DIGITALCOMMONS @WAYNESTATE

Journal of Modern Applied Statistical Methods

Volume 16 | Issue 1 Article 37

5-1-2017

JMASM45: A Computer Program for Bayesian D-Optimal Binary Repeated Measurements Designs (Matlab)

Haftom Temesgen Abebe
Mekelle University, Ethiopia, haftoma@gmail.com

Frans E. S. Tan
Maastricht University, The Netherlands, frans.tan@maastrichtuniversity.nl

Gerard J. P. van Breukelen
Maastricht University, The Netherlands, gerard.vbreukelen@maastrichtuniversity.nl

Martijn P. F. Berger
Maastricht University, The Netherlands, martijn.berger@maastrichtuniversity.nl

Recommended Citation

Abebe, H. T., Tan, F. E. S., van Breukelen, G. J. P., & Berger, M. P. F. (2017). JMASM45: A computer program for Bayesian D-optimal binary repeated measurements designs (Matlab). Journal of Modern Applied Statistical Methods, 16(1), 689-721. doi: 10.22237/jmasm/1493599020

Journal of Modern Applied Statistical Methods
May 2017, Vol. 16, No. 1, 689-721.
doi: 10.22237/jmasm/1493599020

ISSN 1538 − 9472

JMASM45: A Computer Program for Bayesian D-Optimal Binary Repeated Measurements Designs (Matlab)

Haftom T. Abebe
Mekelle University
Mekelle, Ethiopia

Frans E. S. Tan
Maastricht University
Maastricht, The Netherlands

Gerard J. P. van Breukelen
Maastricht University
Maastricht, The Netherlands

Martijn P. F. Berger
Maastricht University
Maastricht, The Netherlands

Planners of longitudinal studies of binary responses in applied sciences have not yet benefitted from optimal designs, which have been shown to improve precision of model parameter estimates, due to absence of a computer program. An interactive computer program for Bayesian optimal binary repeated measurements designs is presented for this purpose.

Keywords: Bayesian optimal designs, logistic mixed effects models, subject-to-measurement cost ratio, relative efficiency, number of time points, autocorrelation

Introduction

Longitudinal study designs are used in different disciplines of science to study the change of a particular outcome variable over time. In smoking prevention studies, for example, pupils in primary and secondary school may be followed up to study the prevalence of smoking as a function of age. The generalized linear mixed model (GLMM) is the most frequently used model for the analysis of longitudinal dichotomous data such as smoking status. Optimal design of longitudinal studies has been shown useful to improve the precision of the model parameter estimates of interest, such as the rate of change, by optimizing the number and timing of repeated measurements. For cross-sectional data, the review of McClelland (1997) provided a good introduction into optimal design for psychologists. Raudenbush and Feng (2001) considered a study with a quantitative outcome in

Haftom T. Abebe is in the Department of Biostatistics, School of Public Health, College of Health Sciences. Email him at: haftom.temesgen@mu.edu.et.

which two groups are followed over time to assess group differences. Optimal design techniques were used to optimize power over feasible designs as a function of duration of a study, frequency of observations, and number of participants. For the GLMM, optimal designs were studied extensively in the literature by Han and Chaloner (2004); Niaparast (2009); Niaparast and Schwabe (2013); and Abebe, Tan, van Breukelen, and Berger (2014a, c), among others.

Unfortunately, optimal designs for nonlinear models depend on the unknown parameter values of interest, that is, on the regression weights that reflect the outcome change over time. Thus, in order to find the optimal design, the model parameter values should be known in advance. However, the parameter values are always unknown as the design is planned to obtained data for estimating them. A common approach to this problem is to use a best guess of the parameter values, which leads to locally-optimal designs, that is, designs which are optimal for a given set of parameter values (see, e.g., Chernoff, 1953). Such designs may not be efficient when the true parameter values differ from those best guesses, that is, the design may not be robust for other parameter values. To overcome this local optimality problem, various methods have been proposed in the literature (see, e.g., Berger & Wong, 2009). The Bayesian approach is one way that has been shown to be useful to take into account the uncertainty of the parameter values (Chaloner & Larntz, 1989; Atkinson, Donev, & Tobias, 2007; Abebe et al., 2014a, b, c; Abebe et al., 2015; among others). The Bayesian design literature is vastly restricted to binary response models. However, no user-friendly software has been developed so far for Bayesian design of longitudinal studies with binary responses.

Due to the absence of a computer program, planners of longitudinal studies in psychology, health sciences, and medicine face the problem of choosing the best number and timing of the repeated measurements. Usually the number and the allocation of the time points at which the measurements are taken are determined by non-statistical criteria. As an example, consider the Dutch smoking prevention study, where smoking and other data were collected from 3735 children in 156 elementary schools by means of a questionnaire at six time points between September 1997 and September 2000: September 1997, February 1998, June 1998, May 1999, February 2000 and September 2000 (Ausems, Mesters, van Breukelen, & De Vries, 2002).

Another example is the attention deficit hyperactivity disorder (ADHD) study (Lahey et al., 1998; Hartung et al., 2002). It was a longitudinal study on 255 children that sought to identify risk and prognostic factors in early childhood for ADHD symptoms, diagnoses, and functional outcomes across childhood,

adolescence, and early adulthood. All participants were followed over seven annual visits after the baseline. The question is whether these designs are efficient, in terms of the number and the timing of the measurements, for estimating the change in smoking and ADHD prevalence over the total follow-up time. This question can be answered by optimal design theory, which is part of the field of statistics.

For the linear random effects model optimal designs were discussed by Tan and Berger (1999) and Tekle, Tan, and Berger (2009), among others. They showed that regardless the underlying polynomial regression model, the number of repeated measures should be chosen as close as possible to the number of regression parameters. Ouwens, Tan, and Berger (2006) and Tekle et al. (2008) extended the work on optimal designs for logistic models with random effects using a maximin approach to handle the local optimality problem, without considering the cost of sampling and measuring. They have kept constant the number of subjects and the number of repeated measures per subject. But in a longitudinal study, costs are associated with the inclusion of patients (subjects) as well as with each repeated measurement.

Further, Bayesian designs are an increasingly popular alternative to maximin design as a method to overcome the local optimality problem. The Bayesian approach takes the uncertainty of the parameter values of the statistical model into account by using a prior distribution on the unknown parameters rather than single-value guesses. This will give more flexibility.

Therefore, a new interactive computer program is presented that computes Bayesian optimal repeated measurements designs for mixed effects logistic models with polynomial time effects under cost constraints, but also allows the user to compute maximin designs. The maximin approach essentially minimizes the largest possible (generalized) variance of the fixed-effect estimators within a user-specified region of the true fixed-effect values, or equivalently, it optimizes among worst possible efficiencies (see, e.g., Tekle et al., 2008; Ouwens et al., 2006).

It computes Bayesian optimal designs for longitudinal studies under cost constraints, thus helping researchers to reduce their study costs. The computer program helps users to identify the optimal number and optimal allocation of time points for a given subject-to-measurement cost ratio. Moreover, it computes the loss in efficiency of equidistant time points compared to the optimal allocation. It produces a plot of optimal allocations of time points under different values of autocorrelation. A separate manual is presented in the appendix and describes the

capabilities of the software, which runs in a Matlab environment (MathWorks, 2010).

The logistic mixed effects model with polynomial time effects is described, and the optimality criterion and the relative efficiency as a measure for the comparison of designs. Thereafter, the smoking prevention study by Ausems et al. (2002) is used to illustrate the application of the program and to discuss the various decisions that the user has to make when determining the most efficient design. The manual can be considered as part of the paper, but can be consulted independently from it. Finally, conclusions and recommendations are provided. The paper ends with a summary and discussion.

The Logistic Mixed Effects Model

Let the $q \times 1$ vector $\mathbf{y}_i = (y_{i1},\ldots, y_{iq})'$ be binary responses y_{ij} of subject i at q time points, $i = 1, 2,\ldots, N$ and $j = 1,\ldots, q$. It is assumed that all subjects have measurements at the same time points, and that, conditional on the subject-specific random effect vector $\mathbf{b}_i$, the binary responses y_{ij} of $\mathbf{y}_i$ are assumed to be Bernoulli distributed with probability of success $\mathrm{p}(y_{ij} = 1 \,|\, \mathbf{b}_i)$. These probabilities are related to the fixed and random effects via the logit link function. The corresponding logistic mixed effects model is given by:

$$\operatorname{logit}\left(\mathrm{p}\left(y_{ij}=1|\,\mathbf{b}_i\right)\right)=\log\left(\frac{\mathrm{p}\left(y_{ij}=1|\,\mathbf{b}_i\right)}{1-\mathrm{p}\left(y_{ij}=1|\,\mathbf{b}_i\right)}\right)=\mathbf{x}'_j\boldsymbol{\beta}+\mathbf{z}'_j\mathbf{b}_i \qquad (1)$$

where the $p \times 1$ vector $\mathbf{x}_j$ is the design vector of the explanatory variables at the j^{th} measurement for subject i, $\boldsymbol{\beta}$ is the corresponding $p \times 1$ vector of fixed polynomial time effects, and $\mathbf{z}_j$ is the $r \times 1$ design vector for the random effects that is usually a subset of vector $\mathbf{x}_j$. The vector $\mathbf{b}_i$ is the corresponding $r \times 1$ vector of random effects, which is assumed to have a multivariate normal distribution with mean zero and covariance matrix $\mathbf{D}$.

