

Efficient Designs for Event-Related Functional Magnetic Resonance Imaging with Multiple Scanning Sessions

MING-HUNG KAO, ABHYUDAY MANDAL,
AND JOHN STUFKEN

Department of Statistics, The University of Georgia,
Athens, Georgia, USA

Event-related functional magnetic resonance imaging (ER-fMRI) is a leading technology for studying brain activity in response to mental stimuli. Due to the popularity and high cost of this pioneering technology, efficient experimental designs are in great demand. However, the complex nature of ER-fMRI makes it difficult to obtain such designs; it requires careful consideration regarding both statistical and practical issues as well as major computational efforts. In this article, we obtain efficient designs for ER-fMRI. In contrast to previous studies, we take into account a common practice where subjects undergo multiple scanning sessions in an experiment. To the best of our knowledge, this important reality has never been studied systematically for design selection. We compare several approaches to obtain efficient designs and propose a novel algorithm for this problem. Our simulation results indicate that, using our algorithm, highly efficient designs can be obtained.

Keywords Compound design criterion; Cyclic permutation; Design efficiency; Genetic algorithms.

Mathematics Subject Classification Primary 62K05; Secondary 62P15.

1. Introduction

ER-fMRI is one of the leading technologies that uses an ultra-fast MR scanner to study human brain activity in response to brief mental stimuli or tasks (e.g., looking at pictures or tapping fingers). This pioneering technology is popular in both medical practice and scientific research, and is arguably the most important advance in neuroscience; see Culham (2006), Josephs and Henson (1999), and Rosen et al. (1998) for overviews of ER-fMRI.

An ER-fMRI design is a sequence of stimuli of one or more types interlaced with a control condition (e.g., rest or fixation). Without a carefully selected

Received December 17, 2008; Accepted January 28, 2009

Address correspondence to John Stufken, Department of Statistics, The University of Georgia, 101 Cedar Street, Athens, GA 30602, USA; E-mail: jstufken@uga.edu

sequence, data collected can easily fail to provide sufficient information for medical investigations or for scientific questions of interest. Designs helping to efficiently render valid and precise statistical inference are therefore very important to avoid wasting precious resources.

However, obtaining good designs for ER-fMRI is an arduous task; it requires careful consideration of statistical models, study objectives, experimental settings and assumptions, and certain practical issues. In addition, the design space, which contains all possible designs, is enormous (e.g., Liu, 2004). Searching over this huge space for good designs is difficult. Moreover, the complex nature of ER-fMRI impels the need for considering multiple, competing objectives in one study. Obtaining designs that efficiently achieve these objectives is necessary, but it increases the complexity of the design problem. Furthermore, answers will hinge on the researcher's interests and experimental conditions. Therefore, we need an efficient, versatile approach that can accommodate a variety of situations to obtain designs best suited to the researcher's needs.

Kao et al. (2009a) proposed an approach to search for efficient ER-fMRI designs. Their approach includes rigorous formulations of statistical models, well-defined multi-objective design criteria, and a search algorithm that incorporates knowledge about the performance of well-known ER-fMRI designs. This approach is shown to outperform methods known hitherto. Designs obtained by their approach can attain much higher efficiencies than designs that had been widely used by practitioners.

While the approach of Kao et al. (2009a) can efficiently yield good designs, they only study experiments with one scanning session. The same applies to other methods in the literature (e.g., Liu and Frank, 2004; Wager and Nichols, 2003). In practice, it can be hard for a subject to maintain a satisfactory performance throughout a long scanning session, and using multiple short sessions is therefore not uncommon. For example, Wang et al. (2007) considered two scanning sessions, each lasting 6 min and 24 s; subjects can rest between sessions. See Brown et al. (2008), Harms et al. (2005), and Harms and Melcher (2002) for other examples. Taking this practical issue into account at the design stage is crucial, but, to our knowledge, there is no previous work that systematically studies this.

In this article, we obtain efficient designs for experiments with multiple scanning sessions. We assume no "pre-scanning" periods so that the stimuli are presented to an experimental subject only within each scanning session; see Sec. 5 for a further discussion. Statistical models, which are natural extensions of widely used models, are formulated, and design criteria for evaluating competing designs are defined. The search algorithm of Kao et al. (2009a) is applied to find designs optimizing these criteria. This approach is compared to various alternative methods. In our simulations, we demonstrate that these alternative methods can perform well for some cases while they can be bad for others. We also propose an algorithm that is restricted to a subclass of designs. The good performance of this algorithm indicates that efficient designs can be found more efficiently over this constraint class.

In the following section, we briefly introduce ER-fMRI designs. Statistical models, design criteria and approaches for obtaining efficient designs are presented in Sec. 3. Simulation results are provided in Sec. 4 and conclusions and a discussion are in Sec. 5.

