

$$\begin{pmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,K-1} & \rho_{1,K} \\ \rho_{1,2} & 1 & \cdots & \rho_{2,K-1} & \rho_{2,K} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{1,K-1} & \rho_{2,K-1} & \cdots & 1 & \rho_{K-1,K} \\ \rho_{1,K} & \rho_{2,K} & \cdots & \rho_{K-1,K} & 1 \end{pmatrix}$$

Thus, it is defined by the values of $K(K+1)/2$ correlation coefficients $\rho_{1,2}, \rho_{1,3}, \ldots, \rho_{K-1,K}$, where ρ_{ij} is the correlation coefficient between the ith and the jth measurement. (We assume that, within a trial, the measurements were made for all patients at approximately the same sequence of time points.) It follows that Model 1 described the variance–covariance structure by using $K(K+1)/2+K = K(K+3)/2$ parameters.

The antedependence correlation structure [5] is defined as follows:

$$\begin{pmatrix} 1 & \rho_1 & \cdots & \rho_1\rho_2\cdot\ldots\cdot\rho_{K-2} & \rho_1\rho_2\cdot\ldots\cdot\rho_{K-2}\rho_{K-1} \\ \rho_1 & 1 & \cdots & \rho_2\cdot\ldots\cdot\rho_{K-2} & \rho_2\cdot\ldots\cdot\rho_{K-2}\rho_{K-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_1\rho_2\cdot\ldots\cdot\rho_{K-2} & \rho_2\cdot\ldots\cdot\rho_{K-2} & \cdots & 1 & \rho_{K-1} \\ \rho_1\rho_2\cdot\ldots\cdot\rho_{K-2}\cdot\rho_{K-1} & \rho_1\rho_2\cdot\ldots\cdot\rho_{K-2} & \cdots & \rho_{K-1} & 1 \end{pmatrix}$$

Thus, it is defined by the values of only $K-1$ coefficients, $\rho_1, \rho_2, \ldots, \rho_{K-1}$, where ρ_i is the correlation coefficient between the ith and the $(i+1)$th measurement. It follows that Model 2 described the variance–covariance structure by using $K-1+K = 2K-1$ parameters.

Model 3 assumed the antedependence correlation structure. Additionally, it assumed that the variance of the measurements of change in visual acuity at time t was equal to $\sigma^2 t^{\delta}$. Thus, it used only two parameters, σ and δ, to express the different K variances. Consequently, it described the variance–covariance structure by using only $K-1+2 = K+1$ parameters.

Finally, Model 4 assumed both the antedependence correlation structure and the power-of-time variance function.

Table I. Summary statistics for the change in visual acuity measurements obtained in the interferon-α trial.

	Interferon-α		Placebo	
Week	N	Mean (SD)	N	Mean (SD)
4	114	−3.51 (8.86)	117	−1.30 (7.71)
12	110	−5.88 (11.62)	117	−2.27 (11.73)
24	102	−9.07 (14.08)	112	−5.70 (13.83)
52	90	−15.48 (15.28)	105	−11.18 (16.42)

N = number of measurements; SD = standard deviation.

Additionally, the correlation coefficients, which defined the antedependence structure, were assumed to depend linearly on time on the scale defined by Fisher's Z-transformation. More specifically, the model assumed that

$$z_i = \log(1+\rho_i) - \log(1-\rho_i) = \gamma_0 + \gamma_1 t_i$$

where ρ_i is the ith of the $K-1$ correlation coefficients, which define the antedependence structure, and t_i is the first time coordinate of the coefficient (recall that ρ_i is the correlation coefficient between the ith and the $(i+1)$th measurement). Consequently, the model described the variance–covariance structure by using only $2+2=4$ parameters. Thus, it was the most parsimonious one among the four considered models.

It is worth noting that Models 1–4 define a sequence of nested models, with Model 4 nested within Model 3, Model 3 nested within Model 2, and Model 2 nested within Model 1.

