

Statistical Appendix to *Modelling Dose-Response in Amblyopia: towards a child specific treatment plan*

David A. Stephens*

Department of Mathematics, Imperial College London

and

Department of Mathematics and Statistics, McGill University, Montreal

1 Introduction

The Monitored Occlusion Treatment of Amblyopia Study (MOTAS) (Stewart et al. (2004)) was the first clinical study aimed at determining the dose-response relationship of occlusion therapy and exhibited several advantages over previous studies. Most importantly, the effects of refraction were differentiated from those of occlusion and concordance (compliance) with treatment was measured objectively. The study is described in more detail in section 2.

The aim of this paper is to study more fully the dose-response relationship of occlusion, and to understand how this relationship is modified by patient characteristics. Previous analyses of these data (Stewart et al. (2004)) have been largely expository and exploratory. For example, in Stewart et al. (2004), the total dose vs. total improvement relationship was computed using semiparametric regression and bootstrap methods over the whole patient cohort. In this paper, we study the dose-response relationship longitudinally and on an interval-by-interval basis, modelling the effect of covariates using linear mixed regression models, in order to gain a quantitative understanding of the therapeutic benefits of occlusion therapy.

In the next section, we give more details of the MOTAS study design and outline the strategy behind our analyses in section 2.3. In section 3 we describe a linear mixed model analysis accounting for the repeated measures nature of the data. An interval-by-interval analysis follows; the results are presented in section 4.1. Bayesian modelling and inference can be found in 5. A Bayesian analysis of dosing strategies is shown in section 5.3.

2 The MOTAS study

The design for this prospective study has been reported in detail elsewhere (Stewart et al. (2002)). It comprised three phases: *baseline*, *refractive adaptation* (or *refraction*) and *occlusion* (see section 2.1). Data, in the form of profile plots from the refraction and occlusion phases, are depicted in Figure 1. Each line tracks a patient's change in visual acuity as measured on the logarithm of Minimum Angle of Resolution (logMAR) scale over successive visits to the clinician; improvement is indicated by a logMAR decrement. The figure indicates the general beneficial effect of occlusion therapy, as the majority of individuals have generally downward sloping trajectories. The amount of occlusion dose within each interval is not reflected in this plot. In the lower panel, profiles for four selected

*Address for Correspondence: Professor D. A. Stephens, Department of Mathematics and Statistics, McGill University, Burnside Hall, 805 Sherbrooke Street, Montreal QC, Canada.

patients are depicted, with their time of entry into the occlusion phase indicated by vertical dotted lines.

2.1 *Study Design and Implementation*

The MOTAS study design and details of the enrollment criteria have been published previously (Stewart et al. (2002)). Prior to study entry, all children had a full ophthalmic assessment. Children who required spectacle correction entered the refractive adaptation phase, whereas those not requiring spectacle correction entered the occlusion phase directly. Children were instructed to wear spectacles (where prescribed) full-time and scheduled to return for vision assessment at six-weekly intervals from onset of spectacle wear until 18 weeks of refractive adaptation was completed. Children remaining eligible, that is, those who still met the study protocol’s operational definition of amblyopia, entered the occlusion phase and were prescribed six hours occlusion per day. Occlusion episodes were recorded to the nearest minute by an occlusion dose monitor (ODM) (Fielder et al. (1994)).

The ODM - a device developed and extensively piloted by the MOTAS cooperative - consists of an eye patch with two small electrodes attached to its under-surface connected to a battery-powered data logger. Both visual function and monitored occlusion dose were recorded at two-weekly intervals until acuity ceased to improve; treatment was concluded in a pragmatic fashion, after two inflexions in an acuity against time plot or identical acuity measurements on three consecutive visits. Participants returned to standard clinical care on completion of the occlusion phase .

The study enrolled a total of 116 children aged between 36 and 100 months over a period of two years. Of these, 87 are suitable for inclusion in the statistical analysis; fifteen did not enter the occlusion phase, but had their visual impairment successfully treated with refractive adaptation - these children are included in the analysis. The remaining 72 although prescribed occlusion for six hours a day received different occlusion doses over different follow-up periods. However, three children entered and were deemed to have completed occlusion, and yet received a total occlusion dose of less than fifteen minutes. Such a dose is considered negligible and thus, in our analysis, these three children were reclassified as not having undergone occlusion. The 29 children excluded from the analysis were either deemed not to meet the inclusion criteria or lost to follow-up after a small number of clinic visits.

A full description of the study base, and those children excluded from the current analysis, is given in Stewart et al. (2004). We will assume that the loss to follow-up is not informative, and that the 87 patients entering the analysis are a suitable random sample from the amblyopic population.

The following data were collected longitudinally: (i) the primary measure of visual function *outcome* measurement was logMAR visual acuity, scored by letter, with the type of test used selected in accordance with the reading ability of the child, and hence was generally age-dependent (the test used at the first session was used throughout the study period); (ii) occlusion *dose* was monitored using an ODM as described above. At the start of the occlusion phase, the investigator explained to the parents and child the practicalities of wearing the monitor. At each subsequent visit, data from the ODM was downloaded to a PC and parents were given the opportunity to review their child’s concordance.

2.2 *The Lack of Randomization and a Control Group*

In the MOTAS study, it was not possible to implement a randomized controlled design for two important reasons. First, it was not considered ethically justified to deny any child the possibility of obtaining a maximum possible beneficial dose. Secondly, it was not possible to propose a sensible and implementable designed experiment, as the level of concordance for any child was not known before the experiment was started. Despite these ethical and practical difficulties, randomized studies of occlusion therapy in amblyopia have recently been approved under strictly regulated protocols (see, for example, PEDIG (2003)). Indeed, the results of the MOTAS study have facilitated a randomized trial (ROTAS), carried out by the MOTAS cooperative, which has recently been completed. In the ROTAS study, children were randomized to two treatment groups prescribed six and twelve hours daily occlusion respectively; preliminary analysis of the completed study data indicates that the target occlusion levels for the higher daily dose group were rarely met; see Stewart et al. (2004) for further details.

2.3 *Analysis Strategy*

The plots in Figure 1 indicate that a piecewise linear model of response maybe a sound foundation for the statistical modelling on which we will embark. Below, we develop increasingly sophisticated statistical models from this starting point. We will also investigate the primary issue of interest, the relationship between the occlusion dose and the amount of improvement in visual acuity.

There are two related models that may be used for the investigation of the dose-response relationship. The first model assumes a repeated measures structure; each child in the study undergoes repeated assessment of their visual acuity over a number of clinic visits. We term the data in this form the *absolute-level* data. The second model focuses on an interval-by-interval analysis, and takes as the response as the change in visual acuity between successive clinic visits. Data in this form will be termed the *interval-level* data. The two models can be made equivalent by specific choice of the conditional modelling structures used in the second approach, but the second approach is generally more flexible as it allows more general covariance structures to be incorporated.

Figure 2 gives an indication that increasing dose within an interval yields an increase in visual acuity. However, the nature of the dose-response relationship is unclear.

The principal hypothesis to be tested by the study is that the amount of occlusion dose that a child receives during an interval between successive clinic visits is directly related to their improvement in vision during that interval. However, several patient-specific characteristics are thought to be effect modifiers or potential confounders: age of the child; degree of visual acuity deficit at the start of the study or at the start of any individual interval; type (anisometropic, mixed, strabismic) of amblyopia, and time on study.

It was hypothesized that younger children improved further for the same occlusion dose than older children (see the discussion in Tan et al. (2003)), and that greater improvement was observed in children with greater visual impairment for the same occlusion dose. Little was known about the relationship between occlusion dose effect and amblyopia type. The nature of the dose-covariate-response relationship is potentially complex.

Time spent in the study is also considered as a potential effect modifier, separate from the therapeutic affect of refraction or occlusion treatment, or the developing age of the child whilst in

the study. There may be a degree of adaption to the visual acuity testing procedures that is a potential confounder for treatment, and so including a suitable time in study variable in the model will allow any differential effect of these two outcome modifiers.

In the study, the ODM data are available on a by-dose basis, that is, the ODM records each separate interval during which the child is wearing the monitor. In this paper, we restrict attention to simple functions of the total dose (in minutes) and the length (in days) of the interval between clinic visits.

In the analysis below we present a number of different analyses and compare the inferences made; specifically, we attempt

- a fixed effects analysis of the absolute-level data, with (a) an independent and (b) a stationary covariance structure for the residual errors,
- a robust analysis on the absolute-level data,
- a mixed effect/semiparametric model analysis of the absolute-level data,
- a fixed effect model analysis of the interval-level data,
- a mixed effect model analysis of the interval-level data.

Finally, we implement a Bayesian analysis, and note that this readily facilitates predictive procedures that are not possible in the non-Bayesian setting. Prediction is an important aspect of this analysis, as it allows treatment regimes to be constructed.

All of the analyses presented in the paper were implemented in the statistical package R, freely available from

<http://cran.uk.r-project.org/>

using code written by the author¹.