2. Background

Before conducting an ER-fMRI experiment, a design sequence consisting of stimuli and the control is prepared. This sequence is presented to an experimental subject while the MR scanner scans his/her brain every few seconds. The time interval between successive MR scans is called time-to-repetition or TR. A blood oxygenation level dependent (BOLD) time series is collected by the MR scanner at each brain voxel (a small region of the brain). This time series indicates the fluctuations of the MR signal intensity, which are linked to the change in the ratio of oxygenated to deoxygenated hemoglobin; see, e.g., Lazar (2008) and Cabeza and Kingstone (2006) for details. The BOLD time series is then used to estimate the hemodynamic response function (HRF; a function describing the change of signal intensity over time that is evoked by a single, brief stimulus) and detect brain activation. Estimation of the HRF and detection of brain activity are two common statistical goals for ER-fMRI. We want an optimal sequence of stimuli so that statistical inference (related to estimation and detection) is most efficient, in some sense.

Following convention, designs are presented as a finite sequence of integers. They look like $\xi = \{101201210 \dots 1\}$, where 0 represents the control and q a type- q stimulus; $q = 1, 2, \dots, Q$ (=total number of stimulus types). When being presented to an experimental subject, each stimulus (e.g., a picture) lasts for a short period of time (several milliseconds) relative to the inter-stimulus interval (ISI). The ISI is the fixed time interval between the onsets of consecutive events; an event can be a stimulus or control. We note that 0s in the sequence are “pseudo-events”. They help to calculate the onset times of the stimuli. For example, with a 0 in the second position of ξ , the first three stimuli (1, 1, and 2) occur at 1ISI, 3ISI, and 4ISI seconds after the outset of the experiment, respectively. The control fills up the time period from the end of a stimulus to the start of the next one.

Well-known ER-fMRI designs include block designs, m -sequence-based designs, and mixed designs. In ER-fMRI, a block design is a sequence where stimuli of the same type appear in clusters. For example, a block design can consist of repetitions of $\{111122220000\}$. This design has two stimulus types and the block size is four. Repetitions of $\{1111000022220000\}$ and of $\{11112222\}$ are other possible patterns. Block designs are known to be good for detecting activation.

The m -sequence-based designs are m -sequences (Barker, 2004; Godfrey, 1993) and designs constructed from m -sequences. The m -sequences can be generated from Galois fields or Reed-Muller codes (cf. Ch. 14 of MacWilliams and Sloane, 1977) and only exist if $Q + 1$ is a prime or prime power. They look rather random with no clear pattern, but are not easy to achieve through a random mechanism (Buračas and Boynton, 2002). In terms of A -optimality (see Sec. 3.2), these designs yield high efficiencies when the BOLD time series is assumed to be uncorrelated, exhibits neither drift nor trend, and the objective of the experiment combines interest in estimating individual HRFs and comparing HRFs for different stimulus types. Except for this particular case, m -sequence-based designs can be significantly outperformed by designs obtained by the approach of Kao et al. (2009a), as demonstrated in that article.

While block designs are good for detection problems, they can be very inefficient for estimating the HRFs. Conversely, designs having high estimation efficiency can perform poorly in detecting activation. There is a trade-off between these two statistical goals; see, e.g., Liu and Frank (2004), Liu et al. (2001), and

Buxton et al. (2000). An intermediate efficiency can be achieved by a mixed design. A design of this kind can be formed by concatenating a fraction of a block design with a fraction of an m -sequence-based design (or a random design). By changing the length of the “blocky” part, and hence that of the “random” part, the resulting designs can move toward having high efficiencies for estimation or high efficiencies for detection.

These well-known designs are incorporated in the search algorithm of Kao et al. (2009a) to efficiently and effectively obtain good ER-fMRI designs for a single scanning session. Their algorithm is briefly introduced in Sec 3.3.1. This algorithm is flexible enough to search for efficient designs for multiple scanning sessions. In the next section, we introduce a variation on popular linear models to take session effects into account. We also describe the design criteria and approaches for finding efficient designs under these criteria.

3. Methodology

3.1. Models

Linear models are popular for modeling the BOLD time series and play an important role in studying designs for ER-fMRI (e.g., Kao et al., 2009a; Liu, 2004; Liu and Frank, 2004; Wager and Nichols, 2003; Liu et al., 2001; Dale, 1999; Friston et al., 1999). However, these models do not take into account the reality that experiments commonly involve multiple scanning sessions. We therefore extend the widely used models for estimation and detection by including session effects. This results in models (1) and (2) below for estimation and detection, respectively:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \\ \vdots \\ \mathbf{X}^{(B)} \end{bmatrix} \mathbf{h} + [\mathbf{I}_B \otimes \mathbf{S}] \begin{bmatrix} \gamma^{(1)} \\ \gamma^{(2)} \\ \vdots \\ \gamma^{(B)} \end{bmatrix} + \mathbf{e}; \quad (1)$$