The fit of each of the four models was assessed by using the scaled residuals [6]. The residuals should be approximately uncorrelated and follow the standard normal distribution. The normality assumption was checked by using the normal Q-Q plot. Between the models, the fit was compared by using the restricted-likelihood-ratio test (LRT) and the Akaike Information Criterion (AIC) [4]. The latter was based on the number of variance–covariance parameters, given that all compared models had the same mean structure.

The calculations were performed using SAS v.9.3 PROC MIXED (Models 1 and 2) and self-written SAS-IML code (Models 3 and 4).

RESULTS

The interferon-α trial

Table III shows the values of minus twice log-restricted-likelihood, the number of variance–covariance parameters, and the resulting values of the AIC for Models 1–4. For none of the Models 2–4, the difference in minus twice log-restricted-likelihood with respect to Model 1 is statistically significant at the 0.05 significance level. This suggests that any of the simpler models offers a fit comparable to (the most general) Model 1.

Table II. Summary statistics for the change in visual acuity measurements obtained in the pegaptanib trial.

	0.3 mg		1 mg		3 mg		Sham	
Week	*N*	Mean (SD)	*N*	Mean (SD)	*N*	Mean (SD)	*N*	Mean (SD)
6	286	−1.56 (9.20)	292	−1.23 (9.82)	286	−2.37 (10.13)	291	−4.08 (9.66)
12	289	−3.14 (11.56)	291	−2.17 (11.21)	283	−4.20 (12.19)	288	−6.36 (11.32)
18	269	−3.69 (12.12)	291	−3.91 (12.15)	281	−5.42 (13.82)	287	−8.72 (12.56)
24	273	−4.06 (13.21)	287	−4.29 (13.33)	283	−6.23 (14.30)	281	−9.85 (13.98)
30	271	−4.94 (14.20)	285	−5.01 (13.90)	278	−7.20 (15.47)	282	−11.59 (15.95)
36	265	−5.84 (14.51)	278	−5.28 (14.79)	273	−7.28 (15.22)	278	−12.42 (15.89)
42	271	−6.49 (14.82)	270	−6.09 (15.14)	267	−7.62 (16.19)	275	−13.16 (16.16)
48	266	−6.83 (15.18)	267	−5.85 (15.37)	259	−8.35 (16.59)	269	−13.48 (16.84)
54	271	−7.55 (15.65)	275	−6.75 (16.11)	264	−10.30 (16.66)	275	−14.78 (17.73)

N = number of measurements; SD = standard deviation.

Table III. Minus twice log-restricted-likelihood (−2REML), number of variance–covariance parameters, and the resulting AIC for Models 1–4.

	Interferon-α trial			Pegaptanib trial		
	−2REML	No. of parameters	AIC	−2REML	No. of parameters	AIC
Model 1	6400.1	10	6420.1	68 067.4	45	68 157.4
Model 2	6405.6	7	6419.6	68 366.3	17	68 400.3
Model 3	6407.0	5	6417.0	68 379.4	10	68 399.4
Model 4	6407.4	4	6415.4	68 415.6	4	68 423.6