3 Analysis of the Absolute-Level Data

Let the $N = 87$ patients in the study be indexed by i and the n_i+1 clinic visits by j , so that V_{ij} is the visual acuity for patient i on visit j at day t_{ij} into the study, with $j=0$ the start of refraction or occlusion. The logMAR scale is actually a discrete measurement scale (based on letter-by-letter measurement), but as measurements are recorded to two decimal place accuracy, we shall throughout presume a continuous response. Similarly, let D_{ij} be the (random) occlusion dose (in hours) observed in interval j ; for those patients who enter the study in the refraction phase, but do not require occlusion, $D_{ij} = 0$ is identically zero for all j . Similarly, $D_{ij} = 0$ for the first observation in the occlusion for those children who entered that phase. Let A_{ij} be the child's age in months at the beginning of interval j . Let L_i , P_i and S_i denote the visual acuity at the start of interval, start of phase and start of study respectively, and T_i denote the amblyopia type, for patient i .

In the absolute-level response model, it is not sensible to relate the absolute level of visual acuity at clinic visit j only to the occlusion dose in the preceding interval. Instead, it is preferable to use

¹Code available on request: please email d.stephens@math.mcgill.ca.

a cumulative measure for dose for that individual up to visit j , and to regard cumulative dose as a time-varying covariate. We denote the total cumulative dose by the end of interval j for child i by D_{ij}^c . In section 4, an interval-level analysis is carried out, relating occlusion dose received in a given interval directly to the change in visual acuity during that interval, is carried out.

In this preliminary model-fitting stage, a number of models and model refinements that employ various combinations of the variables identified above were considered. The most complex model that we study here is detailed below; it contains main effects and a limited number of interactions thought likely to be influential. Specifically, the most complex model contains the following main effect covariates:

- Amblyopia type, T ,
- Visual Acuity at Start of Study, S ,
- Age at start of interval, A ,
- Cumulative Occlusion dose, D^c ,
- Time in Refraction, t_R ,
- Time in Occlusion, $t_O = \max\{0, t - t_0\}$, where t_0 represents the start of occlusion, for those children that enter occlusion.

as well as interactions between D^c and A , D^c and S , D^c and t_O , t_O and A , and t_O and S . Higher order interactions were not found to improve the fit of the model.

Below, we report the fit of this model, termed the Full Model, and various submodels under different random error structures. We begin with a Normal linear regression analysis that ignores the repeated measures component. The parameterization used attempts to detect differences between refraction and occlusion phases; for example, a model including Age at Interval, A , is incorporated using a switchpoint contrast parameterization that contributes a term βA in the refraction phase, and a term $(\beta + \delta)A$ in the occlusion phase. Such a parameterization is used for any variable fitted in both phases.

3.1 *A Normal Linear Regression Model*

The model outlined above was fit under the assumption of independent and identically distributed Normal random errors for each of the 684 observations, without any by-individual grouping.

This model had an adjusted R^2 value of 0.918; the estimated residual error variance is 0.021 ($\sigma = 0.144$). A residual plot revealed that, apart from a small number of outlying observations, the model fit was adequate. Inspection of coefficients related to dose in the model appears to confirm conventional ophthalmological wisdom; for cumulative dose, D^c , the interaction between dose and age at intercept, $D^c.A$, and the interaction between dose and time in occlusion, $D^c.t_O$, the estimates (standard errors) are -9.174e-4 (2.528e-4), 1.156e-5 (3.560e-6) and 1.521e-6 (3.413e-7) respectively. Thus it appears that visual acuity improves with increasing cumulative dose, although this improvement is moderated by the age of the child and the time spent in occlusion.

It appears that not all terms are influential in the model; for example, it seems that amblyopia type is not influential in the occlusion phase, but that the Mixed type amblyopes have inferior visual

acuity during refraction. Thus the next stage of analysis is to omit terms from the model that do not improve fit. In doing this, we apply the principle of marginal coherence, that is, any term that appears in a significant interaction is also included as a main effect. We carry out model comparison using a Bayes Information Criterion (BIC).

The BIC is one of a general class of well-established model selection criteria that takes allows the determination of the optimal model from a list of competitors. These criteria take the generic form

$$Criterion = Fidelity\ to\ Data\ +\ Penalty\ for\ Model\ Complexity.$$

Fidelity to the data, or goodness of fit of the model, in parametric statistical analysis is measured using the likelihood, typically

$$-2\log(maximized\ likelihood)$$

where the maximization is over the parameters in the model. This quantity is **large** and **negative** if the model fit is good; for regression models such as those fitted in this paper, the fidelity term often involves a function of a residual sum of squares from the model fit. The penalty for model complexity usually takes the form

$$k(n)p$$

where $k(n)$ is a factor that is potentially dependent on the sample size n , and p is the number of parameters in the fitted model. In the case of the BIC, $k(n)$ takes the form

$$k(n) = \log n.$$

It has been shown that this choice of penalty function leads to good theoretical performance of the model selection criterion. See Sakamoto et al. (1986) for a comprehensive discussion of such model selection criteria.

The BIC values for a number of models nested in the Full Model were compared, and by dropping three terms from the model the BIC value changed from -599.1089 to -622.7042. While the model which omits age from the refractive adaption phase of treatment attains a lower BIC value (-626.4423), we also retain for consideration the model that includes age in both phases of treatment as this model is more coherent and comprehensible. A summary of these final two models are presented in Table 1.

An inspection of residual plots for the selected models proved satisfactory; removing the redundant variables from the model did not adversely affect the fit. Three-way interaction terms $D^c.t_O.A$ and $D^c.t_O.S$ were also tested, but did not decrease the BIC measure further.

3.2 A Multivariate Normal Model

The next stage of the modelling strategy involves acknowledging the repeated measures aspect of the visual acuity data. We assume a multivariate Normal structure for the data, that is, for individual $i = 1, \dots, N$, $V_i \sim N_{n_i+1}(X_i\beta, \Sigma_i)$, where X_i is a $(n_i + 1) \times p$ covariate matrix, β is a $p \times 1$ coefficient vector, and Σ_i captures the covariance in the random observation errors ϵ_i .

The specification of a suitable covariance structure is hindered by the fact that the observation times are not common across children. The parametric form of covariance structure used here is the

stationary exponential decay model, where

$$\text{Cov}[V_{ik}, V_{il}] = \lambda \exp \left\{ - \left| \frac{t_{ik} - t_{il}}{\phi} \right|^\alpha \right\} \quad (1)$$

with common parameters $\lambda, \phi, \alpha > 0$ for all individuals in the study. This model uses time into study, t_{ij} , in the metric to define the degree of covariance. An alternative model replaces time t_{ij} by visit number j in (1), and uses a discrete autoregressive AR(1) model; this model is used for the interval-level data analysis in section 4.

The multivariate Normal log-likelihood (up to an additive constant) is

$$-\frac{1}{2} \sum_{i=1}^N \log |\Sigma_i| - \frac{1}{2} \sum_{i=1}^N (V_i - X_i \beta)^T \Sigma_i^{-1} (V_i - X_i \beta) \quad (2)$$

where the elements of Σ_i are determined by equation (1).

We compare the covariate coefficient parameter estimates and standard errors derived from the multivariate Normal model under maximum likelihood (ML) and restricted maximum likelihood (REML) criteria for the estimation of the observation covariance matrix. Here the two procedures differ in a simple fashion. As usual, ML requires maximization of the log-likelihood function, l , where

$$2l(\Sigma) = -\log |\Sigma| - (v - X\hat{\beta})^T \Sigma^{-1} (v - X\hat{\beta})$$

where $\hat{\beta} = (X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1} v$ and Σ is the block-diagonal matrix $(\Sigma_1, \Sigma_2, \dots, \Sigma_N)$, is a function of (λ, ϕ, α) . In comparison, REML requires maximization of l_{REML} , where

$$2l_{REML}(\Sigma) = 2l(\Sigma) - \log |X^T \Sigma^{-1} X|$$

and is a procedure that is thought to produce estimates that are less biased for small samples. See Diggle et al. (2001), section 4.5, for a discussion. Maximization of l and l_{REML} can be achieved numerically.

We retain the models selected by the BIC comparison approach of the previous section, that is the two models reported in Table 1. The ML and REML estimates of the covariate coefficients and the covariance parameters λ , ϕ and α are presented in Table 2.

The standard errors are slightly larger under REML, but there is no substantive difference in the inference for either model. Again, using the BIC criterion, the model with A omitted from the refraction phase is preferred, and the multivariate models that account for the repeated measures nature of the data are in general much preferred under BIC, as predicted (BIC values -1057.044 and -1050.708 for the two models). A standardized residual plot for the model with A omitted is depicted in figure 3.