For example, if a quadratic ($p = 3$) time effect is assumed, the design vector is $\mathbf{x}'_j = \begin{pmatrix}1 & t_j & t_j^2\end{pmatrix}$ and $\boldsymbol{\beta} = (\beta_0\ \beta_1\ \beta_2)'$, where t_j is the time point of the j^{th} measurement, $j = 1,\ldots, 6$, and β_0, β_1, and β_2 are the fixed effects. Suppose that a random intercept and random linear slope are assumed. Then the design vector is $\mathbf{z}'_j = (1\ t_j)$ and $\mathbf{b}_i = (b_{0i}\ b_{1i})'$, where b_{0i} and b_{1i} are the corresponding random (subject-specific) deviations from these fixed effects, $i = 1,\ldots, 3735$. Then,

according to model (1), the log-odds of a positive response (smoking) for subject i at time t_j is given by:

$$\text{logit}\left(\text{p}\left(y_{ij}=1|\mathbf{b}_i\right)\right)=\left(\beta_0+b_{0i}\right)+\left(\beta_1+b_{1i}\right)t_j+\beta_2 t_j^2 \qquad (2)$$

To prevent misunderstanding about the flexibility of this model, note that it can handle U-shaped as well as monotonic trends over time. For the average subject (i.e. if the random effects are zero), the derivative of (2) with respect to time t is 0 if

$$t=\frac{-\beta_1}{2*\beta_2}$$

The time variable is bounded by the follow-up period of the longitudinal study, and so equation (2) reaches its maximum or minimum inside or outside the time interval, depending on the values of β_1 and β_2. So model (2) can handle monotonic as well as non-monotonic trends.

For example, in the Dutch smoking prevention study, a quadratic ($p = 3$) time effect will be needed if smoking prevalence on the logodds scale increases nonlinearly over time. For the sequel, it is important to note that in this paper and software, the time interval is scaled as $t \in [-1, +1]$. This can be translated into any suitable time scale by linear transformation, and vice versa. For instance, the time scale of the smoking prevention study, with its baseline of September 1997 as the origin, its last measurement in September 2000, and a month as the unit of measurement, is obtained by the transformation $t^* = 18(t + 1)$. Likewise, our present time scale is obtained as $t = (t^* - 18)/18$. The repeated measurements of smoking were made at time points $t^* = 0$, 5, 9, 20, 29, and 36 months, which in terms of the present time scale gives as time points t = -1.00, -0.72, -0.50, 0.11, 0.61, and 1.00, respectively.

Due to the random effects in model (1) and (2), the log-likelihood cannot be written down in closed form. Hence, either numerical methods or approximations to the log-likelihood must be used. Numerical methods require large computational resources and more importantly they require full knowledge of the data (Moerbeek, Van Breukelen, & Berger, 2003; Han & Chaloner, 2004), making them computationally inconvenient for optimal design procedures. To overcome this problem, approximation methods are employed. There is a large statistical literature on various approximation methods, but here, for the purpose

of obtaining optimal designs, we will focus on the two most frequently used ones, which are implemented in commercially available software packages: first order penalized quasi-likelihood (PQL1) and an extended version of generalized estimating equations (GEE).

First Order Penalized Quasi-Likelihood

The PQL1 variances and covariances of the fixed parameter estimates are calculated using the first-order Taylor expansion around the fixed and random effects. An advantage is that the method performs well in terms of point estimates since it produces the smallest mean squared error and the bias of the estimators decreases as the sample size increases (Breslow & Clayton, 1993; Moerbeek et al., 2003; Jang & Lim, 2009). A disadvantage is that design optimization based on PQL1 is very time consuming. This is due to the fact that the covariance matrix of the binary responses, which must be inverted at each iteration of the optimization process, is very large because it depends on the random effects (which in the design stage are sampled from a multinormal distribution). The variance-covariance matrix of the estimator $\hat{\boldsymbol{\beta}}$ of the parameter $\boldsymbol{\beta}$ for the logistic mixed effects models (1) is approximated in PQL1 by:

$$\operatorname{var}\left(\hat{\boldsymbol{\beta}}\right) \approx \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1} \tag{3}$$

where $\mathbf{X}$ is the $Nq \times p$ design matrix formed by stacking $\{\mathbf{x}'_j\}$ for N subjects and q time points, and $\mathbf{V}$ is the $Nq \times Nq$ block-diagonal matrix with N blocks of $q \times q$ variance-covariance matrices given by:

$$\mathbf{v}_i \approx \mathbf{w}_i^{-1/2}\mathbf{R}\left(\rho\right)\mathbf{w}_i^{-1/2} + \mathbf{ZDZ}' \tag{4}$$

The $q \times q$ matrix $\mathbf{R}(\rho)$ is the residual correlation matrix, $\mathbf{Z}$ is the $q \times r$ design matrix with rows $\mathbf{z}'_j$, $j = 1,\ldots, q$, the $r \times r$ matrix $\mathbf{D}$ is the variance-covariance matrix of the random effects, and $\mathbf{w}_i^{-1}$ is the diagonal matrix of the conditional variances of the transformed responses given the random effects $\mathbf{b}_i$, which is equal to the inverse of the diagonal matrix of the conditional variances of the untransformed responses given the random effects $\mathbf{b}_i$ (See for detail Moerbeek et al., 2001; Molenberghs & Verbeke, 2005, p. 270). Note that, under conditional independence, $\mathbf{R}(\rho)$ is an identity matrix and equation (4) becomes

$$\mathbf{v}_i \approx \mathbf{w}_i^{-1} + \mathbf{ZDZ}' \tag{5}$$

The diagonal matrix of the conditional variances of the untransformed responses given the random effects $\mathbf{b}_i$, is given by:

$$\mathbf{w}_i = \text{diag}\left(w_{i1}^{\mathbf{b}_i}, \ldots, w_{iq}^{\mathbf{b}_i}\right) \tag{6}$$

where $w_{ij}^{\mathbf{b}_i} = \text{var}\left(y_{ij} \mid \mathbf{b}_i\right)$, for $i = 1,\ldots, N$, $j = 1,\ldots, q$. Since the random effects are unknown in the design stage, we will generate $\mathbf{b}_i$ from a multivariate normal distribution with mean zero and variance-covariance $\mathbf{D}$.

Extension of Generalized Estimating Equations

The extended GEE is an alternative method which is not likelihood-based. It has been extended by Zeger, Liang, and Albert (1988) and Molenberghs and Verbeke (2005) to include autocorrelations of the errors in the standard formulation of GLMM. The covariance matrix of the binary responses is expressed conditional on the random effects being zero, which makes the calculations much faster. The asymptotic variance-covariance matrix of $\hat{\boldsymbol{\beta}}$ for the logistic mixed effects models (1) with autocorrelation, based on the extension of the GEE approach, is approximated by:

$$\text{var}\left(\hat{\boldsymbol{\beta}}\right) \approx \left(\sum_{i=1}^{N} \frac{\partial^2 \mathbf{P}_i'}{\partial \boldsymbol{\beta}} \mathbf{u}_i^{-1} \frac{\partial^2 \mathbf{P}_i}{\partial \boldsymbol{\beta}'}\right)^{-1} \tag{7}$$

where $\hat{\boldsymbol{\beta}}$ is the estimator of $\boldsymbol{\beta}$ for model (1), $\mathbf{P}_i = (\text{p}(y_{i1} \mid \mathbf{b}_i),\ldots, \text{p}(y_{iq} \mid \mathbf{b}_i))'$ and the working variance-covariance matrix of the responses is given by:

$$\mathbf{u}_i \approx \mathbf{w}_i^{1/2} \mathbf{R}\left(\rho\right) \mathbf{w}_i^{1/2} + \mathbf{w}_i \mathbf{ZDZ}' \mathbf{w}_i \tag{8}$$

When there are no residual correlations in $\mathbf{R}(\rho)$, a conditional independence model or purely random effects model results and equation (8) reduces to

$$\mathbf{u}_i \approx \mathbf{w}_i + \mathbf{w}_i \mathbf{ZDZ}' \mathbf{w}_i \tag{9}$$

where $\mathbf{w}_i$ is the diagonal matrix of the conditional variances of untransformed responses given the random effects $\mathbf{b}_i = 0$, which is given by:

$$\mathbf{w}_i = \text{diag}\left(w_{i1}^{\mathbf{b}_i=0}, \ldots, w_{i1}^{\mathbf{b}_i=0}\right) \tag{10}$$

where $w_{ij}^{\mathbf{b}_i=0} = \text{var}\left(y_{ij} \mid \mathbf{b}_i = 0\right)$ for $i = 1, \ldots, N$, $j = 1, \ldots, q$ (Molenberghs & Verbeke, 2005, p. 443).

Time-structured data are naturally correlated (Berger, 1986). In this paper, a first order auto regressive (AR1) is considered, i.e., $\rho^{|t_j - t_l|}$, where $j, l = 1, \ldots, q$, and so ρ is the autocorrelation coefficient between two responses at a time distance of one, that is, $\rho = \text{Corr}(y_{ij}, y_{il})$ for which $|t_j - t_l| = 1$. This autocorrelation structure implies that repeated measurements closer in time are more highly correlated and that the correlation decreases as the distance between the time points increases.

Bayesian D-Optimal Design and Relative Efficiency

To introduce the notation for the optimality criterion, suppose that the study to be designed will have q ordered time points $t_1, t_2, \ldots, t_q$ at which measurements are taken for all N subjects. The design space Ξ then contains all designs of the form

$$\Xi = \left\{ \begin{pmatrix} t_1 & t_2 & \ldots & t_q \\ w_1 & w_2 & \ldots & w_q \end{pmatrix} : t_j \in [a, b], t_1 < t_2 < \ldots < t_q \right\} \tag{11}$$

with weight w_i indicating per time point what proportion of all observations is obtained at that point (see also, e.g., Bunke & Bunke, 1986, p. 506) and $q \geq p$ to make these fixed effects identifiable with p being the number of fixed parameters of the model. Although in general the weights (w_i) at the different time points can be different, in this paper we make the restriction of all weights equal to 1 ($w_1 = w_2 = \ldots = w_q = 1$) at all q ordered time points, i.e., measurements are taken on all N subjects at each time point, because we consider longitudinal designs and so all q repeated measurements are obtained from the same individuals. The time interval $[a, b]$ is assumed to be fixed by substantive constraints within the field of application, for example, the total follow-up time in the cohort study of smoking prevention is $b - a = 3$ years, or 36 months. A design ξ_q is an element of the design space Ξ if it has q time points within the time interval $[a, b]$.