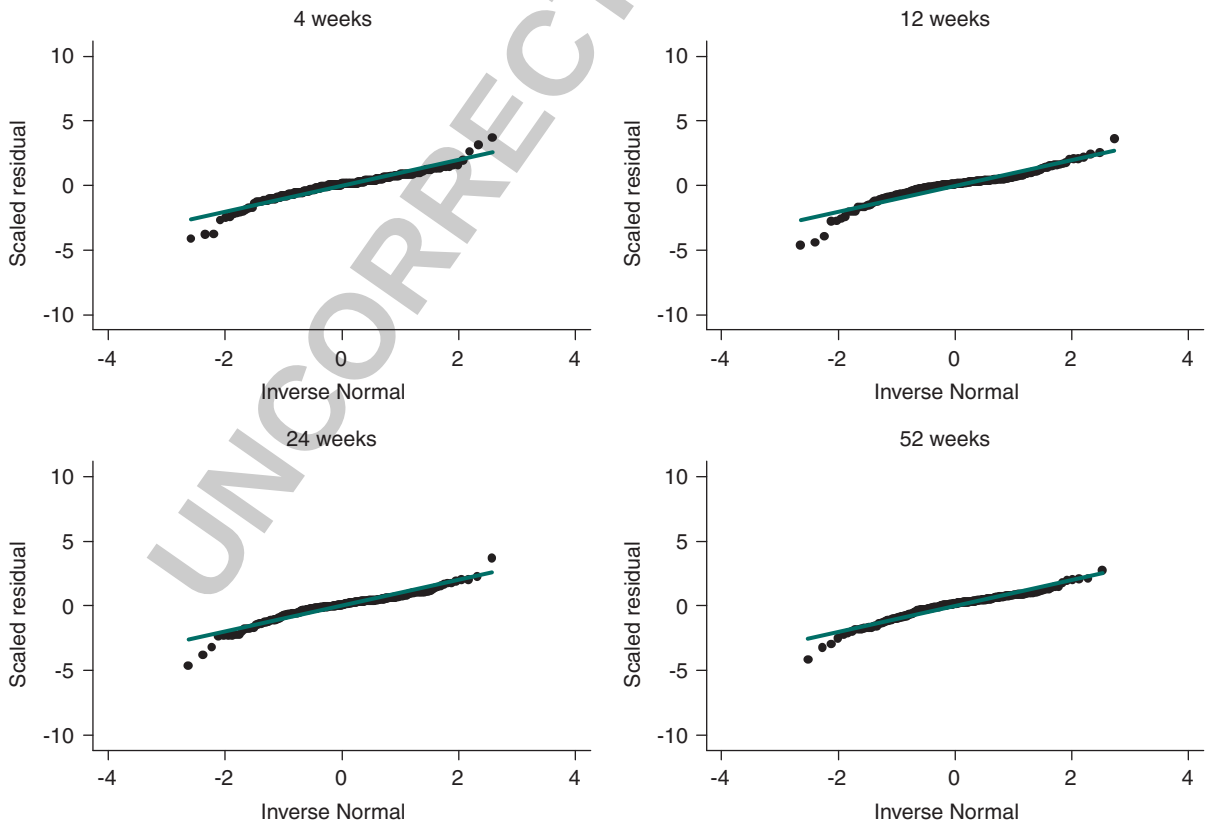

Figure 1. Normal *Q-Q* plots for the scaled residuals for Model 3 fitted to the interferon-α trial data.

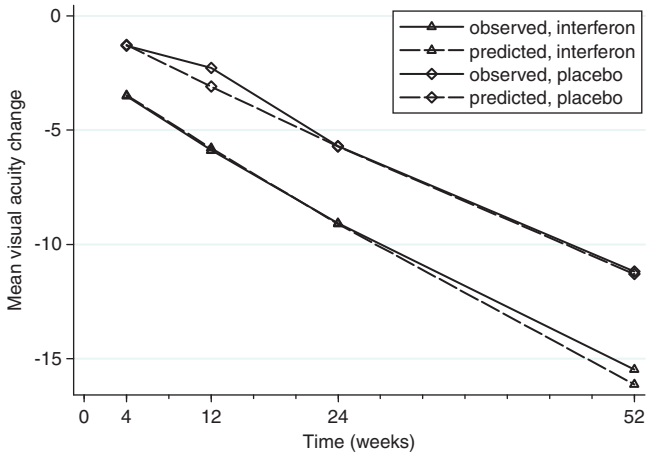

Figure 2. Observed and predicted mean values of the change in visual acuity for the interferon-α trial, using Model 3.

The observed AIC values indicate that all four models present a similar fit to the data. Formally speaking, the lowest AIC value (6417.0), indicating the best fit, is obtained for Model 4. However, the AIC value (6415.4) for Model 3 is not much different.

Figure 1 shows the normal *Q-Q* plots for the scaled residuals at each measurement time point for Model 3. The plots suggest a satisfactory fit of the model. This conclusion is confirmed by Figure 2, which presents the observed and predicted means of the visual acuity change for the two treatment arms of the interferon-α trial. The plot indicates a linear mean trend for both arms. This is confirmed by the 95% confidence intervals of the estimated coefficients, which are shown in Table IV. The intervals for the time-squared coefficient include 0, suggesting that a linear function of time is adequate to describe the mean trend.