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Z}^{(1)} \\ \mathbf{Z}^{(2)} \\ \vdots \\ \mathbf{Z}^{(B)} \end{bmatrix} \boldsymbol{\theta} + [\mathbf{I}_B \otimes \mathbf{S}] \begin{bmatrix} \gamma^{(1)} \\ \gamma^{(2)} \\ \vdots \\ \gamma^{(B)} \end{bmatrix} + \boldsymbol{\eta}, \quad (2)$$

where \mathbf{Y} is a $T \times 1$ vector of a BOLD time series from a brain voxel, $\mathbf{h} = (\mathbf{h}'_1, \dots, \mathbf{h}'_Q)'$ is the parameter vector for the HRFs of the Q stimulus types, $\mathbf{X}^{(b)} = [\mathbf{X}_1^{(b)} \dots \mathbf{X}_Q^{(b)}]$ is the design matrix for the b th session, $b = 1, \dots, B$ (=the number of scanning sessions), $\mathbf{X}_q^{(b)}$ is the design matrix for the q th stimulus type in that session, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_Q)'$ represents the response amplitudes, $\mathbf{Z}^{(b)} = [\mathbf{X}_1^{(b)} \mathbf{h}_0 \dots \mathbf{X}_Q^{(b)} \mathbf{h}_0]$ is the convolution of stimuli with an assumed basis, \mathbf{h}_0 , of the HRF, $\mathbf{S}\gamma^{(b)}$ is a nuisance term describing the trend or drift of \mathbf{Y} within the b th session, \mathbf{I}_B is the $B \times B$ identity matrix, and \mathbf{e} and $\boldsymbol{\eta}$ are noise. A known whitening matrix, $(\mathbf{I}_B \otimes \mathbf{V})$, is assumed so that $(\mathbf{I}_B \otimes \mathbf{V})\mathbf{e}$ and $(\mathbf{I}_B \otimes \mathbf{V})\boldsymbol{\eta}$ are white noise; i.e., $E\{(\mathbf{I}_B \otimes \mathbf{V})\mathbf{e}\} = E\{(\mathbf{I}_B \otimes \mathbf{V})\boldsymbol{\eta}\} = \mathbf{0}$, $\text{Cov}\{(\mathbf{I}_B \otimes \mathbf{V})\mathbf{e}\} = \text{Cov}\{(\mathbf{I}_B \otimes \mathbf{V})\boldsymbol{\eta}\} = \sigma^2 \mathbf{I}$, where $\sigma^2 > 0$ is unknown and \otimes is the Kronecker product. We assume that each session is of the same length, so that the $\mathbf{X}^{(b)}$ s have the same number of rows. Moreover,

by using the same \mathcal{S} , the trend (or drift) in each session is assumed to be of the same functional shape, but the parameters, $\gamma^{(b)}$ s, are allowed to be different.

Note that when $B = 1$ these models reduce to the ones considered in Kao et al. (2009a). For estimation problems, \mathbf{h} is the parameter of interest, whereas detection problems focus on studying $\boldsymbol{\theta}$. In general, estimable parametric functions $\mathbf{C}_x \mathbf{h}$ and $\mathbf{C}_z \boldsymbol{\theta}$ are investigated. We omit details regarding the construction of each design matrix $\mathbf{X}^{(b)}$ and model parametrization since this is identical as in Kao et al. (2009a); see also Kao et al. (2009b).

3.2. Design Criteria

To evaluate the quality of a design, we consider the A -optimality criterion (e.g., Atkinson et al., 2007); A -optimality aims at minimizing the average variance of estimators of parametric functions $\mathbf{C}_x \mathbf{h}$ and $\mathbf{C}_z \boldsymbol{\theta}$. We formulate these design criteria as “larger-the-better” criteria, which have the form of:

$$\begin{aligned} & r_c \left\{ \text{trace} \left\{ \mathbf{C} \left[\mathbf{W}' (\mathbf{I}_B \otimes \mathbf{V}') (\mathbf{I}_T - \mathbf{P}_{(\mathbf{I}_B \otimes \mathbf{V})(\mathbf{I}_B \otimes \mathcal{S})}) (\mathbf{I}_B \otimes \mathbf{V}) \mathbf{W} \right]^{-1} \mathbf{C}' \right\} \right\}^{-1} \\ & = r_c \left\{ \text{trace} \left\{ \mathbf{C} \left[\sum_{b=1}^B \mathbf{W}^{(b)'} \mathbf{V}' (\mathbf{I}_{T_B} - \mathbf{P}_{V\mathcal{S}}) \mathbf{V} \mathbf{W}^{(b)} \right]^{-1} \mathbf{C}' \right\} \right\}^{-1}. \end{aligned}$$

Here, $\mathbf{W} = [\mathbf{W}^{(1)'} \dots \mathbf{W}^{(B)'}]'$ with $\mathbf{W}^{(b)} \equiv \mathbf{X}^{(b)}$ for estimation problems, and $\mathbf{W}^{(b)} \equiv \mathbf{Z}^{(b)}$ for detecting activation; $T_B = T/B$ and $\mathbf{P}_A = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ is the orthogonal projection matrix onto the vector space spanned by the column vectors of \mathbf{A} , \mathbf{A}^- is a generalized inverse matrix of \mathbf{A} , \mathbf{C} is a matrix of linear combinations of the parameters, and r_c is the number of rows of \mathbf{C} .

We denote this design criterion for estimation by F_e and for detection by F_d . The value of F_e is referred to as “estimation efficiency” and the F_d -value is called “detection power”. It is not uncommon to consider both of these statistical objectives in one experiment. As in previous studies, we consider the following family of multi-objective design criteria (MO-criteria):

$$\{F^* = wF_d^* + (1 - w)F_e^*; w \in [0, 1]\}, \quad (3)$$

where w is a weight selected based on the researcher’s emphasis on the detection problem and F_d^* and F_e^* are standardized criteria:

$$F_i^* = \frac{F_i}{\max(F_i)}, \quad i = d, e.$$

The standardization is used to ensure scale comparability between the two criteria. Note that $\min(F_i) = 0$, which corresponds to the case when some of the parametric functions are non-estimable; see also Kao et al. (2009a).