Optimal designs are usually selected by minimizing a real-valued function of the variance-covariance matrix of the parameter estimators, here of the estimators of the three regression weights in (2), which is known as optimality criterion (see, e.g., Silvey, 1980). In this way the precision of the estimators and the power of their significance tests are maximized. Various optimality criteria have been proposed in the literature, such as the D-, A-, or G-optimality criteria. In this paper, we will focus on the best-known and most popular optimality criterion, i.e., the D-optimality criterion. This optimality criterion has two nice properties: 1. It minimizes the volume of the asymptotic confidence ellipsoid for the parameters, for instance for the fixed effects in model (2), thus giving the multivariate generalization of the familiar confidence interval for a single parameter; and 2. It does not depend on the coding used for the endpoints of the chosen time interval [a, b], for instance, on whether we code the time predictor in equation (2) as running from 0 to 1, or from -1 to +1, or use the original time scale in days or months. This means that if the coding for the time interval is transformed linearly, a D-optimal design for the new time interval is obtained by applying the same linear transformation to the D-optimal design for the old interval (see Ouwens et al., 2006).

For example, in the smoking study, the measurements were taken between September 1997 and September 2000 (a period of three years), and by linearly transforming the measured time points into the interval [-1, +1], the actual design of the smoking study ξ_6 becomes (-1 -0.72 -0.50 0.11 0.61 1). Likewise, if e.g. the D-optimal allocation of the time points for the smoking study is -1, -0.5, 0, 0.5, and 1 on the time interval [-1, +1], then it is after 0, 9, 18, 27, and 36 months respectively on the original time scale of [0, 36] months.

The D-optimal design ξ_q^* is the design among all possible designs ξ_q with q time points for which the determinant of the variance-covariance matrix of parameter estimators, for instance, the covariance matrix of $\left(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2\right)$ in model (2), is minimized (Berger & Wong, 2009). It should be noted that, for some studies, other criteria could be more obvious. Using the D-optimality criterion, all fixed-model parameters are considered to be equally important. If, for example, only some of the model parameters are of interest and others are considered to be nuisance, then a D_A-criterion will be more relevant, indicating that only a subset or m linear combinations of the p regression parameters ($p \geq m$) are of interest and specified by an $m \times p$ design matrix $\mathbf{A}$ (see, e.g., Tan, 2011). Nevertheless, the concentration here will be on this D-optimality criterion, because it is expected all fixed effects in model (1) will be of interest.

The variance-covariance matrix of the fixed-effects estimators $\hat{\boldsymbol{\beta}}$ depends on the unknown parameter vector $\boldsymbol{\beta}$ (see Abebe et al., 2014a, b, c; Abebe et al., 2015), which makes design optimization dependent on the very same parameters that have to be estimated with the study to be designed, thus creating a vicious circle. The Bayesian approach resolves this dependency problem by taking the expectation of a function of the variance-covariance matrix over a prior distribution for the unknown parameter vector $\boldsymbol{\beta}$. Thus, the Bayesian D-optimality criterion is defined as follows:

$$\begin{aligned}\phi_D(\xi \mid \pi) &= E_{\boldsymbol{\beta}} \log\left|\left(\operatorname{var}\left(\hat{\boldsymbol{\beta}}\right)\right)^{-1}\right| \\ &= \int_{\boldsymbol{\beta}} \log\left|\left(\operatorname{var}\left(\hat{\boldsymbol{\beta}}\right)\right)^{-1}\right| \pi(\boldsymbol{\beta})\, d\boldsymbol{\beta}\end{aligned} \tag{12}$$

where $\pi(\boldsymbol{\beta})$ is the prior distribution for $\boldsymbol{\beta}$ and $\operatorname{var}\left(\hat{\boldsymbol{\beta}}\right)$ is the variance-covariance matrix of $\hat{\boldsymbol{\beta}}$ for the logistic mixed effects models based on approximation methods (see Abebe et al., 2014a, b, c; Abebe et al., 2015 for details). In fact, the design criterion (12) follows from maximizing the expected Kullback-Leibner (KL)-distance between the prior and posterior distributions, measuring how much information can be gained when moving from prior to posterior. When the normal approximation is used for the posterior distribution, then a design that maximizes the KL-distance is equivalent to maximizing expression (12) and is called Bayes D-optimal. It should be mentioned that expression (12) does not represent the full Bayesian design criterion, but only approximately by ignoring the additional effect of the prior information about the fixed effects. However, for large sample sizes, the contribution of the prior information to the posterior variance is usually negligible (for further details, see Chaloner & Verdinelli, 1995; Sebastiani & Settimi, 1998). Note that maximization of (12) comes down to minimization of the expected log determinant of the covariance matrix, where the expectation is taken over the prior (Atkinson et al., 2007).

The precision of estimating the fixed-effects parameters $\boldsymbol{\beta}$ increases by taking more measurements and sampling more subjects (Moerbeek et al., 2001). However, the addition of subjects and of measurements per subject will increase the costs of the study and these are usually limited by budget constraints. Therefore, it is reasonable to take into account the costs of a longitudinal study when designs are compared with each other. There are two main components of

these costs. These are the costs for recruitment of subjects and the costs of the measurements once a subject has been recruited. Let the cost of recruiting a subject be C_1 and the cost of one measurement per subject be C_2. Then the total cost of a longitudinal study with q time points and N subjects, excluding overhead cost, is given by the linear cost function:

$$\begin{aligned} C &= C_1 N + C_2 N q \\ &= N C_2 \left(k + q\right) \end{aligned} \tag{13}$$

where $k = C_1/C_2$ is the ratio of the cost of adding a new subject to the cost of an additional measurement per subject.

To compare different designs, we will use their relative efficiencies while fixing the total costs C. This means that the designs can differ in terms of the number of subjects N and the number and timing of the measurements q. First, we compute the Bayesian D-optimal designs using fixed N and then we correct for costs and different q and N as follows: Let $\phi_{\mathrm{D}}\left(\xi_q^* \mid \pi\right)$ denote the value of design criterion (12) for the optimal design ξ_q^* with q time points, given the prior distribution π for the fixed effects. Then the relative efficiency (RE) of an arbitrary design ξ_s with s time points relative to the optimal design ξ_q^* is defined as:

$$\mathrm{RE}\left(\xi_s; \xi_q^*\left(\pi\right) \mid \pi\right) = \frac{k+q}{k+s}\left[\exp\left\{\frac{\phi_{\mathrm{D}}\left(\xi_s \mid \pi\right) - \phi_{\mathrm{D}}\left(\xi_q^* \mid \pi\right)}{p}\right\}\right] \tag{14}$$

where π is the prior distribution for the fixed effects and p is the number of fixed effects, that is, $p = 3$ for model (2). If the value of this relative efficiency is close to unity, then the design ξ_s is about equally efficient as the optimal design ξ_q^* for a given prior π. The inverse of this relative efficiency is the number of times that a design ξ_s must be replicated to have the same efficiency as the optimal design ξ_q^*. Note that the term between squared brackets on the right side of equation (14), so without the $(k + q)/(k + s)$ term, is the RE under the assumption of an equal number of subjects N for both designs, which then differ only in the number and timing of the repeated measures. This fixed N-situation, i.e. $N_s = N_q$, underlies the RE formula as given by Chaloner and Larntz (1989). However, if we keep the

total budget *C* instead of *N* the same for all designs, then it follows from equation (13) that we can have

$$\frac{N_s}{N_q} = \frac{k+q}{k+s}$$

as many subjects in design ξ_s as in design ξ_q^*. Since $\text{var}\left(\hat{\boldsymbol{\beta}}\right)$ is inversely proportional to the sample size *N*, it then follows from equations (12) and (13) that the RE of both designs obeys equation (14). See the Appendix for details on the derivation of RE in (14).

Method of Optimization

The Bayesian D-optimal designs for the logistic mixed effect model are found by our computer program numerically by maximization of the criterion value (12) among all candidate designs for a given prior distribution of the parameters. Details of this will be given in the next sub-sections.

Sampling Parameter Values from Priors to Compute the Criterion

To construct Bayesian designs for continuous prior distributions, all candidate designs must be evaluated in terms of their criterion values as defined by (12). However, evaluation of the integration over the prior distribution is very complicated and cannot easily be done analytically. A numerical approximation of the integral is necessary. Numerical approximations can be done by sampling parameter values from the prior distribution and then by replacing the integral in (12) with a summation over the sample (Atkinson et al., 2007; Chaloner & Verdinelli, 1995). Estimating (12) using the traditional sampling (pseudo Monte Carlo) method requires very large samples from the prior to reduce the sample-to-sample variability to the point where different samples do not lead to different design choices. Thus, this approach is costly in terms of computing time. In our computer program, we will use an Adaptive Rejection Metropolis Sampling (ARMS) algorithm (Gilks & Wild, 1992; Gilks, Best, & Tan, 1995), which is a more efficient sampling algorithm that requires a smaller sample to obtain a good approximation of the design criterion (12). ARMS is a generalization of the method of adaptive rejection sampling (ARS) (Gilks, 1992), which was itself a development of the original method proposed by Gilks and Wild (1992). The

ARMS generalization includes a Metropolis step to accommodate non-concavity in the log density. ARMS is a Markov chain Monte Carlo (MCMC) scheme for generating samples from high dimensional target distributions and widely used within Gibbs sampling, where automatic and fast samplers are often needed to draw. It can deal with (intrinsic) non-linear functions as often used in, for instance, pharmacokinetics. For the present log-linear model, the ARMS works very well and is much faster than the Gibbs sampling method.