According to Model 3, the variance of the change in visual acuity at time t is approximately equal to $(5.7)^2 t^{1/2} = 32.5\sqrt{t}$. The 4×4 correlation matrix of the four visual acuity

Table IV. Estimated mean trend coefficients for Model 3.

	Interferon-α trial		Pegaptanib trial			
	Interferon-α	Placebo	0.3 mg	1 mg	3 mg	Sham
Intercept	−2.28 (−4.14,−0.42)	−0.33 (−2.17,1.50)	−1.22 (−2.63,0.18)	−0.82 (−2.22,0.59)	−0.03 (−1.43,1.37)	−1.66 (−3.06,−0.26)
Time	−0.30 (−0.50,−0.10)	−0.23 (−0.43,−0.04)	−0.21 (−0.33,−0.10)	−0.19 (−0.31,−0.07)	−0.21 (−0.33,−0.09)	−0.42 (−0.54, −0.31)
Time^2	0.0006 (−0.002,0.004)	0.0005 (−0.002,0.004)	0.0009 (−0.001,0.003)	0.001 (−0.001,0.003)	0.001 (−0.0004,0.003)	0.003 (0.001,0.005)

Table V. Estimated variance–covariance structure parameters for Models 3 and 4 (for Model 4, parameters $\rho_1, \rho_2, \ldots, \rho_8$ are computed from the estimated values of γ_0 and γ_1).

	Interferon-α trial		Pegaptanib trial	
Parameter	Model 3	Model 4*	Model 3	Model 4*
σ	5.73 (4.96,6.63)	5.73 (4.95,6.62)	6.27 (5.79,6.78)	6.57 (6.07,7.10)
δ	0.28 (0.22,0.33)	0.28 (0.23,0.33)	0.25 (0.22,0.27)	0.23 (0.21,0.26)
ρ_1	0.60 (0.51,0.67)	0.58 (0.50,0.66)	0.73 (0.70,0.75)	0.77 (0.75,0.79)
ρ_2	0.66 (0.58,0.72)	0.67 (0.62,0.72)	0.81 (0.79,0.83)	0.81 (0.79,0.82)
ρ_3	0.78 (0.72,0.83)	0.77 (0.71,0.82)	0.85 (0.84,0.87)	0.84 (0.82,0.84)
ρ_4	—	—	0.88 (0.86,0.89)	0.86 (0.85,0.87)
ρ_5	—	—	0.88 (0.87,0.89)	0.88 (0.87,0.89)
ρ_6	—	—	0.91 (0.90,0.92)	0.90 (0.89,0.91)
ρ_7	—	—	0.92 (0.91,0.93)	0.92 (0.91,0.92)
ρ_8	—	—	0.93 (0.92,0.93)	0.93 (0.92, 0.94)
γ_0	—	1.19 (0.91, 1.48)	—	1.87 (1.77, 1.96)
γ_1	—	0.036 (0.018, 0.054)	—	0.030 (0.027, 0.033)