For a given w , we consider the approaches introduced in the following section to search for designs optimizing F^* .

3.3. Searching for Optimal Designs

3.3.1. *Knowledge-Based Genetic Algorithms.* The algorithm of Kao et al. (2009a) for searching for efficient ER-fMRI designs is based on the genetic algorithm

(GA) technique. GAs (Holland, 1975, 1992) are popular for solving optimization problems, in which good solutions (parents) are used to generate better ones (offsprings). The GA proposed by Kao et al. (2009a) takes advantage of well-known results about good ER-fMRI designs so that the search over the huge design space can be carried out more efficiently.

Their algorithm starts with G initial designs consisting of random designs, an m -sequence-based design, a block design and mixed designs. With probability proportional to fitness (in terms of the objective function), $G/2$ pairs of designs are then selected to produce G offsprings through crossover—randomly selecting a cut-point and exchanging the corresponding fractions of the paired designs. A portion, α_m , of the combined events from the G offspring designs are randomly selected to mutate; these selected events are replaced by randomly generated ones. To help escape local optima and to provide building blocks (or good traits), Kao et al. (2009a) added to the GA population another $\alpha_r \times G$ (α_r is a small fraction) designs drawn from random designs, block designs and mixed designs. The enlarged population is then pruned to maintain a constant population size of G . Only the fittest survive to the next generation. The best solution is kept track of along generations and the process is repeated until a stopping rule is met, e.g., until a pre-specified number of generations has been reached.

The approach for finding efficient designs for our current problem is to use the GA search of Kao et al. (2009a) under the models in that article (i.e., without session effects) and simply split the resulting design into B “short designs”, each corresponding to one session. We will denote this approach by GA-L1. A second possibility is to use the same GA to search for models (1) and (2), so that session effects are included. We denote this approach by GA-L2.

As a third approach, we propose a new GA search over a smaller design space. The idea of this algorithm is to find a “short design” for the first session, and obtain designs for the other sessions by permuting the stimulus types. While other permutations can also be considered, we focus on cyclic permutations. The cyclic permutation is easy to implement and, in our experience, it yields efficient designs. The design for the b th session is constructed by replacing the symbol q in the short design for the first session by $q + b - 1 \pmod{Q}$ (we use Q to represent a residue of zero), $q = 1, 2, \dots, Q$, $b = 1, 2, \dots, B$. For example, with $Q = 3$, stimulus types of 1, 2, and 3 in the design for the first session are replaced, in order, by 2, 3, and 1 for the second session. By juxtaposing the short design and its permuted versions, we obtain a “juxtaposed design” for the entire experiment.

Our algorithm searches for a short design that yields a juxtaposed design maximizing the MO-criterion F^* . We follow Kao et al. (2009a) to incorporate well known ER-fMRI designs. To allow a more effective use of knowledge-based immigrants, we replace the worst parents by immigrants before selecting good parents into a mating pool. This increases the chance to make good use of potential building blocks supplied by immigrants. Our algorithm, which is referred to as GA-L3, is detailed below.

Step 1. (Initial designs) For the first session, generate G initial short designs consisting of random designs, an m -sequence-based design, a block design and mixed designs.

Step 2. (Permutation) For each short design, cyclically permute the stimulus types to create $B - 1$ additional short designs. Form the juxtaposed design and calculate its fitness (e.g., F_e , F_d , or F^* -values).

Step 3. (Immigration) Replace the worst $\alpha_I \times G$ designs in the current generation by immigrants drawn from random designs, block designs and mixed designs. Here, α_I is a small fraction. Calculate the fitness of these immigrants as in Step 2.

Step 4. (Crossover) With probability proportional to fitness, select with replacement G designs and form $G/2$ pairs of distinct parents. Use single-point crossover to generate G offspring designs. That is, randomly select a position in the design sequence and exchange the sub-sequences prior to the selected position of paired parents.

Step 5. (Mutation) Randomly select a portion α_m of the combined events from the G offspring designs. Replace these events by randomly generated ones. As in Step 2, obtain the fitness scores of the resulting G offsprings.

Step 6. (Natural Selection) Out of the $2G$ designs from Steps 3 and 5, keep the best G designs according to their fitness to form the next generation. Discard the others.

Step 7. (Stop) Repeat Steps 3–6 until a stopping rule is met (e.g., after M_g generations). Keep track of the best design over generations.

As do Kao et al. (2009a), m -sequence-based designs or m -sequences are generated following Liu (2004). We include the m -sequence-based design that yields the highest estimation efficiency as one of the initial designs. When these designs are not available, a random design will be used instead.

The initial block design has the highest detection power among designs of different numbers of blocks and of two different patterns. In this pool of candidate block designs, the number of blocks for each stimulus type ranges among 1 to 5, 10, 15, 20, 25, 30, and 40, whenever possible. The two patterns include repetitions of NABC and NANBNC, where N is a block of rests and A, B, and C represent blocks of stimuli of different types.