3.3 Robust Inference in the Multivariate Model

We can compute robust estimates for β in the multivariate Normal model - see, for example, Diggle et al. (2001), p70. The generalized least-squares estimator for β , denoted $\hat{\beta}_W$, is defined by

$$\hat{\beta}_W = (X^T W X)^{-1} X^T W V = M V, \quad (3)$$

say, where $V = (V_1, V_2, \dots, V_N)^T$ is the stacked observation vector, $X = (X_1, X_2, \dots, X_N)^T$ is the stacked design matrix, and W is a “working” precision matrix reflecting the covariance structure

present in the data. Under this model, the estimator is presumed to have an approximate normal distribution, $\hat{\beta}_W \sim N(\beta, \hat{R}_W)$, where, $\hat{R}_W = M\hat{\Sigma}M^T$, and where $\hat{\Sigma}$ is a consistent estimator of Σ , the block diagonal matrix $(\Sigma_1, \Sigma_2, \dots, \Sigma_N)$. The estimate $\hat{\Sigma}$ was formed as the block diagonal matrix $(\hat{\Sigma}_1, \hat{\Sigma}_2, \dots, \hat{\Sigma}_N)$, with the blocks for each individual defined by $\hat{\Sigma}_i = (v_i - X_i\hat{\beta}_W)(v_i - X_i\hat{\beta}_W)^T$. For our data, the seemingly most appropriate results (after examining residual plots) were obtained with W set equal to the block precision matrix assuming an exponential decay in covariance with scale parameter $\phi = 150$ and $\alpha = 0.6$.

Estimates, estimated standard errors and Wald-type t-statistics were computed under the robust specification. The results of this analysis are omitted here for brevity, but confirmed the results obtained in the model-based analysis reported in previous sections; in particular, the therapeutic effect of occlusion dosing was confirmed.

3.4 A Linear Mixed Effects Model

We now utilize a linear mixed effects (LME) model for the absolute level of visual acuity. We assume that the variation in visual acuity has both a systematic and random component. Specifically, we use a random intercepts model, where there is an individual-specific random intercept at interval zero, that is

$$V_{ij} = X_{ij}^T\beta + Z_i + \epsilon_{ij}, \quad (4)$$

and a presumed autocorrelation in the residual errors ϵ . The model was fitted in R using the `nlme` library.

The results of the REML fit of the mixed effects to the data using the models identified above are presented in table 3. The presumed correlation structure specified in R for the grouped data, and chosen by BIC is an AR(1) structure in visit number; this structure ignores the calendar time between visits, but no improvement was found in the fit when a continuous time model was fitted. When refitted using ML, the models with A omitted and included, gave very similar estimates and standard errors, and had BIC measures -1086.919 and -1080.583 respectively.

Estimates of the random effects and residual standard deviations were similar for the two models: (1.377e-3, 0.147) for the preferred model, (1.407e-3, 0.147) for the model with A included. Again, the pattern of inference is confirmed.

The correlation structures available in the `nlme` package in R are limited. By experimentation, we found that the R continuous AR(1) model, with autocorrelation at lag t equal to ρ^t , and rescaling of the time axis by a factor between 100 and 200, yielded a BIC value around -1010.900. We return to multivariate modelling with random effects in section 5.

3.5 A Semiparametric Additive Mixed Effect Model

Finally in our efforts at modelling the absolute-level data, we fit a semiparametric additive mixed effect (SPAME) model; see, for example Ruppert et al. (2003), and for full details of the inferential procedures, see the Appendix.

In this model, we replace components of the fixed effect models identified above with additive

model components, that is, we model

$$V_{ij} = X_{ij}^T \beta + Z_i + \sum_{k=1}^K f_k(X_{ij}) + \epsilon_{ij}, \quad (5)$$

where the f_k , $k = 1, \dots, K$ are functions of the covariates modelled semiparametrically. We use truncated spline basis functions on the covariates individually; generically, for scalar x varying across a data-dependent range, we specify (fixed but data-dependent) knot positions $\kappa_{k1}, \dots, \kappa_{kM}$, and model function f_k as

$$f_k(x) = \gamma_{k0} + \sum_{m=1}^{M_k} \gamma_{km} (x - \kappa_{km})_+^q \quad (6)$$

where $\gamma_{k0}, \dots, \gamma_{kM}$ are coefficients to be estimated, and $(x - \kappa_{km})_+^q = \max\{0, (x - \kappa_{km})^q\}$. In our analysis, we take $q = 1$, and use the `default.knots` specification of Ngo and Wand (2004) that places up to 35 knots at the covariate quantiles. This model can be fitted in a straightforward fashion using `lme` in R; the function f_k in equation (6) has vector of coefficients γ_k of length M_k , which are assumed to be independent random effects with common variance σ_k^2 . We also assume independence between $\gamma_1, \dots, \gamma_K$, and thus retain a block diagonal structure for the entire random effect matrix. The residual errors are assumed to follow an AR(1) structure, as in section 3.4.

In the analysis, the fixed effect contribution is specified by using two intercepts for the phase of the data; the type of amblyopia in the refraction phase, T ; and the visual acuity at the start of study, S . All other main effects and interactions described in the analysis of section 3.1 are fitted using the semiparametric model. For convenience, the Age at Interval covariate, A , was transformed by location shift to $A - 36$.

First, the preferred model from 3.2 was fitted semiparametric components as outlined above; this yielded a BIC measure of -1060.776 using ML. On inspection of the plots of the eight semi-parametric fits, it was deemed that the interaction between cumulative dose, D^c , and time in occlusion, t_O did not contribute significantly, and was dropped from the model. This yielded a BIC value of -1076.397. On inspection of the estimated fixed effects coefficients, it was deemed that Amblyopia Type could be omitted from the refraction phase model, as could the switch point parameter (occlusion intercept) indicating presence in the occlusion phase. On omitting these terms, the BIC value was -1087.290. The resulting fixed effect model retained only three terms: an intercept, visual acuity at the start of study, S , and a switch point modification for S in the occlusion phase, that is, $X_{ij}^T \beta = \beta_0 + \beta_1 S_{ij} + (\beta_1 + \delta) S_{ij} I(t_{ij} > t_0)$.

As established earlier Age at Interval, A , does not appear to be an influential covariate in refraction, but was initially retained for comparison with the occlusion phase model; the BIC value of the model with A in refraction omitted was -1093.818. In fact, using the semiparametric analysis, it is evident that A only influences visual acuity via an interaction with dose. The six remaining additive model functions are depicted in figure 4. The dependence on cumulative dose D^c is depicted in figure 5, where error bars for the dose effect are also included. Note that the interactions with A and S moderate the dose effect, but that this figure adequately captures the occlusion/improvement dose-response relationship.

4 Modelling the Interval-Level data

We now adopt an alternative modelling strategy. Rather than modelling the entire observation vector using a multivariate model, we attempt to model the vector of first differences of the visual acuity data, as we wish to study the change in visual acuity as a function of occlusion dose and the other covariates. The dependence on covariate values is modelled in a more straightforward fashion. Unlike in previous analyses, the dose in an interval, D , is used as a covariate, rather than cumulative dose.²

In terms of the notation above, let $Y_{ij} = V_{ij} - V_{ij-1}$, for $j = 0, 1, \dots, n_i$, $i = 1, 2, \dots, N$, be the change visual acuity in interval j between visit $j - 1$ and visit j for patient i . In this analysis, we also fit “current” visual acuity, that is, the acuity measured at the $j - 1$ st visit, as a covariate. Thus, we use the following terms in the model:

- Dose during the Interval, D
- Age at start of Interval, A
- Visual Acuity at Start of Phase, P
- Visual Acuity at Start of Interval, L
- Amblyopia Type, T

These variables, and interactions between them were fitted using similar techniques to those used in section 3. Time on study, t , time in refraction, t_R , and time in occlusion, t_O , were also tested as potentially influential covariates, but proved to add little to the fit of the model.

4.1 A Normal Linear Regression Model

First, we fit a Normal linear regression with the covariates listed above. Backwards step-wise regression was implemented using BIC to test whether a more complicated model significantly improved the fit of the model to the data. Using the BIC measure, the models for the two phases that were found to be most appropriate were, for the Refraction phase, $L + P + T$, and for the Occlusion phase, $D + A + L + P + D.A + D.L + A.L$. Time on study, in refraction, or in occlusion was not found to improve the fit in any model.

Table 4 gives parameter estimates for the terms in the final models. The fit of this model yielded an estimate of the residual error standard deviation of $\sigma = 0.0901$, with an R^2 value of 47.0%. This level of measurement error is higher than expected from prior ophthalmological experience. When the regression models were fitted separately to refraction and occlusion phase data, the estimates for σ were similar (0.1112 and 0.0788).

The Normal linear regression of the refraction phase visual acuity measurements suggests that prior to occlusion, the vision of anisometropic children given spectacles decreases on the logMAR scale (i.e. improves) on average by 0.066 (0.036, 0.096) between each visit, that strabismic children exhibit a similar improvement, but that children classified as mixed type do not exhibit significant improvement. The stepwise procedure also reveals that visual acuity at the start of the phase of the study and visual acuity at the start of the interval are both influential in the visual improvement.

²In fact, cumulative dose, D^c , was tested as a covariate, both in conjunction with and in place of D , but did not prove to be an influential term in the model for change in visual acuity.

The Normal linear regression model for the occlusion phase suggests a complex relationship between the degree of improvement and the covariates. The beneficial effect of occlusion is clear, but it is modified by both the age of the child and the visual acuity at the start of the interval. The dose (in hours) main effect coefficient estimate is $-7.630\text{e-}4$ ($3.778\text{e-}4$). Children who were younger, and/or had higher logMAR at the start of occlusion and at the start of an interval all improved further for the same occlusion dose.