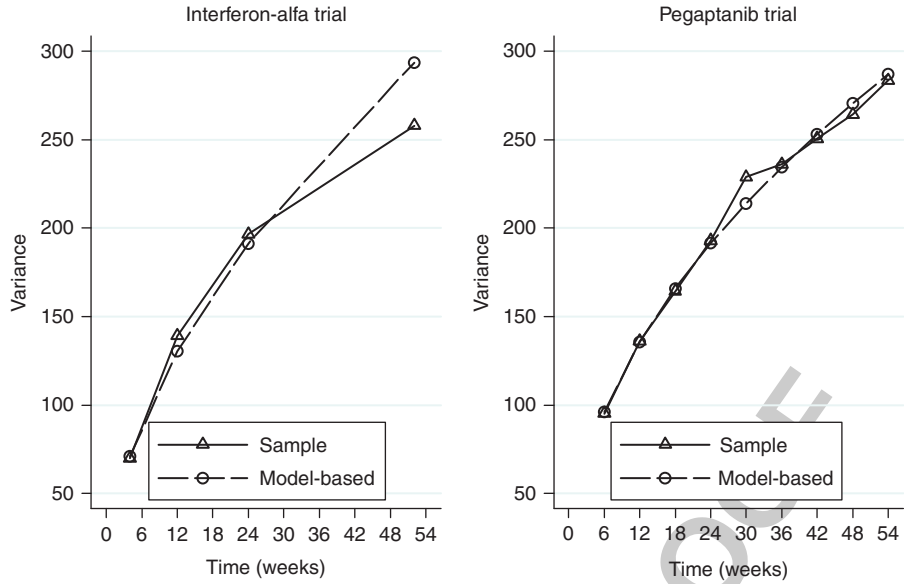

Figure 3. Sample variances and model-based estimates for Model 3.

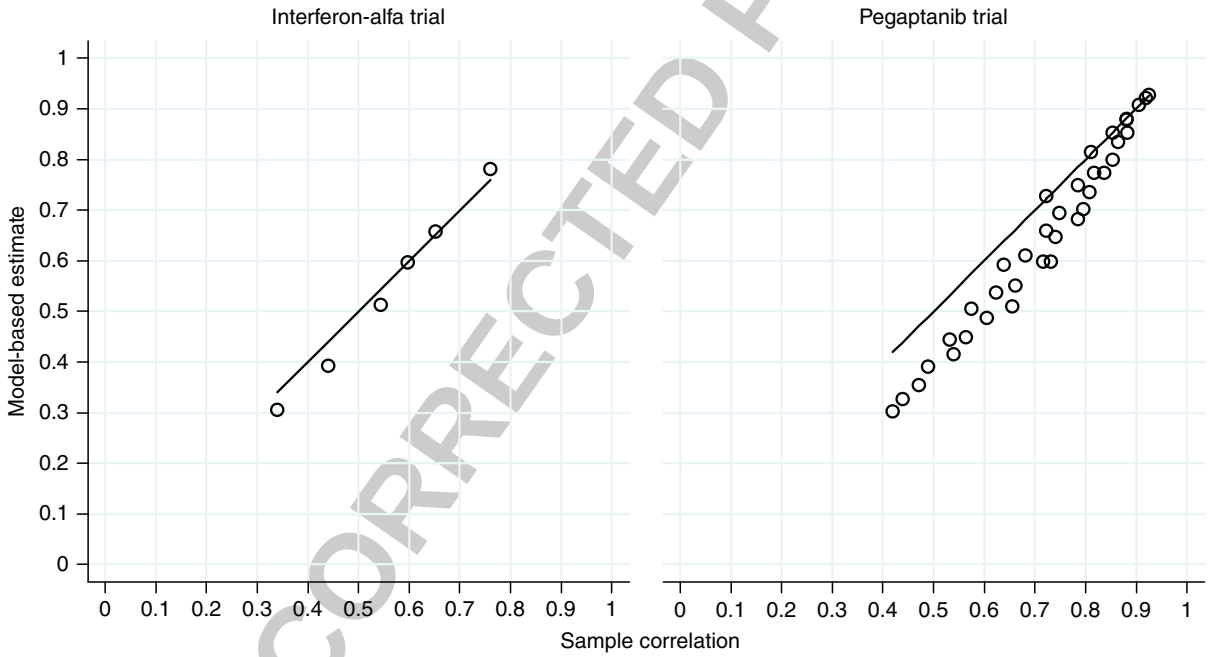

Figure 4. Sample correlations and model-based estimates for Model 3 (with 45° lines).

change measurements can be expressed by using only three correlation coefficients ρ_1, ρ_2, and ρ_3. The estimated values of the coefficients are displayed in Table V. The table also contains the estimates obtained for Model 4.

The panel on the left-hand side of Figure 3 presents the sample variances and the corresponding model-based estimates for Model 3. Apart from a slight overestimation at 52 weeks, the model-based estimates are in a very good agreement with the sample variances.