The combination of a block design with an m -sequence-based design or a random design is obtained through crossover. These mixed designs constitute a portion, e.g., one-third, of the initial designs. The remaining initial designs are formed by random designs.

In Step 3, each immigrant consists of a portion of a block design and a portion of a random design; the relative length of the two parts is randomly decided. When the length for the random part is too short (e.g., less than ten events), a block design is used. Similarly, an immigrant can be a random design. The block-design portion is randomly selected to have one to ten blocks (as possible) and one of the two aforementioned patterns.

When searching for efficient designs, we follow Kao et al. (2009a) to use G (size of a generation) = 20, α_m (rate of mutation) = 0.01, α_I (proportion of immigrants) = 0.2 and stops after $M_g = 10,000$ generations. A larger M_g does not seem to lead to significantly better designs.

Note that, to use an MO-criterion as the objective function in the algorithm, the maxima of F_e and F_d are needed. Theoretical values of $\max(F_e)$ and $\max(F_d)$ are generally not available. They are approximated numerically by the GA using the non-standardized functions F_e and F_d as the objective functions, respectively.

A Matlab program implementing this algorithm can be found in <http://www.stat.uga.edu/~amandal/>.

3.3.2. *Heuristic Approaches.* Some heuristic approaches can also be used. These approaches might provide “short-cuts” for finding efficient designs for the entire experiment. The idea is to find an efficient short design for a single session and, in some way, combine B such designs for B sessions. More specifically, we apply the GA of Kao et al. (2009a) to find a short design that maximizes the design criteria for a single session, namely (for estimation and detection):

$$r_c \{ \text{trace} \{ C [W^{(1)'} V' (I_{T_B} - P_{V_S}) V W^{(1)}]^{-1} C' \} \}^{-1}.$$

The following approaches are then considered for a design for the B sessions:

GA-S1. Use the same short design for each session.

GA-S2. Use different designs for different sessions; these designs, with similar efficiencies, are generated by multiple runs of the GA.

GA-S3. Permute the stimulus types of the design for the first session to obtain designs for other sessions.

For GA-S3, we consider again cyclic permutations. Note that GA-S3 searches for an efficient short design for a single session, and the short design obtained is then used to generate a juxtaposed design for all sessions through cyclic permutations. This is different from GA-L3, which aims for a short design that results in an efficient juxtaposed design with respect to an MO-criterion F^* .

4. Simulations

Through simulations, we demonstrate the performance of approaches introduced in the previous section. These approaches are summarized in Table 1. We focus on $Q = 2, 3,$ and $4,$ and designs of lengths $L = 242, 255,$ and $624,$ respectively. The number of sessions B is set to the number of stimulus types $Q.$ The ISI and TR are both set to two seconds.

Two cases are considered. For models (1) and (2), Case I assumes white noise, i.e., $V = I,$ and S is a vector of ones. In Case II, we assume that, for each session, e and η are stationary AR(1) processes with a correlation coefficient of 0.3, and S corresponds to a second-order Legendre polynomial. While Case I is frequently studied in the literature, Case II is closer to the settings used by practitioners. We consider two parametric functions, which are individual stimulus

Table 1
Algorithms compared in the simulations

Name	Algorithm description
GA-S1	Use the same efficient short design for all sessions
GA-S2	Use different efficient short designs for different sessions
GA-S3	Cyclically permute stimulus types in an efficient short design to create designs for other sessions
GA-L1	Search for efficient designs under models without session effects
GA-L2	Search for efficient designs under models with session effects
GA-L3	Search for the short design that, with its cyclically permuted designs, yields an efficient juxtaposed design for all sessions

effects and pairwise contrasts. Specifically for detection problems, where $C_z \theta$ is of interest, $C_z = I_Q$ for individual stimulus effects, and $C_z = D$ for pairwise contrasts. Here, each row of D corresponds to the difference between θ_{q_1} and θ_{q_2} ; $1 \leq q_1 \neq q_2 \leq Q$. For estimation problems, we use $C_x = C_z \otimes I_k$. The value of k , which is the length of h_q in model (1), is decided by the selected basis h_0 of the HRF for model (2). A common choice of h_0 is the canonical HRF of SPM2 (<http://www.fil.ion.ucl.ac.uk/spm>), a popular software for analyzing fMRI data. The duration of the canonical HRF is 32s, counting from the stimulus onset to the complete return of the HRF to baseline. With both ISI and TR equal to two, the corresponding k is 17 ($=1 + \frac{32}{2}$); see Kao et al. (2009b) for details.