4.2 A Linear Mixed Effects Model

Using the covariates identified by the step-wise procedure in the Normal linear regression above to model the mean change in visual acuity, the residuals were explored to help determine a suitable covariance structure for the data. An empirical covariance matrix did not support a random slopes model, as the variance in residuals did not vary greatly from visit to visit.

Successively more complex covariance structures were fit to the data. All models for the refraction and occlusion phases assume random intercepts at the individual child grouping level, as in equation (4). For the occlusion phase, the addition of a random slopes model was also considered. For a single covariate X , a generic random slopes model takes the form

$$Y_{ij} = \beta_0 + (\beta_1 + \xi_i)X_{ij} + \epsilon_{ij}$$

where ξ_i is the random effect contribution of covariate X for individual i over the fixed contribution β_1 . Autoregressive correlation structures were considered for the observed data, as in section 3.4; here, we retain the AR(1) model, where $\text{Corr}[Y_{ij}, Y_{ij'}] = \rho^{|j'-j|}$.

The models considered were as follows:

MODEL 1: Random intercepts with no correlation,

MODEL 2: Random intercepts and slopes with no correlation,

MODEL 3: Random intercepts with AR correlation,

MODEL 4: Random intercepts and slopes with AR correlation.

The BIC measure was used to select the most appropriate model, and it was apparent that MODEL 3 provided a fit to the data that was comparable to models with random intercepts, AR correlation, and random slopes in (i) Dose, (ii) Age at Interval during the occlusion phase, (iii) visual acuity at the start of the refraction phase, or (iv) at the start of the occlusion phase. We conclude that random slopes are not necessary to model the data. However the inclusion of random intercepts and an autoregressive correlation structure does provide a better fit to the data than a Normal linear model alone. Table 5 gives parameter estimates for the terms in the final models using REML. The fit of this model yielded estimates of the residual error standard deviation and the correlation of $\sigma = 0.0735$ and $\rho = -0.1708$.

The parameter estimates from the mixed effect model are similar to those observed in the Normal linear regression. In the refraction phase visual acuity measurements suggests that prior to occlusion, the vision of anisometric children given spectacles decreases on the logMAR scale (i.e. improves) on average by 0.083 (0.012,0.154) between each visit. Strabismic children exhibit a lesser degree of

improvement while children of mixed type do not exhibit significant improvement. As in the Normal linear analysis, we see that children who were younger, and/or had higher logMAR at the start of occlusion and at the start of an interval all improved further for the same occlusion dose.

5 Bayesian Inference and Prediction

The results of the likelihood-based analysis above have identified the key predictors in the model for changes in visual acuity. However, the interactions between the covariates render the selected model difficult to interpret. Therefore, to study the impact of different dosing strategies, we now implement a Bayesian analysis, as this propagates the uncertainty in the inference in a fully coherent fashion.

In the case of the fixed effects model, under the assumption of independent residual errors, computation of the posterior quantities of interest can proceed analytically in both the absolute-level and interval-level data. We focus on the correlated response model of section 3.2 for absolute-level data and perform inference using Markov chain Monte Carlo. We note that, in fact, any of the models fitted in earlier sections could be implemented in a Bayesian MCMC setting; the most complex model, the SPAME model of section 3.5 is richly parameterized but is still routinely implementable using MCMC.

5.1 Bayesian Inference for the Absolute-Level Data

For illustration, we implement a Bayesian analysis for the model in the multivariate Normal model with the improper “Jeffreys” prior on the positive parameters in the autocovariance function (1), that is, we take $p(\beta, \lambda, \phi, \alpha) = (\lambda\phi\alpha)^{-1}$, and derive the posterior distribution which takes the form

$$p(\beta, \lambda, \phi, \alpha | V) = p(\lambda, \phi, \alpha | V) p(\beta | V, \lambda, \phi, \alpha)$$

where

$$p(\lambda, \phi, \alpha | V) \propto \frac{|M_3|^{-1/2}}{\prod_{i=1}^N |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} [M_1 - M_2^T M_3^{-1} M_2] \right\} \frac{1}{\lambda\phi\alpha}$$

where

$$M_1 = \sum_{i=1}^N V_i^T \Sigma_i^{-1} V_i \quad M_2 = \sum_{i=1}^N X_i^T \Sigma_i^{-1} V_i \quad M_3 = \sum_{i=1}^N X_i^T \Sigma_i^{-1} X_i$$

and $\beta | V, \lambda, \phi, \alpha \sim N_p(M_3^{-1} M_2, M_3)$. The posterior distribution is not available analytically, and neither are the posterior marginal distributions, but inference may be carried out using Markov chain Monte Carlo (MCMC) on the three parameter joint posterior. We use a Metropolis update on a sweep of the conditionals, reparameterized onto the log scale, and jointly on the block of the three parameters. The conditional posterior for β given (λ, ϕ, α) can be sampled directly.

For the covariate model, to specify the design matrices, we fitted the model from section 3.2, with Age at Interval included in the model; this term is included for the purposes of the prediction for different dosing strategies.

Fitting a random effects model, similar to that in equation (4), is straightforward using a Gibbs sampler. Denoting by $Z = (Z_1, \dots, Z_N)$ the vector of child-specific random effects, the posterior of interest becomes the joint distribution $p(\beta, \lambda, \phi, \alpha, Z, \sigma_Z^2 | V)$, where σ_Z^2 is the (unknown) random

effect error variance, which is included in the MCMC cycle; we assign an Inverse Gamma prior with parameters 2.5 and 0.25. Then

- Conditional on Z , the posterior for $(\beta, \lambda, \phi, \alpha)$ is updated as in the fixed effect only model, with data V_{ij} replaced by $V_{ij} - Z_i$.
- Conditional on $(\beta, \lambda, \phi, \alpha)$ and σ_Z^2 , the posterior for Z_i is univariate normal.
- Conditional on all other parameters, the posterior for σ_Z^2 is Inverse Gamma.

Full details of the conditional distributions are omitted here. The MCMC code was written in R, and in version 2.0.1 took around 10 minutes to implement on a 1.6GHz machine. After appropriate construction of an independence Metropolis kernel for the parameters in the autocovariance model using short pilot runs, it was possible to produce samples from the marginal posterior distribution for parameters $(\beta, \lambda, \phi, \alpha)$ that were virtually uncorrelated serially. The numerical summaries of the marginal posterior distribution for each of the parameters are given in table 6.

The sample median and 95% posterior credible intervals for the residual error standard deviation in the two analyses were similar; 0.147 (0.135,0.161) for fixed effects model, 0.135 (0.121,0.155) for the random effects. The posterior samples from the fixed-effect analysis will be used later in section 5.3 to examine the impact of different dosing strategies.

5.2 Bayesian Inference for the Interval-Level Data

A Bayesian analysis of the interval level data is again facilitated using MCMC for the multivariate response data, with or without random effects. In this analysis, again, a range of autocovariance structures were tested; for illustration, we present the results for a model where $\text{Corr}(Y_{ij}, Y_{ij'}) = \rho^{|j' - j|}$, as in section 4.2.

The earlier detected pattern of inference is confirmed. The posterior median estimate (credible interval) for the residual error standard deviation is 0.091 (0.086,0.096), similar to the result in section 4.2. Full results are omitted here.

5.3 Bayesian Prediction: The Impact of Different Dosing Strategies

The Bayesian analysis of the previous sections is useful as it readily facilitates a study of the impact of different dosing strategies. This is especially the case when sampling-based analyses are implemented. Once a large sample from the posterior distribution has been obtained, an equal-sized sample from the posterior predictive distribution can be obtained for fixed values of the modelled covariates by sampling from the likelihood part of the model conditional on each element of the posterior sample in turn. Forward sampling from the model in this way is a more interpretable method for assessing the potential worth of occlusion dosing strategies.

Here, for illustration, we examine the impact of different dosing strategies on two hypothetical children, one aged 48 months at the start of study, the other aged 72 months. Each child will enter the study with a visual acuity of 1.0 logMAR, and be diagnosed as an anisometrope. They will spend four months in refraction, then be switched to occlusion. In occlusion, they could receive 2, 6 and 9 hours of occlusion per day, over a 30 day month period, and then be followed for a further six visits, when the study will end.

The impact of the different dosing strategies for the two children is depicted in figure 6, where the posterior predictive median response profiles are plotted, where the posterior samples are generated in the Bayesian analysis of the absolute-level data in section 5.1. The difference between the response profiles under different strategies is evident (the expected profile under a zero dose strategy is included for comparison). Also, comparing the results for the two children, the impact of age is striking, with the younger child improving to a greater degree than the older child under each strategy. This analysis is illustrative; the impact for children with varying characteristics can readily be studied in the same way.

Acknowledgements: The statistical analysis in this paper was carried out whilst the author was Visting Scholar at the Department of Statistics, University of Washington, Seattle, with the assistance of Dr Erica Moodie, Department of Biostatistics, University of Washington. The author was partly supported by funding from The Guide Dogs for the Blind Association, UK, who funded the MOTAS study.