Similarly, the panel on the left-hand side of Figure 4 presents the sample correlations and their corresponding model-based estimates for Model 3. Again, a very close agreement between the two sets of estimates can be observed.

The pegaptanib trial

The values of minus twice log-restricted-likelihood, the number of variance–covariance parameters, and the resulting values of the AIC for Models 1–4 are shown in Table III. The difference in minus twice log-restricted-likelihood between Models 1 and 2 is statistically significant at the 0.05 significance level. Also, the lowest AIC value, indicating the best fit, is obtained for Model 1.

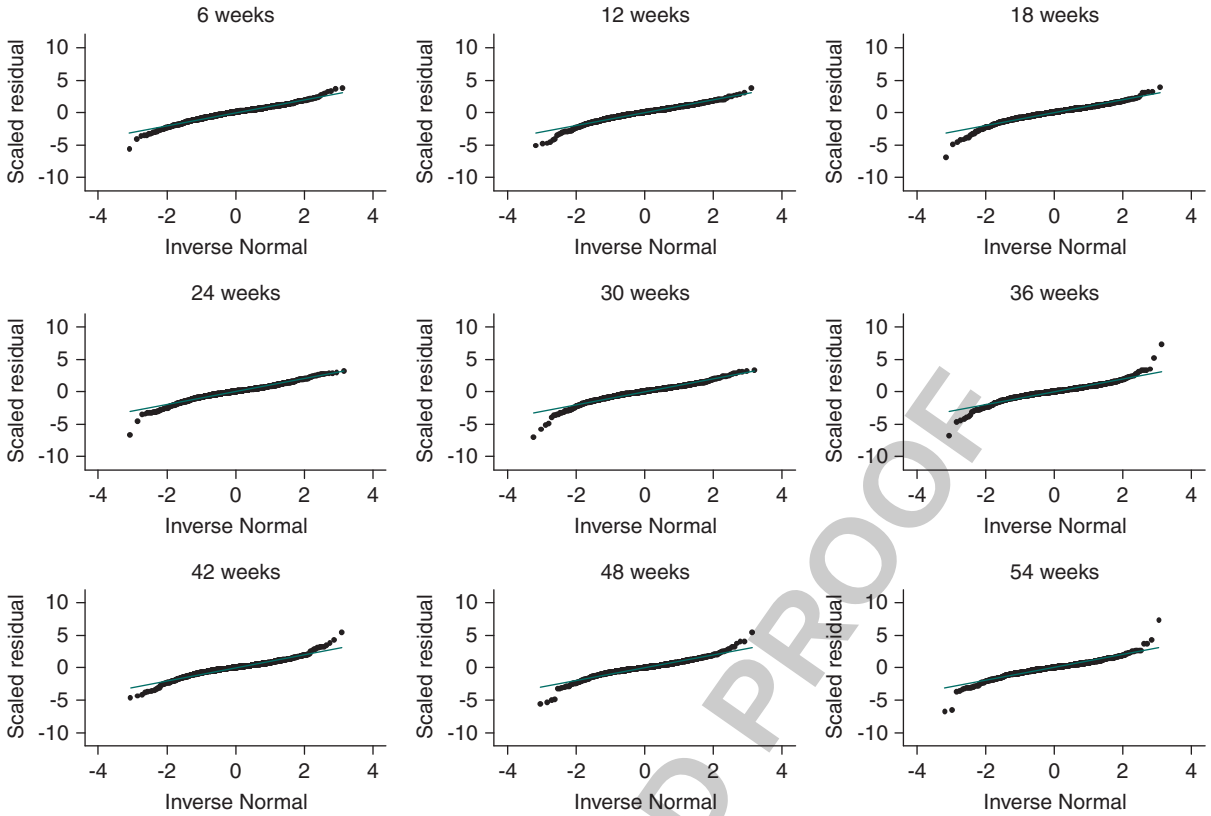

Figure 5. Normal *Q-Q* plots for the scaled residuals using Model 3 fitted to the pegaptanib trial.