Optimization Algorithm for Finding an Optimal Design

To find candidate designs and in particular the optimal design, the program uses the FMINCON function of MATLAB version 7.10.0499 (R2010a). This function performs constrained non-linear optimization and requires an initial design ξ_0. Without loss of generality, the time interval was coded as [-1, +1], and equally-spaced time points were used as initial designs. There is no need to start with non-equally spaced time points because our experience is that Bayesian optimal designs for our model do not depend on the spacing of the initial design. According to Firth and Hinde (1997), the Bayesian criterion may only lead to different optimal designs for different starting values when very dispersed prior distributions are considered. In fact, the Bayesian D-optimal designs as obtained with our program can deviate a lot from equidistance, thus showing that equidistance as initial design does not constrain the final design (see, e.g., Abebe et al., 2015).

The following global search algorithm is used to find the Bayesian D-optimal designs for a given multivariate normal prior distribution of the parameters:

1. Take samples from the prior distribution of the parameters using ARMS.
2. Compute the Bayesian D-optimal allocation of q time points, using $q = p$ equidistant time points as initial design, where p is the number of fixed parameters of the model. Note that the final optimal allocation does not need to be equally spaced (see, e.g., Abebe et al., 2014a).
3. Increase the number of time points q by one and perform step 2 again to find the Bayesian optimal design (allocation) for the new value of q. Repeat step 2 and 3 until the maximum number of time points q (user specified) is reached.

4. Thereafter, select the optimal number of time points q for the Bayesian D-optimal design by computing the relative efficiencies of designs with different numbers of time points against each other for a user-specified subject-to-measurement cost ratio. Do this for each cost ratio considered to obtain one optimal design per cost ratio for a chosen prior distribution.

An Example: The Dutch Smoking Prevention Study

As an illustration of the various decisions that the user has to make when determining the most efficient design, consider the Dutch smoking prevention study as described in the introduction section. A logistic mixed-effects model with quadratic time effect was found to give an adequate fit to the repeated measures of smoking status (0 = no, 1 = yes). Therefore, this model was adopted to illustrate the application of the BODMixed_Logistic program in guiding researchers for a similar future study. After starting the BODMixed_Logistic program, all the steps will be reviewed that are necessary to obtain the optimal design, starting with the specification of the model and the various input values. See the program manual for a description of the graphical user interface offered per step.

Choice of the Model

The first step is to choose the statistical model; the optimal design depends on the underlying statistical model and is different for a quadratic model than for a linear one. For the fixed model part, we choose a quadratic growth function, both in view of its fit to the smoking data and because it is more flexible than a linear one and can handle monotonic trends as well as U-shaped trends due to the finite time interval. For the random model part, we assume a random intercept as well as a random linear slope. This can be specified in the program by choosing nonzero variances for the intercept and linear slope and zero variance for the quadratic slope, with or without slope-intercept covariance.

To the program user it may be reassuring to know that Abebe et al. (2014c) found that the Bayesian D-optimal designs are hardly affected by the choice of a covariance structure for the random effects, at least in case of a non-zero autocorrelation and the presence of a random intercept or random slope. Further, the autocorrelation between the repeated measures must be specified. Fortunately, the maximum loss in efficiency incurred by misspecification of the autocorrelation appears to be less than 5% (Abebe et al., 2014c), excepting the

case of a zero autocorrelation which gives very different allocations of time points than nonzero values. For illustration purpose, we will assume the default value of 0.1 for the autocorrelation, remembering that this is the correlation between two measurements with a time interval of 1 on the time scale [-1, +1]. Of course, the program user is free to try out different covariance structures and autocorrelations to check the dependence of the optimal design on these values for his/her specific study.

Approximation Method

Next, the user has to choose between the two approximations of the likelihood that are implemented in the program: PQL1 and extended GEE. If computation time is not an issue, then we would recommend using the PQL1 approximation. The extended GEE, however, is computationally much faster and often produces similar Bayesian D-optimal designs as the PQL1 approximation (Abebe et al., 2014c). In this example, we choose the extended GEE.

Choice of Optimality Criterion

At this stage, the model and the necessary parameter values have been specified. The program offers three different optimization criteria.

a. The option 'Bayesian D-optimal' maximizes the criterion in equation (12), thus minimizing the generalized variance of the fixed effects estimators, for a user specified prior distribution of those fixed effects. Abebe et al. (2014b, c) showed that it is best to choose a prior distribution with a large variance (uninformative prior) to express the degree of uncertainty about the 'true' parameter values. The prior means then have little impact on the optimal design, provided that the autocorrelation is not too close to zero ($\rho > 0.001$).
b. The option 'locally D-optimal' criterion can be chosen if the user wants to check the optimal design for specific values of the fixed effects regression parameters. Note that this comes down to assuming a prior with zero variance. This option is in general not recommended, because it will often lead to a sub-optimal design.
c. The option 'Maximin D-optimal design' essentially minimizes (among all possible designs) the largest possible (generalized) variance of the fixed-effect estimators within a user-specified region of the true fixed-effect values, or equivalently, it maximizes the minimum efficiency within this

> region (see, e.g., Tekle et al., 2008; Ouwens et al., 2006). Using this criterion, the user remains on the safe side, and will furthermore obtain a design that is optimal for at least one combination of likely parameter values. A disadvantage of this criterion is that the maximin design is often optimal for some points on the boundary of the region ("parameter space") for the true fixed effects, and these boundary points are less likely than values within the region (Atkinson et al., 2007, p. 258).

For this illustration, Bayesian D-optimal design is selected with, as input prior distribution for the fixed effects, an independent normal with prior means μ = [1, 2, 3] and a prior variance $\sigma^2 = 5$ for both fixed effects. Abebe et al. (2014c) showed that the Bayesian D-optimal designs with such large prior variance are hardly affected by the choice of prior means, provided that the autocorrelation is not too close to zero ($\rho > 0.001$).

Optimal designs can be determined now in either of two ways: By fixing the number of time points q and finding the optimal q allocations, or by finding the optimal number and allocation of time points for a given subject-to-measurement cost ratio k.

Computing the Optimal Allocation for a Given Number of Time Points *q*

For this illustration, we use $q = 6$ time points as the design in the smoking example had 6 repeated measurements. The resulting optimal time points are, according to Figure 1 (see the 4th design in it), [-1, -0.6080, -0.2063, 0.1875, 0.5465, 1]. Translated into the scale of the smoking study period in months, that is, into the time interval [September 1997, September 2000], this gives as optimal design points September 1997, April 1998, November 1998, June 1999, January 2000, and September 2000. To compare, the actual time points were September 1997, February 1998, June 1998, May 1999, February 2000, and September 2000. In this example we fixed the number of time points, but it may be of interest to find the optimal number of time points for a given subject-to-measurement cost ratio, which will now be discussed.

Finding the Optimal Design for a Given Subject-to-Measurement Cost Ratio *k*

As mentioned previously, the user can choose between fixing the number of time points q and fixing the subject-to-measurement cost ratio. The second option will

now be illustrated assuming a cost ratio $k = 1$, that is, equal costs for recruiting a subject and for a single measurement on a single subject. A maximum of seven time points were chosen, which covers the number of time points in most longitudinal studies. The minimum is three because the model has $p = 3$ fixed effects and is thus not identifiable with less than three time points.

The results are given in Figure 1, showing the Bayesian optimal designs for each of the number of time points $q = 3, 4,\ldots, 7$, and the relative efficiency of each Bayesian optimal design compared to the Bayesian optimal design with $q = 7$ time points for the chosen cost ratio, here $k = 1$. The optimal number of repeated measures q for that cost ratio is $q = 4$, giving a relative efficiency of 1.2324 compared to $q = 7$. Further, the relative efficiency of an equidistant design with $q = 4$ time points compared to the optimal design with $q = 4$ is 0.9770, and so equidistance is highly efficient here, although it is not optimal. Finally, to show the effect of the chosen cost ratio on the optimal design, Figure 2 gives a plot of the relative efficiencies of the Bayesian optimal designs with different numbers of time points compared to the optimal design with the maximum number of time points, for each of several cost ratios k. Clearly, the optimal number of time points increases as the subject-to-measurement cost ratio becomes large. The practical implication of this is that, if the user is uncertain about the cost ratio, he or she should try several cost ratios within the plausible range.

The efficiencies of the actual design of the smoking design relative to the Bayesian optimal design increase with an increasing cost ratio k, and the relative efficiency is large for cost ratios $k \geq 2$. For small cost ratios k, the loss in efficiency for the actual design relative to the Bayesian design with 4 time points is at most 25%, which can be compensated by sampling about 33% more children. For large cost ratios ($k \geq 10$), the loss in efficiency for the actual design is at most about 4%, which can be compensated by sampling about 4% more children.

Plotting the Bayesian Optimal Design for Different Values of the Autocorrelation

In the example it was assumed there is a single value 0.1 for the autocorrelation. However, the autocorrelation is rarely known in the design stage. The program therefore offers as a last option a plot of the effect of the autocorrelation value on the Bayesian D-optimal design for a user specified number of time points q and range of autocorrelation. Figure 3 shows such a plot for $q = 6$ time points (horizontal axis) against the autocorrelation (vertical axis) within the range from 0.001 to 0.90 for the random intercept logistic model with

quadratic time effects. From this plot we see that the Bayesian D-optimal allocation for $q = 6$ is fairly independent of the size of the autocorrelation, at least within the chosen range from 0.001 to 0.9. As mentioned before, a zero autocorrelation usually gives quite different optimal allocations which are far from equidistant.