For MO-criteria F^* , we allow the weight w to increase from 0 to 1 in steps of 0.1, thereby gradually shifting emphasis from the estimation problem to the detection problem. A total of 11 designs are generated from each approach for each case. In Fig. 1, we present the F_e^* -values against F_d^* -values for the designs

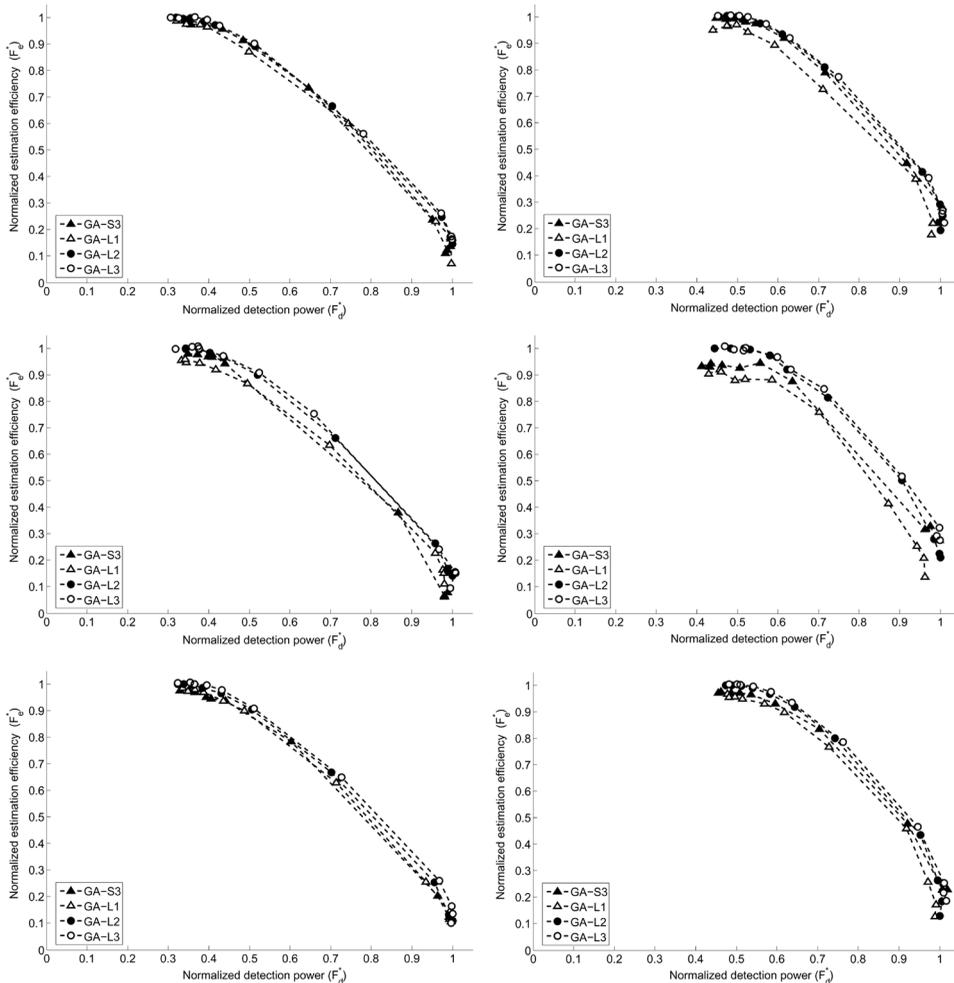


Figure 1. F_e^* -values versus F_d^* -values of designs obtained with $C = I$.

obtained with $Q = 2, 3,$ and 4 and $C = I$. The figure for $C = D$ is omitted since it provides similar information. For clarity, we only include in the figure the designs obtained by GA-L1, GA-L2, GA-L3, and GA-S3 (GA-S3 performs best among the three heuristic approaches). Note that, to calculate F_d^* and F_e^* , we approximate $\max(F_e)$ and $\max(F_d)$ via GA-L2.

In the figure, GA-S3 performs relatively well for $Q = 2$ and 4 . However, the designs generated by this approach are less efficient for $Q = 3$, especially when more weight is assigned to F_e^* . On the other hand, GA-L2 and GA-L3 consistently obtain the most efficient designs. We note that GA-L3 searches over a constrained design space compared to GA-L2. The result indicates that efficient designs can be found over this smaller space. In addition, designs obtained by GA-L1, which ignores session effects, can be inefficient, especially for Case II.

For a single session, the short designs found by GA-L3 are less efficient than those of GA-S3. However, juxtaposed designs obtained from GA-L3 become more efficient than those of GA-S3. To demonstrate this, we present in Table 2 the ratios of F_e -values of the designs obtained from GA-L3 to those from GA-S3. For example, with Case II, $Q = 3,$ and $C_z = D$, the short design found by GA-L3 attains only 69% of the estimation efficiency of that found by GA-S3. However, the juxtaposed design corresponding to GA-L3 reaches 127% of the F_e -value achieved by the GA-S3 design for all sessions.

We implement our simulations by using Matlab (version 7.3) on a Linux cluster with 64-bit AMD Opteron, dual-processor, mix of single-core node and dual-core node; each core has 2GB RAM and the Linux operating system is 2.6.9-78.0.5.ELsmp. The CPU time spent by the six methods for obtaining the 11 designs under different cases are presented in Table 3. Since GA-S1 to GA-S3 obtain designs via different manipulations of short designs, we only present the total time spent for achieving the 11 short designs. For GA-S2, different short designs are generated by assigning different random seeds to start the GA. These designs can be obtained in parallel. As indicated by Table 3 and Fig. 1, GA-L3 uses less CPU time than GA-L2 and still yields efficient designs.