Appendix A Details of the Semiparametric Model Fitting Procedure

This section describes how to fit a semiparametric additive mixed effect model that allows flexible modelling of the response function. Following Ruppert et al. (2003), we use the linear mixed model formulation,

$$Y = X\beta + Zu + \epsilon \quad (\text{A-1})$$

where

$$E \begin{bmatrix} u \\ \epsilon \end{bmatrix} = \mathbf{0} \quad \text{Var}[\theta] = \begin{pmatrix} G & 0 \\ 0 & R \end{pmatrix}$$

where the matrix X contains the fixed effects predictors, matrix Z is the (basis function) design matrix in the semiparametric representation of the function of $f_1, \dots, f_K$. We give brief details in the following sections.

A.1 Inference for the Linear Mixed Model

Inference for the linear mixed model in equation (A-1) is achieved using penalized least-squares³, or likelihood procedures under a (model-based) assumption of Gaussianity; we give this version for ease of interpretation. Suppose, in conjunction with equation A-1, $u \sim N(0, G)$ and $\epsilon \sim N(0, R)$ with u and ϵ independent. This model can be interpreted as $Y|\beta, u \sim N(X\beta + Zu, R)$, $u \sim N(0, G)$, yielding (on integrating out u) the marginal model $Y|\beta \sim N(X\beta, ZGZ^\top + R)$. Let $V = ZGZ^\top + R$. Then the maximum penalized likelihood estimates of β and u given G and R are given by

$$\hat{\theta} = \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = (C^\top R^{-1}C + B)^{-1}C^\top R^{-1}y \quad (\text{A-2})$$

where $C = [X \ Z]$ and B is the block diagonal matrix with blocks $\mathbf{0}$ and G^{-1} . The variance of the estimators are given by $\text{Var}[\hat{\theta}] = (C^\top R^{-1}C + B)^{-1}$. Fitted values are obtained routinely as

$$\hat{y} = C(C^\top R^{-1}C + B)^{-1}C^\top R^{-1}y$$

³The results are also justified under the (model-free) paradigm of Best Linear Unbiased Prediction (BLUP).

whereas predictions under this model at new design point $c_0 = [x_0 \ z_0]$ are obtained from

$$\hat{y}_0 = c_0(C^\top R^{-1}C + B)^{-1}C^\top R^{-1}y$$

with variance $c_0(C^\top R^{-1}C + B)^{-1}c_0^\top$. The quantities R and G that together define V and B can be estimated using maximum profile (integrated) likelihood

$$l_P(V) = \text{constant} - \frac{1}{2} \left[\log |V| + y^\top V^{-1}(I - X(X^\top V^{-1}X)^{-1}X^\top V^{-1})y \right]$$

obtained from the likelihood plugging in $\hat{\beta} = (X^\top V^{-1}X)^{-1}X^\top V^{-1}y$, or REML, using the restricted likelihood

$$l_R(V) = l_P(V) - \frac{1}{2} \log |X^\top V^{-1}X|.$$

obtained by first integrating out β from the likelihood $Y \sim N(X\beta, V)$.

This model has a (model-based) Bayesian interpretation where the unknown parameters β and u are assigned independent prior distributions, with β having an improper uniform prior on $\mathbb{R}^{\text{ncol}(X)}$, and u assigned the Gaussian prior described above. In a fully Bayesian approach, G is set as a fixed hyperparameter, or assigned an informative prior distribution. Here, an empirical Bayes approach is used where G and the parameters in R are estimated using ML/REML.

A.2 Model Selection

Model selection can be carried out using BIC for linear mixed models, but an adjustment to the BIC calculation must be made. The conventional BIC is defined by

$$BIC = -2\hat{l} + (\text{Number of Parameters} \times \log(\text{Sample size}))$$

where $\hat{l}$ is the maximized log-likelihood. In a standard linear regression problem, the number of parameters fitted is given by the trace of the “hat” or smoother matrix, $\mathbb{S}$, that projects the data onto the fitted values, and that convention is followed here; we have

$$\hat{y} = C(C^\top R^{-1}C + B)^{-1}C^\top R^{-1}y = \mathbb{S}y,$$

say, giving that the number of parameters is $\text{tr}(\mathbb{S}) = \text{tr}(C(C^\top R^{-1}C + B)^{-1}C^\top R^{-1})$.

A.3 Specification of the Semiparametric Design: The Truncated Spline Basis

In the semiparametric additive model, the matrix Z contains the truncated spline basis terms, with columns corresponding to knots $\kappa_{k1}, \dots, \kappa_{kM}$, for $k = 1, \dots, K$. Typically, the random effects coefficients for function k are assigned a common Gaussian distribution, so that the matrix G is diagonal; however, this is not necessary - a more informative prior may be chosen.

We use truncated spline basis functions to construct the semiparametric specification. Generically, for scalar x varying across a data-dependent range, we specify (fixed but data-dependent) knot positions $\kappa_{k1}, \dots, \kappa_{kM}$, and model function f_k as

$$f_k(x) = \sum_{m=1}^M u_{km}(x - \kappa_{km})_+^q \tag{A-3}$$

where $u_{k1}, \dots, u_{kM}$ are (random effects) coefficients for function k , and the basis function component $(x - \kappa_{km})_+^q = \max\{0, (x - \kappa_{km})^q\}$, so that a typical row of Z (an $N \times KM$ matrix) in equation (A-1) takes the form

$$[(x - \kappa_{11})_+^q \quad (x - \kappa_{1M})_+^q \quad \dots \quad (x - \kappa_{KM})_+^q].$$

In our analysis, we take $q = 1$, and use ten knots at the covariate quantiles, with a knot also placed at zero, giving $M = 11$. For convenience, we transform (by translation) the covariates such that they are non-negative. The function f_k in equation (6) has vector of coefficients $\mathbf{u}_k$ of length M , which are assumed to be independent random effects with common variance σ_k^2 , $k = 1 \dots, K$. We also assume independence between $\mathbf{u}_1, \dots, \mathbf{u}_K$, and thus retain a block diagonal structure for the entire random effect matrix.

The semiparametric model can be fit in a straightforward fashion using `lme` in R for certain choices of the residual error covariance R , and more generally using numerical procedures for general covariance specifications. For the analysis in this paper we wrote code in R to perform the ML/REML estimation.

References

- Diggle, P. J., P. Heagerty, K.-Y. Liang, and S. L. Zeger (2001). *Analysis of Longitudinal Data* (2 ed.). Oxford University Press.
- Fielder, A. R., R. Auld, M. Irwin, K. D. Cocker, H. S. Jones, and M. J. Moseley (1994). Compliance monitoring in amblyopia therapy. *Lancet*, 343–547.
- Lavori, P. W. (2000). A design for testing clinical strategies: biased adaptive within-subject randomization. *Journal of the Royal Statistical Society, Series A* 163, 29–38.
- Lavori, P. W. and R. Dawson (2004). Dynamic treatment regimes: practical design considerations. *Clinical Trials* 1, 9–20.
- Moodie, E. E. M., T. S. Richardson, and D. A. Stephens (2005). Demystifying optimal dynamic treatment regimes. UW Biostatistics Working Paper 240, University of Washington, <http://www.bepress.com/uwbiostat/paper240>.
- Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society, Series B* 65, 331–366.
- Murphy, S. A. (2004). An experimental design for the development of adaptive treatment strategies. *Statistics in Medicine*. In Press.
- Ngo, L. and M. P. Wand (2004). Smoothing with mixed model software. *Journal of Statistical Software* 9(1).
- PEDIG (2003). A randomized trial of patching regimens for treatment of moderate amblyopia in children. *Archives of Ophthalmology* 121(5), 603–611.
- Rahi, J. S., S. Logan, C. Timms, I. Russell-Eggitt, and D. Taylor (2002). Risk, causes and outcomes of visual impairment after loss of vision in the non-amblyopic eye: a population-based study. *Lancet* 360(2), 597–602.

- Robins, J. R. (2004). Optimal structural nested models for optimal sequential decisions. In D. Y. Lin and P. Heagerty (Eds.), *Proceedings of the Second Seattle Symposium on Biostatistics*, New York. Springer.
- Ruppert, D., M. P. Wand, and R. J. Carroll (2003). *Semiparametric Regression*. Cambridge University Press.
- Sakamoto, Y., M. Ishiguro, and G. Kitagawa (1986). *Akaike Information Criterion Statistics*. D. Reidel Publishing Company.
- Stewart, C. E., A. R. Fielder, D. A. Stephens, and M. J. Moseley (2002). Design of the monitored occlusion treatment of amblyopia study (MOTAS). *British Journal of Ophthalmology* 86, 915–919.
- Stewart, C. E., M. J. Moseley, A. R. Fielder, and D. A. Stephens (2004). Optimization of the dose-response of occlusion therapy for amblyopia: the ROTAS study. *Investigative Ophthalmology and Visual Science (suppl.)* 45, 2579.
- Stewart, C. E., M. J. Moseley, D. A. Stephens, and A. R. Fielder (2004). Treatment dose-response in amblyopia therapy: the monitored occlusion treatment of amblyopia study (MOTAS). *Investigations in Ophthalmology and Visual Science* 45, 3048–3054.
- Tan, J. H. Y., J. R. Thompson, and I. Gottlob (2003). Differences in the management of amblyopia between European countries. *British Journal of Ophthalmology* 87, 291–296.