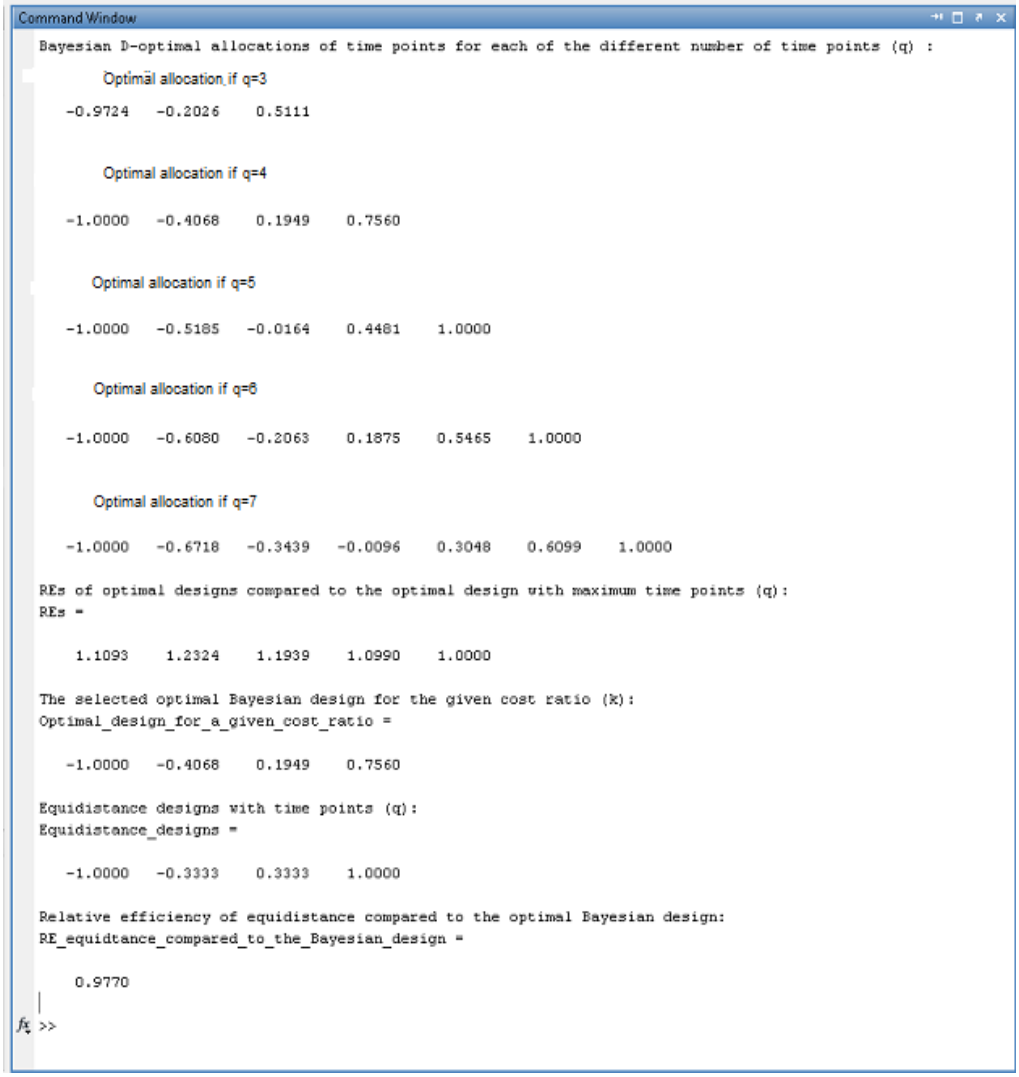

Figure 1. Bayesian optimal allocations of time points for cost ratio $k = 1$ with a maximum number of time points $q = 7$ for the logistic mixed model with quadratic time effects, assuming a random intercept and random linear slope logistic model with quadratic time effects, and autocorrelation 0.1

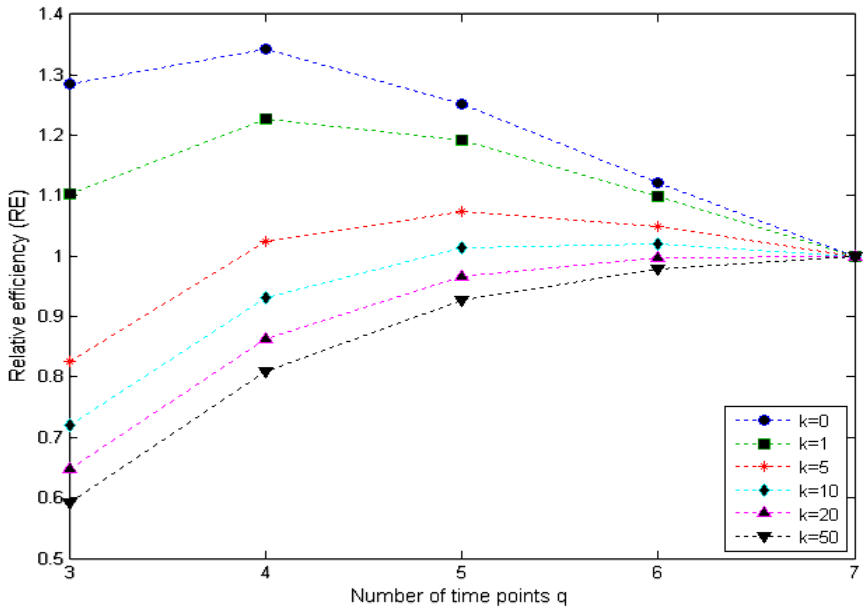

Figure 2. Relative efficiency of Bayesian optimal designs compared to the Bayesian optimal design

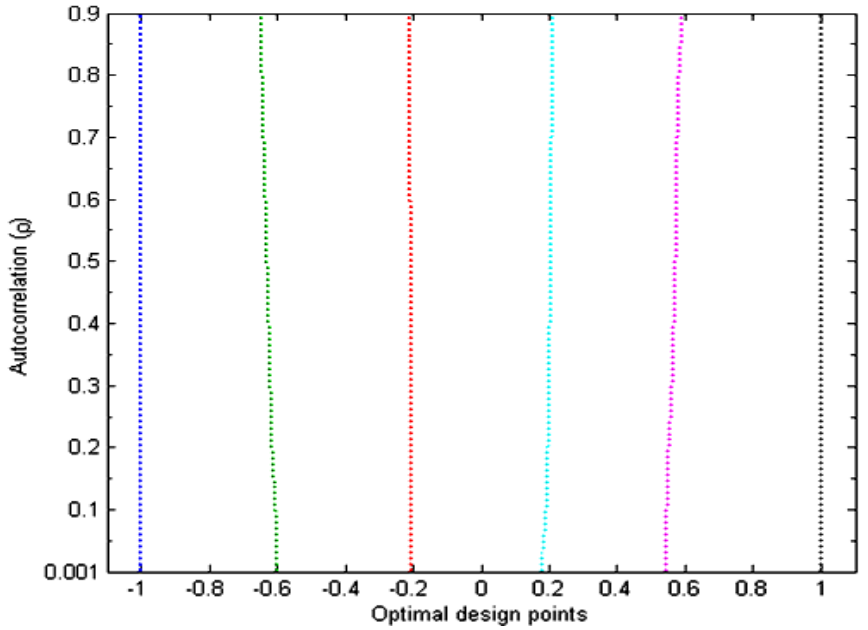

Figure 3. Bayesian D-optimal allocation of $q = 6$ time points as a function of the autocorrelation, for the logistic mixed model with quadratic time effects

Table 1. Optimal time points for a quadratic model, fixed effects, number of time points $q = 6$

Prior variance σ^2	Prior mean ($\beta_1, \beta_2, \beta_3$)	Autocorrelation (ρ) 0.0	0.01	0.9
0.5	[0, 0, 0]	(-1, -1, 0, 0, 1, 1)	(-1,-0.60,-0.19, 0.19,0.60,1)	(-1,-0.63,-0.22,0.22,0.64,1)
	[1, 2, 3]	(-1,-1, -0.27, -0.27, 0.47,0.47)	(-1,-0.66,-0.30,0.04,0.40,0.71)	(-1,-0.69, -0.25,0.21,0.53,1)
5	[0, 0, 0]	(-1, -0.36, 0.06, 0.06, 0.56, 1)	(-1, -0.60, -0.20, 0.20, 0.60, 1)	(-1,-0.67,-0.23,0.25,0.67, 1)
	[1, 2, 3]	(-1,-0.60,-0.24,0.13,0.46,0.91)	(-1, -0.60, -0.21, 0.18, 0.54, 1)	(-1,-0.65,-0.21,0.21, 0.60,1)

Table 2. Optimal time points for a quadratic model, random intercept, intercept variance $\tau_0^2 = 1$, number of time points $q = 6$

Prior variance σ^2	Prior mean ($\beta_1, \beta_2, \beta_3$)	Autocorrelation (ρ) 0.0	0.01
0.5	[0, 0, 0]	(-1, -1, 0, 0, 1, 1)	(-1, -0.58, -0.18, 0.18, 0.58, 1)
	[1, 2, 3]	(-1, -1, -0.26, -0.26, 0.51, 0.51)	(-1, -0.66, -0.29, 0.06, 0.41, 0.73)
5	[0, 0, 0]	(-1, -1, -0.28, 0.29, 1, 1)	(-1, -0.60, -0.20, 0.20, 0.60, 1)
	[1, 2, 3]	(-1, -0.62, -0.22, 0.19, 0.50, 0.98)	(-1, -0.60, -0.21, 0.18, 0.55, 1)