We also note that more CPU time is consumed for most cases when $C_z = D$. The additional time mainly comes from the check for estimability. For $C_z = I$, we need $\mathcal{M} = \sum_{b=1}^B \mathbf{W}^{(b)'} \mathbf{V}' (\mathbf{I}_{T_b} - \mathbf{P}_{VS}) \mathbf{V} \mathbf{W}^{(b)}$ to be nonsingular. For $C = D$, the parametric function can still be estimable when \mathcal{M} is not of full rank. To ensure estimability for the latter case, we examine the equality $\mathbf{D} \mathcal{M}^- \mathcal{M} = \mathbf{D}$. The parametric

Table 2
Ratio of F_e -values (GA-L3 to GA-S3, in percentage)

Q	Design	Case I		Case II	
		$C_z = I$	$C_z = D$	$C_z = I$	$C_z = D$
2	All sessions	103	104	102	107
	Single session	91	96	92	94
3	All sessions	109	124	118	127
	Single session	70	69	64	69
4	All sessions	108	112	109	112
	Single session	97	79	65	80

Downloaded By: [Stufken, John] At: 00:19 11 December 2009

Table 3
Total CPU time spent for obtaining the 11 designs (in hours)

Q		Case I		Case II	
		$C_z = I$	$C_z = D$	$C_z = I$	$C_z = D$
2	GA-S1 to GA-S3*	1.4	1.6	1.4	1.6
	GA-L3	2.4	2.9	2.3	3.0
	GA-L2	3.1	5.9	4.5	5.9
	GA-L1	3.0	5.9	2.9	6.0
3	GA-S1 to GA-S3*	2.3	2.6	2.4	2.7
	GA-L3	3.2	4.2	3.1	4.3
	GA-L2	4.5	10.1	6.8	9.9
	GA-L1	4.1	9.8	4.1	9.8
4	GA-S1 to GA-S3*	3.7	4.7	3.8	4.5
	GA-L3	16.9	19.2	17.4	19.5
	GA-L2	20.3	39.1	31.5	39.5
	GA-L1	19.1	38.8	19.5	38.2

*GA-S2 uses B GA runs and takes B times the presented CPU time if these runs are performed sequentially.

function is estimable if this equality holds (Seber, 1977). Performing this check requires more CPU time.

5. Conclusions and Discussion

In this article, we obtain efficient experimental designs for ER-fMRI when multiple sessions are included in one experiment. We compare the six approaches listed in Table 1 via simulations. In the simulations, we use $B = Q$ (the number of sessions = the number of stimulus types). Similar results are also observed for $B \neq Q$.

The first three approaches that we compared are GA-S1, GA-S2, and GA-S3. These approaches find (short) designs that are efficient for a single scanning session. These short designs are manipulated to form long designs for the entire experiment. GA-S1 uses the same short design for all sessions. For different sessions, GA-S2 utilizes different short designs with similar design efficiencies. GA-S3 cyclically permutes stimulus types in an efficient short design to form new short designs for subsequent scanning sessions.

The other three approaches search for efficient designs for the entire experiment directly. When evaluating competing designs, GA-L1 considers a model with no session effects, whereas the model utilized in GA-L2 is with session effects; both approaches search the entire design space for good designs. GA-L3 is similar to GA-L2 but searches over a restricted design space. The designs considered in GA-L3 are juxtaposed designs formed by a short design and permutations of that design. As in GA-S3, we consider cyclical permutations.

We evaluate the performance of all competing designs in experiments with multiple sessions. Among the approaches being compared, GA-L2 and GA-L3 consistently obtain the most efficient designs. We note that GA-L3 searches over a subclass of designs and it uses less CPU time than GA-L2 to obtain efficient designs.

This indicates that good designs can be found more efficiently by searching over this restricted design space.

In addition, designs obtained under models without session effects (GA-L1) can be less efficient for experiments involving multiple scanning sessions. This is especially true when correlated noise and a drift in the response are assumed.

Among the heuristic approaches, GA-S3 performs best. This approach obtains a design for all sessions by cyclically permuting the stimulus types in an efficient short design for the first session. For some cases, this approach obtains efficient designs, and it uses much less CPU time than GA-L2 and GA-L3. However, the approach does not perform well for other cases, especially when the length of a design for a session is short relative to the number of parameters of interest. Compared to these heuristic approaches, GA-L2 and GA-L3 are more reliable.

In this study, we consider designs with no pre-scanning periods; i.e., no stimuli are presented outside the scanning sessions. This assumption is also made by, e.g., Liu and Frank (2004) and Liu et al. (2001). Alternatively, one can consider pre-scanning periods and present extra stimuli to the subject before each scanning session as does Aguirre (2007). In that article, a long design is divided into short designs for multiple sessions, and the last few stimuli of the previous session are presented again before each session. The last few stimuli of the last session are presented before the first session. Whereas Aguirre (2007) does not aim for designs yielding maximal efficiencies under multiple sessions, we might apply the GA approaches to search for such designs. However, pursuing this is beyond the scope of this article.