		Model with A omitted from Ref.				Model with A included in Ref.			
Phase	Term	Est.	s.e.	t.stat	p	Est.	s.e.	t.stat	p
Ref.	Int.	-2.057e-2	2.007e-2	-1.025	0.306	-8.616e-2	4.442e-2	-1.940	0.053
	t_R	-8.306e-4	7.528e-5	-11.033	0.000	-8.315e-4	7.519e-5	-11.059	0.000
	T_M	5.507e-2	1.603e-2	3.436	0.001	5.528e-2	1.601e-2	3.454	0.001
	T_S	2.200e-2	1.626e-2	1.353	0.177	2.714e-2	1.653e-2	1.641	0.117
	S	8.705e-1	2.424e-2	35.910	0.000	8.782e-1	2.466e-2	35.620	0.000
	A					9.450e-4	5.711e-4	1.655	0.098
Occ.	Occ.	-2.208e-1	5.679e-2	-3.887	0.000	-1.588e-1	6.796e-2	-2.336	0.020
	t_O	1.458e-4	2.268e-4	0.643	0.521	1.423e-4	2.265e-4	0.628	0.530
	S	-1.194e-1	3.924e-2	-3.044	0.002	-1.262e-1	3.940e-2	-3.204	0.001
	A	2.719e-3	6.937e-4	3.920	0.000	1.796e-3	8.896e-4	2.019	0.039
	D^c	-9.975e-4	1.756e-4	-5.681	0.000	-1.001e-3	1.754e-4	-5.709	0.000
	$D^c.A$	1.121e-5	2.648e-6	4.235	0.000	1.125e-5	2.645e-6	4.255	0.000
	$D^c.t_O$	1.517e-6	3.351e-7	4.526	0.000	1.528e-6	3.347e-7	4.566	0.000
	$t_O.S$	-1.780e-3	2.597e-4	-6.856	0.000	-1.781e-3	2.594e-4	-6.867	0.000
		Adj. R^2 : 0.918 $\hat{\sigma} = 0.145$				Adj. R^2 : 0.918 $\hat{\sigma} = 0.145$			

Table 1: Summary of model fits to the absolute-level data for models selected using BIC assuming independent residual errors.

		ML				REML			
Phase	Term	Est.	s.e.	t.stat	p	Est.	s.e.	t.stat	p
Ref.	Int.	-1.030e-2	6.601e-2	-0.156	0.876	-8.021e-2	6.832e-2	-0.117	0.907
	t_R	-1.046e-3	1.034e-4	-10.084	0.000	-1.052e-3	1.060e-4	-9.919	0.000
	t_O	7.903e-5	2.482e-4	0.318	0.750	7.551e-5	2.525e-4	0.299	0.765
	T_M	5.150e-2	3.117e-2	1.652	0.098	5.121e-2	3.239e-2	1.581	0.114
	T_S	9.699e-3	3.177e-2	0.305	0.760	9.329e-3	3.230e-2	0.283	0.777
	S	8.783e-1	3.755e-2	23.394	0.000	8.783e-1	3.889e-2	22.587	0.000
	A	3.594e-4	8.159e-4	0.440	0.660	3.384e-4	8.441e-4	0.401	0.688
Occ.	Occ.	-1.543e-1	5.688e-2	-2.713	0.007	-1.544e-1	5.733e-2	-2.694	0.007
	S	-6.520e-2	3.106e-2	-2.099	0.036	-6.418e-1	3.128e-2	-2.051	0.040
	A	1.656e-3	7.202e-4	2.280	0.021	1.655e-3	7.258e-4	2.280	0.023
	D^c	-1.343e-3	2.169e-4	-6.194	0.000	-1.348e-3	2.206e-4	-6.112	0.000
	$D^c.A$	1.126e-5	3.160e-6	3.979	0.000	1.261e-5	3.220e-6	3.918	0.000
	$D^c.t_O$	1.941e-6	3.402e-7	5.692	0.000	1.950e-6	3.524e-7	5.622	0.000
	$t_O.S$	-1.738e-3	2.870e-4	-6.057	0.000	-1.734e-3	2.920e-4	-5.940	0.000
Covariance Parameters		$\hat{\lambda} = 0.020$ $\hat{\phi} = 157.748$ $\hat{\alpha} = 0.617$				$\hat{\lambda} = 0.021$ $\hat{\phi} = 170.904$ $\hat{\alpha} = 0.618$			

Table 2: Estimates from the ML and REML model fit of the model with A included to the absolute-level data assuming residual errors have a stationary covariance structure; t.stat is the Wald test statistic, p is the corresponding p -value.

		A Omitted				A Included			
Phase	Term	Est.	s.e.	t.stat	p	Est.	s.e.	t.stat	p
Ref.	Int.	3.32e-2	3.121e-2	1.064	0.288	8.872e-3	6.811e-2	0.130	0.896
	t_R	-9.996e-4	9.914e-5	-10.083	0.000	-1.006e-3	1.005e-4	-10.015	0.000
	t_O	4.512e-5	2.614e-4	0.172	0.863	4.011e-5	2.619e-4	0.153	0.879
	T_M	4.197e-2	3.149e-2	1.332	0.186	4.204e-2	3.154e-2	1.336	0.186
	T_S	1.183e-2	3.127e-2	0.378	0.706	1.426e-2	3.190e-2	0.447	0.656
	S	8.428e-1	3.819e-2	22.070	0.000	8.464e-1	3.930e-2	21.539	0.000
	A					3.416e-4	8.476e-4	0.403	0.687
Occ.	Occ.	-1.790e-1	6.139e-2	-2.916	0.004	-1.725e-1	6.348e-2	-2.718	0.007
	S	-2.592e-2	3.460e-2	-0.749	0.454	-2.691e-2	3.472e-2	-0.775	0.439
	A	1.682e-3	7.630e-4	2.203	0.028	1.590e-3	7.954e-4	2.000	0.046
	D^c	-1.325e-3	2.316e-4	-5.721	0.000	-1.302e-3	2.392e-4	-5.446	0.000
	$D^c.A$	1.213e-5	3.362e-6	3.610	0.000	1.183e-5	3.490e-6	3.367	0.001
	$D^c.t_O$	1.742e-6	3.617e-7	4.613	0.000	1.736e-6	3.643e-7	4.625	0.000
	$t_O.S$	-1.433e-3	3.118e-4	-4.594	0.000	-1.435e-3	3.122e-4	-4.596	0.000

Table 3: Estimates from the LME model fit to the absolute-level data using REML; t.stat is the Wald test statistic, p is the corresponding p -value

Phase	Term	Est.	s.e.	t.stat	<i>p</i>
Refraction	Intercept	-6.573e-2	1.539e-2	-4.271	0.000
	<i>L</i>	-3.571e-1	4.231e-2	-8.441	0.000
	<i>P</i>	2.358e-1	4.330e-2	5.445	0.000
	<i>T_M</i>	6.066e-2	1.724e-2	3.518	0.000
	<i>T_S</i>	6.569e-3	1.755e-2	0.374	0.708
Occlusion	Intercept	-1.229e-3	3.500e-2	-0.035	0.972
	<i>D</i>	-7.630e-4	3.778e-4	-2.019	0.044
	<i>A</i>	-6.041e-4	5.121e-4	-1.180	0.239
	<i>L</i>	-5.231e-1	6.209e-2	-8.425	0.000
	<i>P</i>	1.146e-1	2.209e-2	5.189	0.000
	<i>D.A</i>	1.113e-5	5.930e-6	1.878	0.061
	<i>D.L</i>	5.930e-6	3.472e-4	-2.586	0.010
	<i>A.L</i>	5.247e-3	9.724e-4	5.396	0.000

Table 4: Estimates and standard errors for the parameters in the joint model for refraction and occlusion phases from a Normal linear model fit to the interval level data

Phase	Term	Est.	s.e.	t.stat	<i>p</i>
Refraction	Intercept	-8.509e-2	3.149e-2	-2.702	0.007
	<i>L</i>	-6.376e-1	4.145e-2	-15.383	0.000
	<i>P</i>	4.839e-1	5.405e-2	8.952	0.000
	<i>T_M</i>	8.313e-2	3.640e-2	2.284	0.023
	<i>T_S</i>	1.682e-2	3.818e-2	0.441	0.660
Occlusion	Intercept	-5.728e-3	3.473e-2	-0.165	0.869
	<i>D</i>	-8.645e-4	3.226e-4	-2.680	0.008
	<i>A</i>	-5.738e-4	4.920e-4	-1.166	0.244
	<i>L</i>	-5.113e-1	5.557e-2	-9.237	0.000
	<i>P</i>	1.234e-1	2.144e-2	5.755	0.000
	<i>D.A</i>	1.339e-5	5.063e-6	2.645	0.008
	<i>D.L</i>	-9.705e-4	2.959e-4	-3.279	0.001
	<i>A.L</i>	4.912e-3	9.242e-4	5.315	0.000

Table 5: Estimates and standard errors for the parameters in the joint model for refraction and occlusion phases from a mixed effects model with random intercepts and AR correlation.