Table 3. Optimal time points for a quadratic model, random intercept/slope, random intercept variance $\tau_0^2 = 1$, random slope variance $\tau_1^2 = 1$, number of time points $q = 6$

Prior variance σ^2	Prior mean ($\beta_1, \beta_2, \beta_3$)	Autocorrelation (ρ) 0.0	0.01
0.5	[0, 0, 0]	(-1, -1, 0, 0, 1, 1)	(-1, -0.56, -0.17, 0.17, 0.57, 1)
	[1, 2, 3]	(-1, -1, -0.24, -0.24, 0.51, 0.51)	(-1, -0.64, -0.28, 0.07, 0.42, 0.73)
5	[0, 0, 0]	(-1, -0.66, -0.17, 0.17, 0.63, 1)	(-1, -0.60, -0.20, 0.20, 0.60, 1)
	[1, 2, 3]	(-1, -0.59, -0.20, 0.18, 0.50, 0.97)	(-1, -0.60, -0.21, 0.18, 0.54, 1)

Summarizing the example of Bayesian optimal design with the BODMixed_Logistic program, it can be concluded that when the subject-to-measurement cost ratio k is less than 5, i.e. the cost of an additional subject does not exceed five times the cost of an additional observation on a single subject, then the optimal number of repeated measurements is four time points. Further, the optimal allocation is not equidistant, but equidistance is highly efficient.

Using the suggested Bayesian D-optimal design, the relative efficiency of the optimal number of repeated measures q for the given cost ratio $k = 1$, which is equal to $q = 4$, relative to the $q = 6$ (which is the number of time points in the smoking study) is equal to about RE = 1.2324/1.099 = 1.1213 (see Figure 2). This means that about 10% less budget is needed for the optimal design to reach the same efficiency as compared to the actual design of the smoking prevention study of Ausem et al. (2004), which had six time points.

Finally, to demonstrate the effect of the covariance structure D, prior means and variances, as well as of autocorrelation on the Bayesian D-optimal design, we will show some additional results for a quadratic model with fixed effects, random intercept, random intercept/slope, and for various priors and autocorrelations. We fixed the number of time points to $q = 6$ and used the extended GEE method for these results which are summarized in Tables 1 to 3, which gives the optimal time points for varying parameter values.

Shown in Table 1 are optimal allocations of time points for a quadratic model with fixed effects only, Table 2 for the random intercept model with intercept variance equal to $\tau_0^2 = 1$, and Table 3 for the random intercept/slope model with intercept variance and slope variance equal to $\tau_0^2 = 1$ and $\tau_1^2 = 1$, respectively. It can be seen that when there is no autocorrelation (i.e. $\rho = 0$), the optimal allocation of time points depends strongly on the covariance structure and priors and coinciding time points occur. Further, when the autocorrelation $\rho > 0$, the optimal allocations are never coinciding and are comparable for a prior variance equal to $\sigma^2 = 5$ and all covariance structures D. The effect of a large versus small autocorrelation is only presented for the fixed effects model ($\mathbf{D} = 0$), because Abebe et al. (2014c) already showed this for the random effects models. Finally, the prior means do not have much effect on the optimal allocation. This is in line with the findings of Abebe et al. (2014c) for a large prior variance.

Summary and Discussion

Optimal designs for longitudinal studies have been shown useful to improve the precision of the model parameter estimates of interest. Due to absence of a computer program for the optimal design of longitudinal studies with a binary response, planners of such studies in psychology, health sciences, and medicine have not yet benefitted from optimal design theory. We present a user-friendly computer program that computes Bayesian optimal designs for mixed effects logistic models with polynomial time effects. This computer program helps researchers to identify the optimal number and allocations of time points of measurements for a given subject-to-measurement cost ratio, and computes the loss in efficiency of equidistance compared to the optimal allocation. Moreover, it helps to assess the effect of autocorrelation on optimal allocations of design points. The program was illustrated on a smoking prevention study showing that, when the cost ratio k is less than 5, the optimal number of repeated measurements is 4 time points. Further, the optimal allocation is not equidistant, but equidistance is highly efficient.

The use of a Bayesian design does not force researchers to use Bayesian methods to analyze the data. Once the experimental data is collected by using the Bayesian D-optimal design, researchers can fit their model either with Bayesian or with frequentist methods.

The current version of the MATLAB program BODMixed_Logistic is freely available upon request from the corresponding author, which may be available eventually via the internet. The current version of the program considers designs based on the D-optimality criterion and assumes that all subjects are available over the total study period and that there is no dropout. Further, extensions of the model and software can be made by, e.g., adding a grouping variable or covariates like age or allowing for different types of covariance structures than already described in this paper. Future work may therefore aim at these extensions and at allowing for dropout. Another important issue for future work is Bayesian optimal design for model using non-polynomial (splines) time effects.

References

Abebe, H. T., Tan, F. E. S., van Breukelen, G. J. P., & Berger, M. P. F. (2014a). Bayesian *D*-optimal designs for the two parameter logistic mixed effects model. *Computational Statistics and Data Analysis, 71*, 1066-1076. doi: 10.1016/j.csda.2013.07.040

Abebe, H. T., Tan, F. E. S., van Breukelen, G. J. P., & Berger, M. P. F. (2014b). On the choice of a prior for Bayesian D-optimal designs for the logistic regression model with a single predictor. *Communications in Statistics - Simulation and Computation, 43*(7), 1811-1824. doi: 10.1080/03610918.2012.745556

Abebe, H. T., Tan, F. E. S., van Breukelen, G. J. P., & Berger, M. P. F. (2014c). Robustness of Bayesian D-optimal design for the logistic mixed model against misspecification of autocorrelation. *Computational Statistics, 29*(6), 1667-1690. doi: 10.1007/s00180-014-0512-3

Abebe, H. T., Tan, F. E. S., van Breukelen, G. J. P., & Berger, M. P. F. (2015). Bayesian design for dichotomous repeated measurements with autocorrelation. *Statistical Methods in Medical Research, 24*(5), 594-611. doi: 10.1177/0962280213508850

Atkinson, A. C., Donev, A. N., & Tobias, R. D. (2007). *Optimum experimental designs, with SAS*. Oxford, UK: Clarendon Press.

Ausems, M., Mesters, I., van Breukelen, G., & De Vries, H. (2002). Short-term effects of a randomized computer-based out-of-school smoking prevention trial aimed at Dutch elementary schoolchildren. *Preventive Medicine, 34*(6), 581-589. doi: 10.1006/pmed.2002.1021

Berger, M. P. F. (1986). A comparison of efficiencies of longitudinal, mixed longitudinal, and cross-sectional designs. *Journal of Educational Statistics, 11*(3), 171-181. doi: 10.3102/10769986011003171

Berger, M. P. F., & Wong, W. K. (2009). *An introduction to optimal designs for social and biomedical research*. Chichester, UK: Wiley.

Breslow, N. E., & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association, 88*(421), 9-25. doi: 10.2307/2290687

Bunke, H, & Bunke, O. (1986). *Statistical inference in linear models* (Vol.1). Chichester, UK: Wiley.

Chaloner, K., & Larntz, K. (1989). Optimal Bayesian designs applied to logistic regression experiments. *Journal of Statistical Planning and Inference, 21*(2), 191-208. doi: 10.1016/0378-3758(89)90004-9

Chaloner, K., & Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science, 10*(3), 273-304. doi: 10.1214/ss/1177009939

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *The Annals of Mathematical Statistics, 24*(4), 586-602. doi: 10.1214/aoms/1177728915

Firth, D., & Hinde, J. P. (1997). On Bayesian D-optimum design criteria and the equivalence theorem in non-linear models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology), 59*(4), 793-797. doi: 10.1111/1467-9868.00096

Gilks, W. R. (1992). Derivative-free adaptive rejection sampling for Gibbs sampling. In J. M Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian statistics 4: Proceedings of the Fourth Valencia International Meeting, April 15-20, 1991* (pp.641-649). Oxford, UK: Clarendon Press.

Gilks, W. R., Best, N. G., & Tan, K. K. C. (1995). Adaptive rejection metropolis sampling within Gibbs sampling. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 44*(4), 455-472. doi: 10.2307/2986138

Gilks, W. R., & Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 41*(2), 337-348. doi: 10.2307/2347565

Han, C., & Chaloner, K. (2004). Bayesian experimental designs for nonlinear mixed models with application to HIV dynamics. *Biometrics, 60*(1), 25-33. doi: 10.1111/j.0006-341X.2004.00148.x

Hartung, C., Willcutt, E., Lahey, B., Pelham, W., Loney, J., Stein, M., & Keenan, K. (2002). Sex differences in young children who meet criteria for attention deficit hyperactivity disorder. *Journal of Clinical Child & Adolescent Psychology, 31*(4), 453-464. doi: 10.1207/S15374424JCCP3104_5

Jang, W., & Lim, J. (2009). A numerical study of PQL estimation biases in generalized linear mixed models under heterogeneity of random effects. *Communication in Statistics – Simulation and Computation, 38*(4), 692-702. doi: 10.1080/03610910802627055

Lahey, B., Pelham, W., Stein, M., Loney, J., Irapani, C., Nugent, K.,…Baumann, B. (1998). Validity of DSMIV attention-deficit/hyperactivity disorder for younger children. *Journal of the American Academy of Child & Adolescent Psychiatry, 37*(7), 695-702. doi: 10.1097/00004583-199807000-00008

MathWorks. (2010). MATLAB (Version 7.10.0.499 (R2010a)) [computer software]. Natick, MA: The MathWorks, Inc.