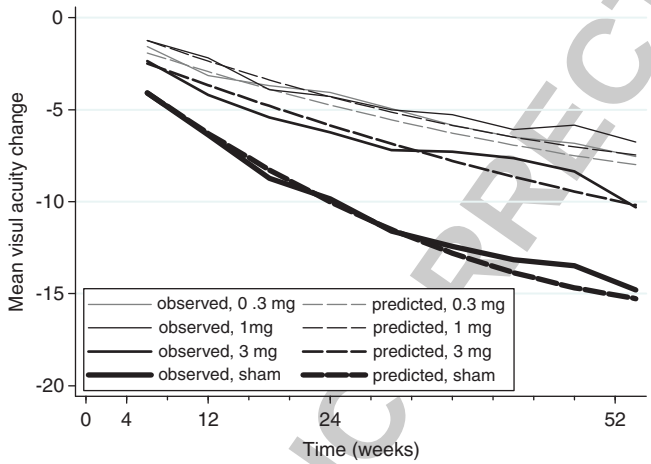

Figure 6. Observed and predicted mean values of the change in visual acuity for the pegaptanib trial, using Model 3.

Given that the data included up to nine visual acuity change measurements for a single patient, this result might be expected, because the variance–covariance structure of the measurements might be potentially complex. However, it is worth noting that, according to the AIC values, Model 3 offers a better fit than both the simpler Model 4 and the more complex Model 2. Also, the result of the LRT comparing Models 2 and 3 is not statistically significant at the 0.05 significance level. Moreover, the fit of Model 3 is quite satisfactory, as can be seen from the normal *Q-Q* plots for the scaled residuals at each measurement time point (Figure 5) and from the plot of the observed and predicted means of the visual acuity change for the four treatment arms (Figure 6). As Model 3 is more parsimonious, it is preferred over Model 1. The plot and the 95% confidence intervals of the estimated coefficients, which are shown in Table IV, suggest a quadratic mean trend for the sham group and linear trends for the active treatment groups.

According to Model 3, the variance of the change in visual acuity at time t is approximately equal to $(6.3)^2 t^{1/2} = 39.7\sqrt{t}$. The 9×9 correlation matrix of the four visual acuity change measurements can be expressed by using only eight correlation coefficients $\rho_1, \rho_2, \ldots, \rho_8$. The estimated values of the coefficients are displayed in Table IV.

The panel on the right-hand side of Figure 3 presents the sample variances and the corresponding model-based estimates for Model 3. Apart from a slight underestimation at 30 weeks, the model-based estimates are in an excellent good agreement with the sample variances. The panel on the right-hand side of Figure 4 presents the sample correlations and their corresponding model-based estimates for Model 3. For correlations below 0.7, i.e. for visual acuity measurements made at more separated time points, the model-based estimates are lower than their sample-based counterparts. Note, however, that the latter do not take into account the differences due to the treatment effects.

Both trials

Figure 7 presents Fisher's *Z*-transformation of the estimates of the antedependence structure defining the correlation coefficients for

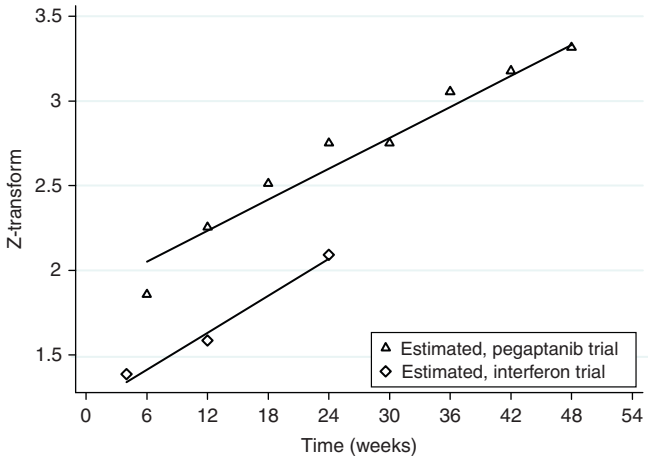

Figure 7. Fisher's *Z*-transformation of the estimated correlation coefficients defining the correlation structure of Model 3. The straight lines are obtained from Model 4 and describe the linear dependence of the transformed coefficients on their first time coordinates (indicated on the horizontal axis).