				Quantiles		
Phase	Term	Mean	S.D.	0.025	0.500	0.975
Ref.	Int.	3.812e-3	9.330e-2	-1.764e-1	4.885e-3	1.847e-1
	t_R	-1.064e-3	1.124e-4	-1.286e-3	-1.062e-3	-8.460e-4
	t_O	1.071e-4	2.548e-4	-3.974e-4	1.120e-4	5.958e-4
	T_M	5.566e-2	4.783e-2	-3.848e-2	5.572e-2	1.474e-1
	T_S	7.216e-3	4.836e-2	-8.763e-2	7.330e-3	1.018e-1
	S	8.747e-1	5.308e-2	7.716e-1	8.747e-1	9.768e-1
	A	1.411e-4	1.137e-3	-2.087e-3	1.275e-4	2.344e-3
Occ.	Occ.	-1.527e-1	5.798e-2	-2.664e-1	-1.526e-1	-3.996e-2
	S	-6.793e-2	3.167e-2	-1.298e-1	-6.779e-2	-6.161e-3
	A	1.697e-3	7.438e-4	2.284e-4	1.706e-3	3.154e-3
	D^c	-1.327e-3	2.249e-4	-1.767e-3	-1.322e-3	-8.838e-4
	$D^c.A$	1.212e-5	3.237e-6	5.825e-6	1.207e-5	1.851e-5
	$D^c.t_O$	1.967e-6	3.510e-7	1.282e-6	1.966e-6	2.655e-6
	$D^c.t_O$	-1.765e-3	2.897e-4	-2.324e-3	-1.762e-3	-1.203e-3

Table 6: Posterior Summaries for the parameters in the Bayesian analysis of the random effect model: 5000 posterior samples collected after a burn-in of 1000 iterations.

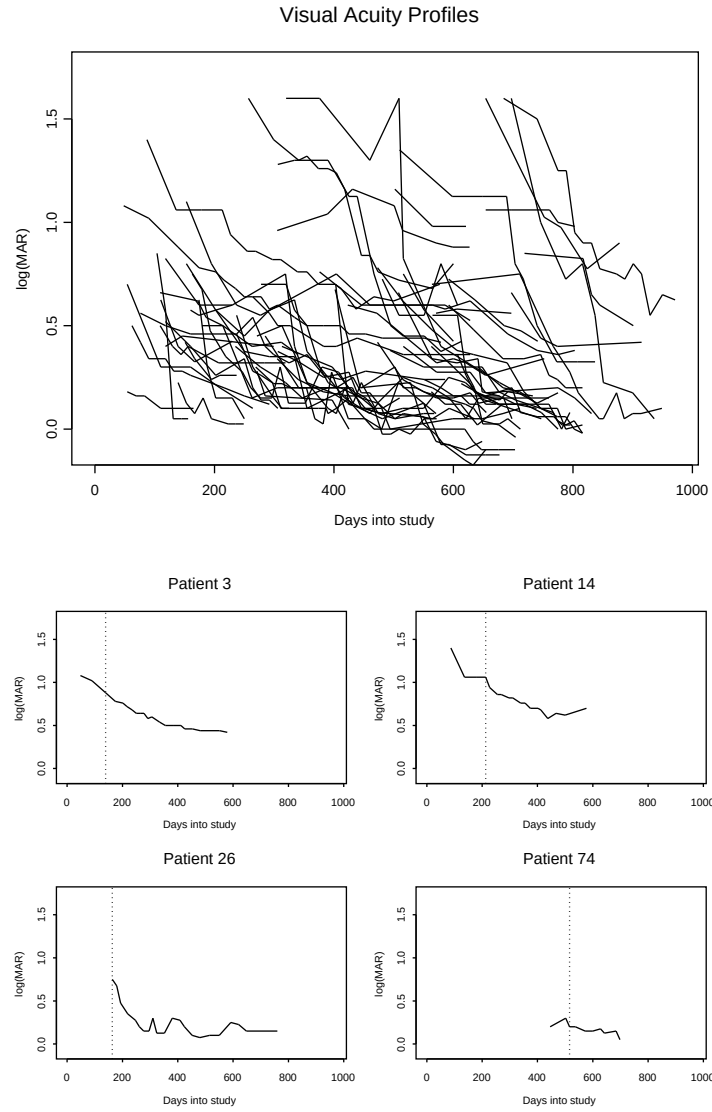


Figure 1: Profile plots for the individuals in the MOTAS study (top) and for four selected patients, with the start of occlusion indicated by the dotted line (bottom).

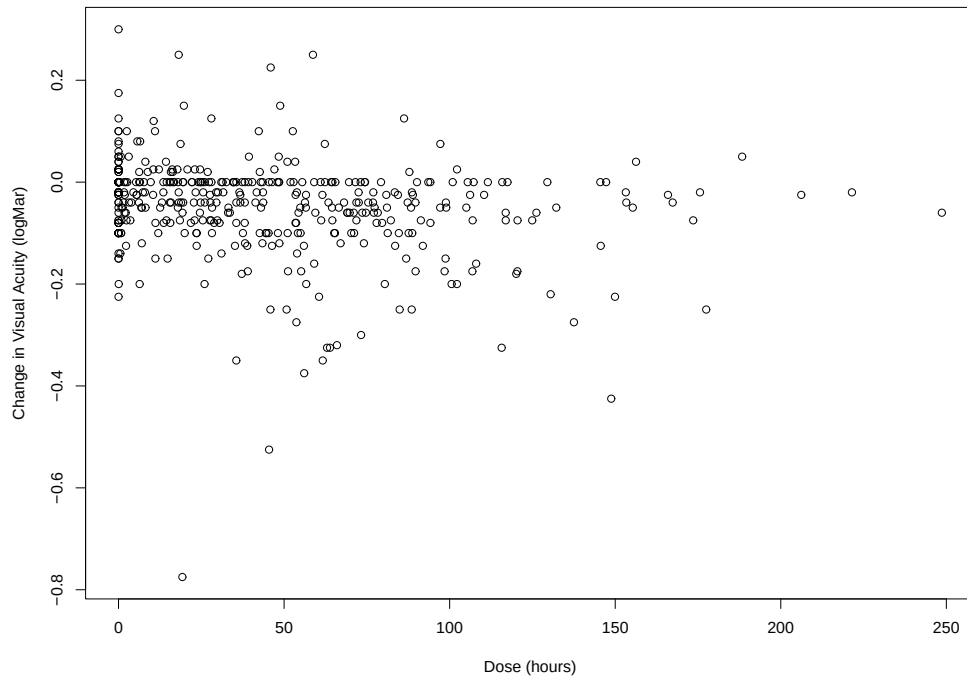


Figure 2: Interval-by-interval change in visual acuity plotted against total interval occlusion dose. One datum omitted with total dose of 439 hours in a 63 day interval.

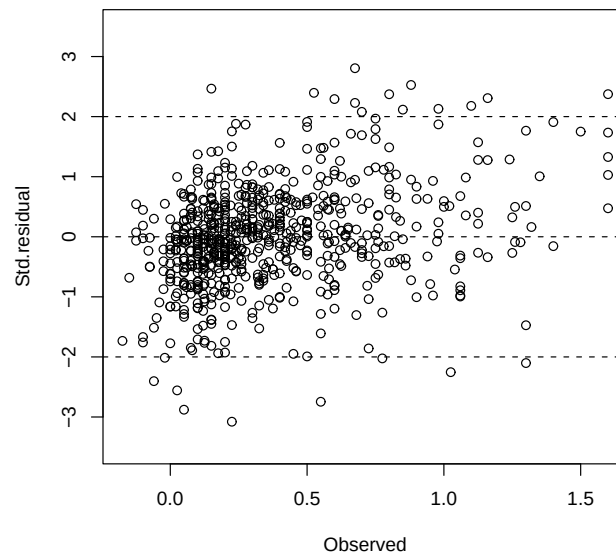


Figure 3: Standardized residual plot for the multivariate model of section 3.2. Horizontal dotted lines are at zero and plus or minus two. One outlying point (1.60,10.61) omitted.