McClelland, G. H. (1997). Optimal design in psychological research. *Psychological Methods, 2*(1), 3-19. doi: 10.1037/1082-989X.2.1.3

Moerbeek, M., Van Breukelen, G. J. P., & Berger, M. P. F. (2001). Optimal experimental designs for multilevel logistic models. *Journal of the Royal Statistical Society. Series D (The Statistician), 50*(1), 17-30. doi: 10.1111/1467-9884.00257

Moerbeek, M., Van Breukelen, G. J. P., & Berger, M. P. F. (2003). A comparison of estimation methods for multilevel logistic models. *Computational Statistics, 18*(1), 19-37. doi: 10.1007/s001800300130

Molenberghs, G., & Verbeke, G. (2005). *Models for discrete longitudinal data*. New York, NY: Springer.

Niaparast, M. (2009). On optimal design for a Poisson regression model with random intercept. *Statistics & Probability Letters, 79*(6), 741-747. doi: 10.1016/j.spl.2008.10.035

Niaparast, M., & Schwabe, R. (2013). Optimal design for quasi-likelihood estimation in Poisson regression with random coefficients. *Journal of Statistical Planning and Inference, 143*(2), 296-306. doi: 10.1016/j.jspi.2012.07.009

Ouwens, M. J. N. M., Tan, F. E. S., & Berger, M. P. F. (2006). A maximin criterion for the logistic random intercept model with covariates. *Journal of Statistical Planning and Inference, 136*, 962-981. doi: 10.1016/j.jspi.2004.07.014

Raudenbush, S. W., & Feng, L. X. (2001). Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change. *Psychological Methods, 6*(4), 387-401. doi: 10.1037/1082-989X.6.4.387

Sebastiani, P., & Settimi, R. (1998). First-order optimal design for non-linear models. *Journal of Statistical Planning and Inference, 74*(1), 177-192.

Silvey, S. D. (1980). *Optimal design: An introduction to the theory for parameter estimation*. New York, NY: Chapman and Hall.

Tan, F. E. S. (2011). Conditions for D_A–maximin marginal designs for generalized linear mixed models to be uniform. *Communications in Statistics-Theory and Methods, 40*(2), 255-266. doi: 10.1080/03610920903411218

Tan, F. E. S., & Berger, M. P. F. (1999). Optimal allocation of time points for the random-effects model. *Communication in Statistics – Simulation and Computation, 26*(2), 517-540. doi: 10.1080/03610919908813563

Tekle, F. B., Tan, F. E. S., & Berger, M. P. F. (2008). Maximin D-optimal designs for binary longitudinal responses. *Computational Statistics & Data Analysis, 52*(12), 5253-5262. doi: 10.1016/j.csda.2008.04.037

Tekle, F. B., Tan, F. E. S., & Berger, M. P. F. (2009). Interactive computer program for optimal designs of longitudinal cohort studies. *Computer Methods and Programs in Biomedicine, 94*(2), 168-176. doi: 10.1016/j.cmpb.2008.11.002

Zeger, S. L., Liang, K-Y, & Albert, P. S (1988). Model for longitudinal data: a generalized estimating equation approach. *Biometrics, 44*(4), 1049-1060. doi: 10.2307/2531734

Appendix A: Derivation for the Relative Efficiency Equation (14)

To compare designs we compute their efficiencies using the concept of equivalent sample size (see Atkinson et al., 2007, p. 152; Berger & Wong, 2009, p. 37). Let $\text{var}\left(\hat{\boldsymbol{\beta}}_{\xi_s}\right)$ and $\text{var}\left(\hat{\boldsymbol{\beta}}_{\xi_q}\right)$ be the variance-covariance matrices of $\hat{\boldsymbol{\beta}}$ for the design ξ_s with s time points and the design ξ_q with q time points, respectively, and let N_s and N_q be the number of subjects for the design ξ_s and ξ_q, respectively. For the D-criterion and a given model with p parameters, the relative efficiency of design ξ_s compared to design ξ_q is given by:

$$\text{RE}\left(\xi_s;\xi_q\right)=\frac{N_s}{N_q}\left[\frac{\det\left\{\left[\text{var}\left(\hat{\boldsymbol{\beta}}_{\xi_s}\right)\right]^{-1}\right\}}{\det\left\{\left[\text{var}\left(\hat{\boldsymbol{\beta}}_{\xi_q}\right)\right]^{-1}\right\}}\right]^{\frac{1}{p}} \tag{15}$$

Where the two determinants in (15) are both based on one subject only, and the factor N_s/N_q takes into account the sample size per design.

This relative efficiency (15) can be rewritten as follows:

$$\begin{aligned}\text{RE}\left(\xi_s;\xi_q\right)&=\frac{N_s}{N_q}\exp\left\{\log\left\{\left[\frac{\det\left\{\left[\text{var}\left(\hat{\boldsymbol{\beta}}_{\xi_s}\right)\right]^{-1}\right\}}{\det\left\{\left[\text{var}\left(\hat{\boldsymbol{\beta}}_{\xi_q}\right)\right]^{-1}\right\}}\right]^{\frac{1}{p}}\right\}\right\}\\&=\frac{N_s}{N_q}\exp\left\{\frac{\log\det\left\{\left[\text{var}\left(\hat{\boldsymbol{\beta}}_{\xi_s}\right)\right]^{-1}\right\}-\log\det\left\{\left[\text{var}\left(\hat{\boldsymbol{\beta}}_{\xi_q}\right)\right]^{-1}\right\}}{p}\right\}\end{aligned} \tag{16}$$

Rewriting N_s and N_q in terms of cost ratio k and number of time points for the same total cost using the cost function equation (13), i.e.,

$$N_s = \frac{C}{C_2(k+s)} \quad \text{and} \quad N_q = \frac{C}{C_2(k+q)}$$

we obtain

$$\mathrm{RE}(\xi_s;\xi_q) = \frac{k+q}{k+s}\exp\left\{\frac{\log\det\left\{\left[\mathrm{var}\left(\hat{\boldsymbol{\beta}}_{\xi_s}\right)\right]^{-1}\right\} - \log\det\left\{\left[\mathrm{var}\left(\hat{\boldsymbol{\beta}}_{\xi_q}\right)\right]^{-1}\right\}}{p}\right\} \quad (17)$$

This relative efficiency (17) is for locally optimal design, i.e., for given parameter values. By generalizing this to Bayesian design, the RE of design ξ_s compared to design ξ_q with prior distribution π for $\boldsymbol{\beta}$ becomes as follows:

$$\mathrm{RE}(\xi_s;\xi_q(\pi)\,|\,\pi) = \frac{k+q}{k+s}\exp\left\{\frac{\mathrm{E}_{\boldsymbol{\beta}}\log\det\left\{\left[\mathrm{var}\left(\hat{\boldsymbol{\beta}}_{\xi_s}\right)\right]^{-1}\right\} - \mathrm{E}_{\boldsymbol{\beta}}\log\det\left\{\left[\mathrm{var}\left(\hat{\boldsymbol{\beta}}_{\xi_q}\right)\right]^{-1}\right\}}{p}\right\} \quad (18)$$

Thus, using the Bayesian D-optimality criterion (12), the RE will be:

$$\mathrm{RE}(\xi_s;\xi_q(\pi)\,|\,\pi) = \frac{k+q}{k+s}\left[\exp\left\{\frac{\phi_{\mathrm{D}}(\xi_s\,|\,\pi) - \phi_{\mathrm{D}}(\xi_1\,|\,\pi)}{p}\right\}\right] \quad (19)$$

When the ratio $(k + q)/(k + s)$ is one, that is, if either $q = s$ or the cost ratio k is very large, this relative efficiency (19) becomes the same as the relative efficiency given by Chaloner and Larntz (1989).

Appendix B: BODMixed_Logistic Manual

Introduction

Bayesian Optimal Design for Mixed effects Logistic models with polynomial time effect (BODMixed_Logistic) is graphical user interface software that computes optimal designs for longitudinal studies with a binary response. The program runs in a MATLAB (32-bit version 7.10.0499 (R2010a)) environment. In any case, the program works on a HP Compaq 8200 Elite PC with Windows 7 Enterprise and configuration i5-2400 CPU, 3.1 GHz, 4 GB RAM memory and 64-bit operating system or comparable systems.

To start the program:

1. Start Matlab.
2. Choose the option Window → Workspace → Current Folder and choose the directory where the software is located.
3. Choose Window → Command window and type BODMixed_Logistic (case sensitive) press the ↵ Enter key.

After starting the BODMixed_Logistic program, the user will find the main menu of the BODMixed_Logistic program as shown in Figure 4. There are five panels that will each be explained in turn. In this paper, a tutorial section is included which discusses the various decisions that the user has to make when using the program to find the most efficient design.

First Panel: Input Values of the Model

- *Choose model type*: The user can choose the degree of the polynomial of the mixed logistic model, i.e., a linear (which is the default value), quadratic or cubic model for the trend over time.
- *Variance-covariance parameters (D)*: The user will find a sub-menu to enter the input values for the variances and covariances (matrix **D**) of the random parameters. Figure 5 shows the sub-menu for a quadratic model. A fixed effects logistic model is obtained by setting all values in **D** to zero. The matrix **D** must be specified for each run, i.e. the values of the previous run are not saved.