For the sake of clarity, we only focus on statistical objectives in this study. In addition to these objectives, Kao et al. (2009a) also considered psychological constraints, and customized requirements. Including these additional objectives in our family of MO-criteria is straightforward. In addition, we can also consider the case when the researcher's interest lies in both individual stimulus effect and pairwise contrasts. As in Kao et al. (2009b), the C_z matrix can then be taken as $C_z = [\delta I, (1 - \delta)D]'$; $\delta \in [0, 1]$.

Acknowledgments

The research of John Stufken was in part supported by NSF Grant DMS-07-06917, and that of Abhyuday Mandal by NSF Grant DMS-09-0573. The authors are thankful to three anonymous referees for their comments and suggestions, which resulted in an improvement of the presentation.

References

- Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *NeuroImage* 35:1480–1494.
- Atkinson, A. C., Donev, A., Tobias, R. (2007). *Optimum Experimental Designs, with SAS*. Oxford: Oxford University Press.
- Barker, H. A. (2004). Primitive maximum-length sequences and pseudo-random signals. *Trans. Instit. Measure. Control* 26:339–348.
- Brown, M. R. G., Vilis, T., Everling, S. (2008). Isolation of saccade inhibition processes: rapid event-related fMRI of saccades and nogo trials. *NeuroImage* 39:793–804.
- Buračas, G. T., Boynton, G. M. (2002). Efficient design of event-related fMRI experiments using m-sequences. *NeuroImage* 16:801–813.

- Buxton, R. B., Liu, T. T., Martinez, A., Frank, L. R., Luh, W. M., Wong, E. C. (2000). Sorting out event-related paradigms in fMRI: The distinction between detecting an activation and estimating the hemodynamic response. *NeuroImage* 11:S457.
- Cabeza, R., Kingstone, A. (2006). *Handbook of Functional Neuroimaging of Cognition*. 2nd ed. Cambridge, MA: MIT Press.
- Culham, J. C. (2006). Functional neuroimaging: Experimental design and analysis. In: Cabeza, R., Kingstone, A., eds. *Handbook of Functional Neuroimaging of Cognition*. 2nd ed. Cambridge, MA: MIT Press.
- Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapp.* 8:109–114.
- Friston, K. J., Zarahn, E., Josephs, O., Henson, R. N. A., Dale, A. M. (1999). Stochastic designs in event-related fMRI. *NeuroImage* 10:607–619.
- Godfrey, K. (1993). *Perturbation Signals for System Identification*. New York: Prentice Hall.
- Harms, M. P., Melcher, J. R. (2002). Sound repetition rate in the human auditory pathway: representations in the waveshape and amplitude of fMRI activation. *J. Neurophysiol.* 88:1433–1450.
- Harms, M. P., Guinan, J. J., Sigalovsky, I. S., Melcher, J. R. (2005). Short-term sound temporal envelope characteristics determine multisecond time patterns of activity in human auditory cortex as shown by fMRI. *J. Neurophysiol.* 93:210–222.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. Ann Arbor, MI: University of Michigan Press.
- Holland, J. H. (1992). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence, Complex Adaptive Systems*. 1st ed. Cambridge, MA: MIT Press.
- Josephs, O., Henson, R. N. A. (1999). Event-related functional magnetic resonance imaging: Modelling, inference and optimization. *Philosoph. Trans. Biolog. Sci.* 354:1215–1228.
- Kao, M. H., Mandal, A., Lazar, N., Stufken, J. (2009a). Multi-objective optimal experimental designs for event-related fMRI studies. *NeuroImage* 44:849–856.
- Kao, M. H., Mandal, A., Stufken, J. (2009b). Optimal design for event-related functional magnetic resonance imaging considering both individual stimulus effects and pairwise contrasts. *Statistics and Applications*. Accepted.
- Lazar, N. A. (2008). *The Statistical Analysis of Functional MRI Data*. New York: Springer.
- Liu, T. T. (2004). Efficiency, power, and entropy in event-related fMRI with multiple trial types: Part II: design of experiments. *NeuroImage* 21:401–413.
- Liu, T. T., Frank, L. R. (2004). Efficiency, power, and entropy in event-related fMRI with multiple trial types: Part I: theory. *NeuroImage* 21:387–400.
- Liu, T. T., Frank, L. R., Wong, E. C., Buxton, R. B. (2001). Detection power, estimation efficiency, and predictability in event-related fMRI. *NeuroImage* 13:759–773.
- MacWilliams, F. J., Sloane, N. J. A. (1977). *The Theory of Error Correcting Codes*. Amsterdam: North-Holland Publishing Company.
- Rosen, B. R., Buckner, R. L., Dale, A. M. (1998). Event-related functional MRI: past, present, and future. *PNAS* 95:773–780.
- Seber, G. A. F. (1977). *Linear Regression Analysis*. New York: Wiley.
- Wager, T. D., Nichols, T. E. (2003). Optimization of experimental design in fMRI: A general framework using a genetic algorithm. *NeuroImage* 18:293–309.
- Wang, Y. P., Xue, G., Chen, C. S., Xue, F., Dong, Q. (2007). Neural bases of asymmetric language switching in second-language learners: An ER-fMRI study. *NeuroImage* 35:862–870.