both trials (Table V). The plot includes the straight lines obtained from Model 4 that describe the linear dependence of the transformed coefficients on their first time coordinates. The plot illustrates that the *Z*-transformation of the correlation coefficients increases linearly in time, with approximately the same slope for the two trials, but with a different intercept. The linear equations are $z = 1.195$ (SE 0.145) $+ 0.036$ (SE 0.009) $\cdot t$ for the interferon-α trial and $z = 1.869$ (SE 0.049) $+ 0.030$ (SE 0.002) $\cdot t$ for the pegaptanib trial. Thus, the difference in slopes does not differ statistically from 0 (95% confidence interval, CI (-0.012, 0.024)), whereas the vertical shift between the lines in Figure 7 is approximately constant and equal to 0.674 (95% CI (0.368, 0.980)).

DISCUSSION

The analysis reveals remarkably consistent results regarding the variance–covariance structure of the longitudinal measurements of the change in visual acuity versus baseline. First, the variance of the measurements can be approximately expressed as a constant (close to 36) times the square root of the time at which the measurement was taken. Second, the correlations of the measurements can be expressed by using the antedependence structure. Moreover, the correlation coefficients, defining the antedependence structure, are linearly increasing with time with the same slope, but with different intercepts (on Fisher's *Z*-transformation scale).

These results can be used in designing experiments, in which the use of longitudinal measurements of the change in visual acuity versus baseline is of interest. In particular, they can be used for sample size calculations [4,7].

It would be of interest to validate our results by using data from randomized trials with other drugs recently approved for ARMD. Interferon-α was shown not to have any significant impact on the visual acuity of patients with ARMD [1]; hence, the correlation structure identified in the interferon-α trial may well reflect the natural history of the disease and the intrinsic variability of visual acuity measurements. On the contrary, pegaptanib had a highly significant effect on visual acuity, but it did not seem to affect the correlation structure of repeated measurements of visual acuity to any great extent. It is tempting, therefore, to hypothesize that treatments for ARMD affect the mean values of the measurements' visual acuity over time, but not their correlation structure. This hypothesis awaits further confirmation.

Acknowledgement

The authors are grateful to the pharmaceutical companies Roche, OSI/Eyetech, and Pfizer for providing individual patient data for the trials used in this analysis. Prof. Burzykowski acknowledges financial support from the IAP research network nr P6/03 of the Belgian government (Belgian Science Policy).

REFERENCES

[1] Beck RW, Maguire MG, Bressler NM, *et al*. Visual acuity as an outcome measure in clinical trials of retinal diseases. *Ophthalmology* 2007; **114**:1804–1809.

[2] Pharmacological Therapy For Macular Degeneration Study Group. Interferon α-IIA is ineffective for patients with choroidal neovascularization secondary to age-related macular degeneration. Results of a prospective randomized placebo-controlled clinical trial. *Archives of Ophthalmology* 1997; **115**:865–872.

[3] Gragoudas AS, Adamis AP, Cunningham ET, *et al*. for the VEGF Inhibition Study in Ocular Neovascularization Clinical Trial Group. Pegaptanib for neovascular age-related macular degeneration. *New England Journal of Medicine* 2004; **351**:2805–2816.

[4] Verbeke G, Molenberghs G. *Linear Mixed Models for Longitudinal Data*. Springer: New York, 2005.

[5] SAS/STAT's User Guide: PROC MIXED. SAS Institute, 2003.

[6] Schabenberger O. Mixed model influence diagnostics. *Proceedings of the Twenty-Ninth Annual SAS® Users Group International Conference* 2004; paper 189-29.

[7] Stroup WW. Mixed model procedures to assess power, precision, and sample size in the design of experiments. *ASA Proceedings of the Biopharmaceutical Section*. American Statistical Association: Alexandria, 1999; 15–24.

AQ1