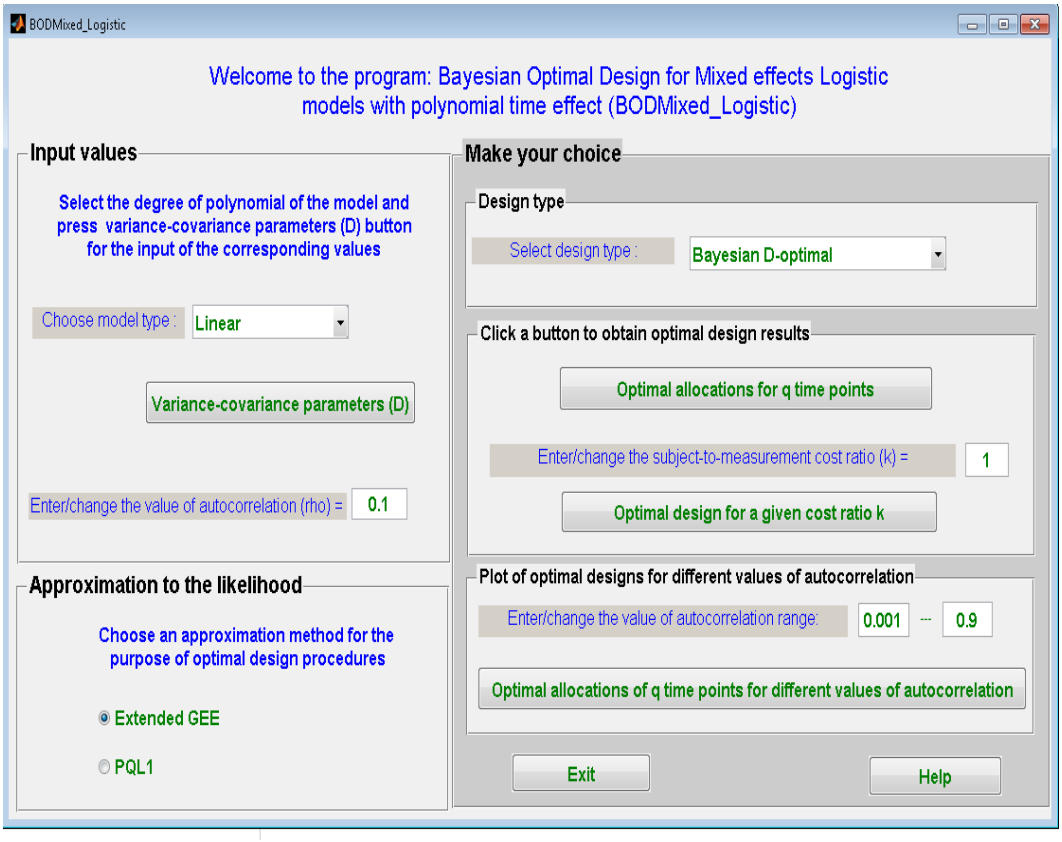

Figure 4. Layout of the main menu with the default input values for BODMixed_Logistic program

---

- *Enter/change the value of autocorrelation (rho)*: This is the size of the autocorrelation coefficient that the user expects between two repeated measurements at a time distance of one, i.e., $\rho = \text{Corr}(y_{ij}, y_{il})$ for which $|t_j - t_l| = 1$, keeping in mind that the total follow-up time is scaled to the interval [-1, +1] so that a time distance of 1 corresponds to half the follow-up time.

Second Panel: Computational Method

- *Approximation to the likelihood*: The user can choose an approximation method for the computation of optimal designs, i.e., either extended GEE or PQL1. The default method is the extended GEE.

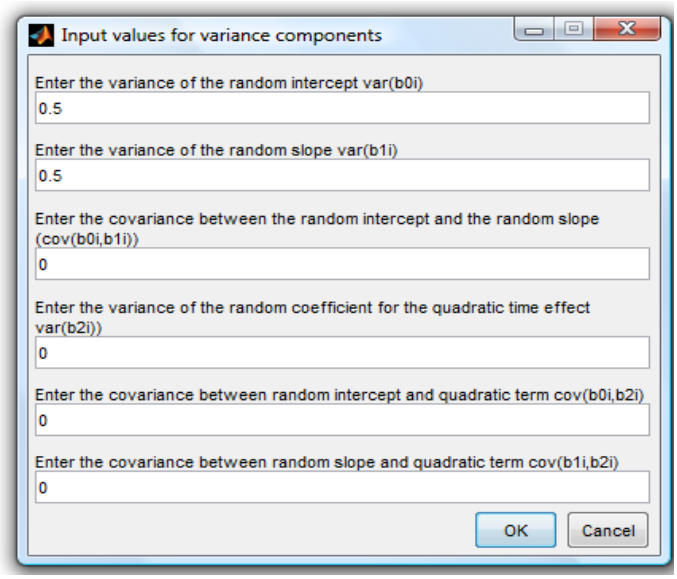

Figure 5. The sub-menu of BODMixed_Logistic for input values for variance components in the **D** matrix for the mixed logistic model with quadratic time effects

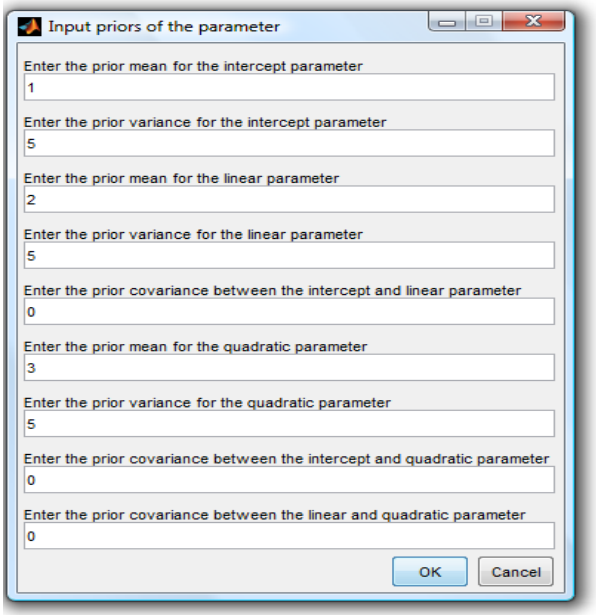

Figure 6. The sub-menu of BODMixed_Logistic for input values for the (normal) priors for the fixed effects parameters of the logistic model with quadratic time effect in the case of Bayesian design

Third Panel: Design Criterion

- *Select design type*: Either Bayesian D-optimal, locally D-optimal, or Maximin D-optimal design. When the user selects a design type, a sub-menu to fill in the input values for the relevant parameters will appear. Figure 6 is an example of a sub-menu for a Bayesian D-optimal design, where the prior means and prior variances can be specified. The input values must be filled in for each run, i.e. the values of the previous run are not saved.

Fourth Panel: Optimal Design Results

In this panel the user can choose between two methods of optimization:

- Fixing the number of time points at some value q to find the optimal allocation of those time points within the time interval [-1, +1],
- or fixing the subject-to-measurement cost ratio and letting the software then find the optimal number of time points as well as the optimal allocation.

- *Optimal allocations for q time points*: A dialog box appears to fill in a specific number of time points q (see Figure 7a). Then, the optimal allocations of time points within the time interval [-1, +1] will be found for the specified number of time points q, and the relative efficiency of equidistant time points compared to the optimal allocation will also be computed

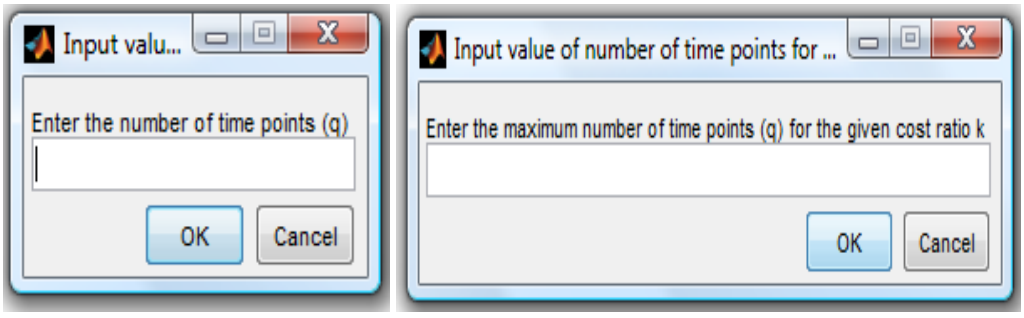

Figure 7. The sub-menu of BODMixed_Logistic to specify the (a) number of time points (q), left, and (b) maximum number of time points, right

- *Enter/change the subject-to-measurement cost ratio (k)*: This is the ratio of the cost of adding a new subject to the cost of an additional measurement per subject. This ratio is assumed to be greater than or equal to zero.
- *Optimal design for a given cost ratio k*: determines the optimal number (q) of repeated measurements as well as the optimal allocation of the q time points for a given subject-to-measurement cost ratio k. The user must specify the maximum allowable number of time points (see Figure 7b). Note that the minimum number of time points is two for a linear, three for a quadratic, and four for a cubic polynomial time effect model. These minima have been implemented in the program already.

Fifth Panel: Plot of Optimal Designs for Different Values of Autocorrelation

The optimal allocations within the time interval [-1, +1] for a given number of time points q can be computed for each autocorrelation value and plotted against the autocorrelation within the range chosen by the user.

- *Enter/change the value of autocorrelation range*: The user can enter a lower and upper bound for the autocorrelation parameter.
- Optimal allocations of q time points for different values of autocorrelation: The user gets the sub-menu of Figure 7a to choose the number of time points (q). Any value with $q \geq p$ can be filled in, where p is the number of fixed parameters of the model (p = 2, 3, or 4 for the linear, quadratic, or cubic model, respectively).

The user can change input values or obtain results by pressing the corresponding buttons on the main menu (Figure 4) as many times as he/she wishes. A ‘Help’ button is also available for guidance. The ‘Exit’ button in the main menu stops the program.