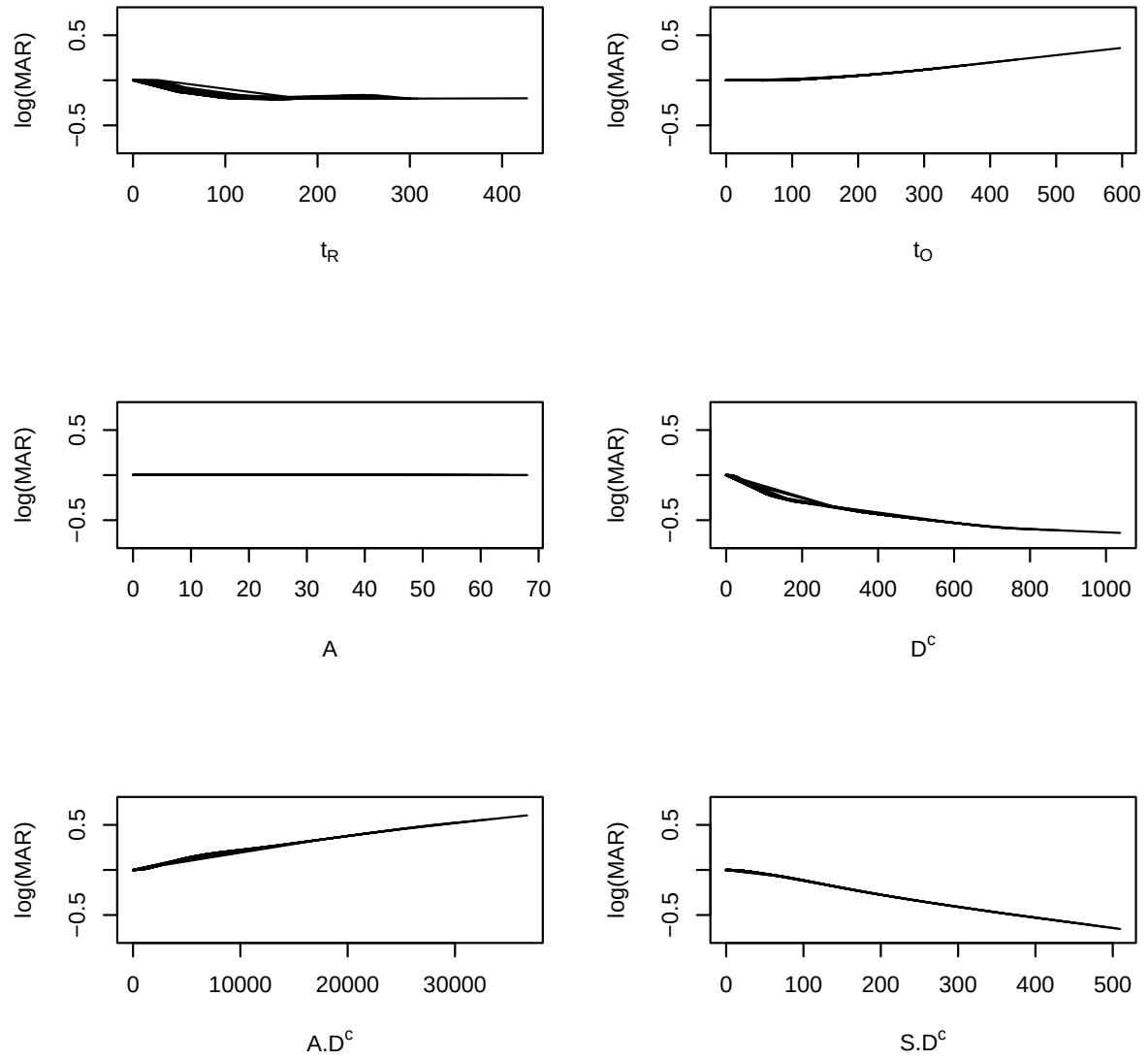


Figure 4: Plots of the additive model fits for the six terms in the SPAME analysis from section 3.5. Estimated response functions for all children in study are plotted.

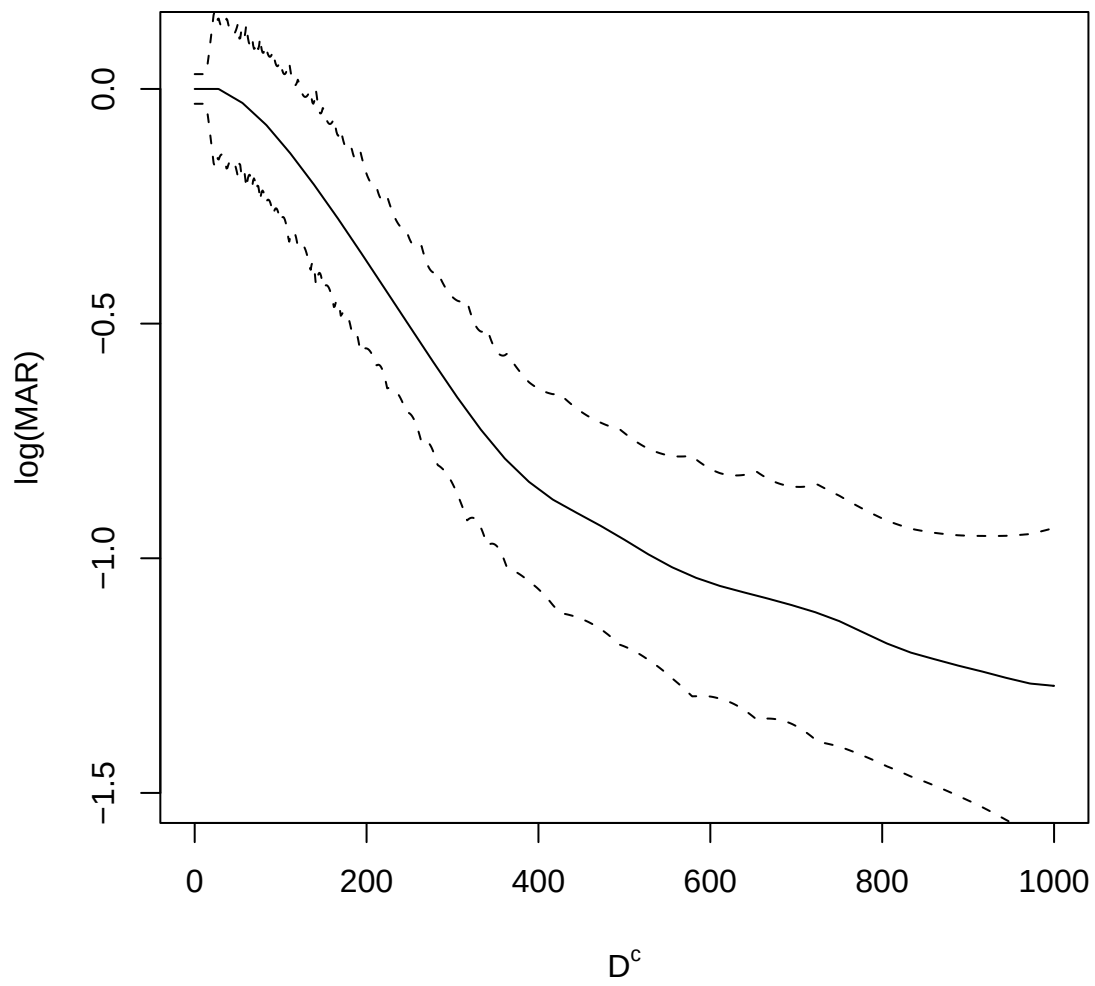


Figure 5: The estimated (marginal) effect of cumulative dose, with 95% confidence bands, from the SPAME analysis from section 3.5

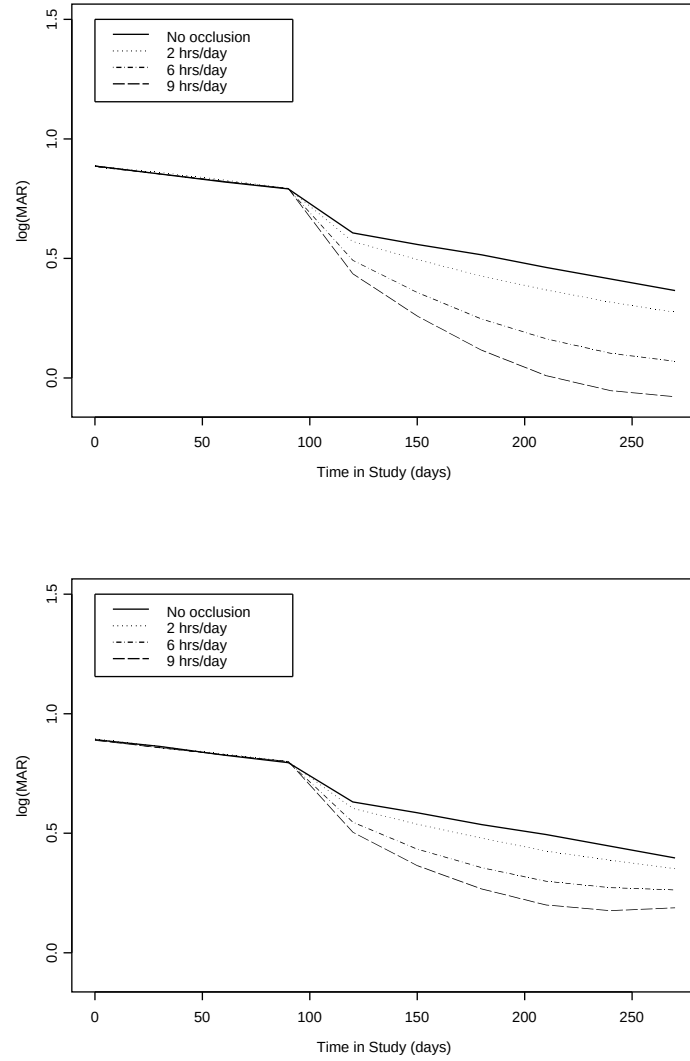


Figure 6: Predicted response for different occlusion dosing strategies in the model for longitudinal response. Top panel: Profiles for an anisometropic child aged 48 months at the start of study. Bottom panel: profiles for an anisometropic child aged 60 months at the start of study.