Model-based analysis of rapid event-related functional near-infrared spectroscopy (NIRS) data: A parametric validation study

M.M. Plichta, a,c,* S. Heinzel, b A.-C. Ehlis, a P. Pauli, c and A.J. Fallgatter a

Department of Psychiatry and Psychotherapy, University of Würzburg, Laboratory for Psychophysiology and Functional Imaging, Fuechsleinstraße 15, 97080 Würzburg, Germany
Theodor-Boveri-Institute, Department of Genetics and Neurobiology, University of Würzburg, Germany
Institute of Psychology, Biological Psychology, Clinical Psychology, and Psychotherapy, University of Würzburg, Germany

Received 1 September 2006; revised 27 October 2006; accepted 16 November 2006
Available online 25 January 2007

To validate the usefulness of a model-based analysis approach according to the general linear model (GLM) for functional near-infrared spectroscopy (NIRS) data, a rapid event-related paradigm with an unpredictable stimulus sequence was applied to 15 healthy subjects. A parametric design was chosen wherein four differently graded contrasts of a flickering checkerboard were presented, allowing directed hypotheses about the rank order of the evoked hemodynamic response amplitudes. The results indicate the validity of amplitude estimation by three main findings (a) the GLM approach for NIRS data is capable to identify human brain activation in the visual cortex with inter-stimulus intervals of 4–9 s (6.5 s average) whereas in non-visual areas no systematic activation was detectable; (b) the different contrast level intensities lead to the hypothesized rank order of the GLM amplitude parameters: visual cortex activation evoked by highest contrast>moderate contrast>lowest contrast>no stimulation; (c) analysis of null-events (no stimulation) did not produce any significant activation in the visual cortex or in other brain areas.

We conclude that a model-based GLM approach delivers valid NIRS amplitude estimations and enables the analysis of rapid event-related NIRS data series, which is highly relevant in particular for cognitive NIRS studies.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Functional near-infrared spectroscopy; NIRS; Event-related; GLM; Two-stage ordinary least square; OLS; Parametric design; Model-based approach; Methods

Introduction

Event-related (ER) experimental designs are useful and often a precondition in examining cognitive processes since they are much more flexible than blocked task paradigms (Burock et al., 1998; Rosen et al., 1998; Pollmann et al., 2000). By them (a) a randomized stimulus order can be presented to the subject providing for a class of paradigms which is “naturally” event-related (e.g. oddball paradigm), (b) trials can be analyzed based on the co-registered behavioral data (e.g. ratings of stimuli, false/correct responses to stimuli) and (c) dynamics of the evoked activity can be analyzed.

Compared to EEG studies of event-related potentials (ERPs), techniques like functional magnetic resonance imaging (fMRI) or functional near-infrared spectroscopy (fNIRS) deal with hemodynamic signals which are inherently slow. The hemodynamic response (HR) evoked by a brief stimulus returns to baseline after 10–12 s or more (Boynton et al., 1996; Buckner et al., 1996) compared to the millisecond range of signals in EEG. As a consequence of these slow signals, stimuli that are separated less than one HR cycle result in a temporal overlap of the recorded HR. This, in turn, hampers a straightforward time-locked averaging procedure as it can be performed in the majority of evoked potential studies (overlap problems in ERP studies are discussed in Woldorf, 1993). Therefore, using sufficiently long inter-stimulus intervals (ISIs) is one solution to prevent overlapping HRs (Buckner et al., 1996) and thereby allows event-related averaging analogously to the analysis of non-overlapping ERPs. However, this approach is inefficient regarding its utilization of experimental time: less stimuli than a complete HR cycle can be presented in a fixed time period which results in a loss of statistical power. Furthermore, uncontrollable confounding states (boredom) or activities (exploration behavior) of the subject might occur during these long delays (Serences, 2004).

In order to perform rapid ER designs which apply ISIs shorter than a complete HR cycle, adequate analysis procedures which handle temporal overlaps of the signals have to be applied. Provided that successive HRs summarize linearly (sufficient linearity of the HRs has been demonstrated with ISIs of 2 s in Dale and Buckner, 1997) and are time invariant, overlapping responses from rapid ER protocols can be analyzed in terms of the general linear model (GLM) by explicitly molding overlapping responses. Therefore, a reasonable hemodynamic response function (HRF) is convolved with the stimulation protocol (delta function) and taken as the predictor for the functional time series.
This model-based analysis approach according to the GLM has become the standard strategy for analyzing functional data in the fMRI domain (Bullmore et al., 1996; Friston et al., 1995; Worsley and Friston, 1995).

Regarding the fNIRS domain, such a standard strategy for data analysis, as first proposed by Schroeter et al. (2004b), has not been fully established until today although there are several similarities to fMRI: the data basis (hemodynamic responses) and the possible experimental designs are highly comparable. Mirroring the development of fMRI research, the traditional fNIRS study designs of the past (block-design protocols) have been recently complemented by ER designs (e.g. Horovitz and Gore, 2004; Izzetoglu et al., 2005; Jasdzewski et al., 2003; Kennan et al., 2002; Plichta et al., 2006a,b; Schroeter et al., 2002). This development will result in a need of an adequate statistical analysis procedure for rapid ER-fNIRS designs. Of course the GLM approach is only one possibility beside many other reasonable analysis approaches, and the particular research question has to dictate the use of either GLM or other approaches. Anyhow, in the event that the GLM approach is adequate for data analysis, the fNIRS community can profit from the rich experiences of the fMRI domain and utilize the general framework of a model-based analysis approach. However, until today rapid ER-fNIRS studies which employ a GLM for data analysis are rare. Izzetoglu et al. (2005) performed a model-based ER analysis of fNIRS data where the ISIs were shorter than an HR cycle. In that study, participants had to solve anagrams while frontal brain activation was recorded by multi-channel fNIRS. The results indicate a parametric modulation of the fNIRS response due to the task difficulty. However, only oxyhemoglobin (O₂Hb) data are described. No results are reported in terms of the regional specificity of the hemodynamic response and, thus, systemic effects cannot be excluded (Francescini et al., 2003). Furthermore, the stimulus order was not randomized and no null control (i.e. no stimulation condition) was included. In another study of Boas et al. (2003), an event-related motor paradigm was conducted where the mean ISI was 8 s (range 2 s–3 s) and a GLM approach was applied. However, a validation of the estimated beta weights, e.g. by conducting a parametric experimental design, was not the scope of the latter study.

The aim of the present work is to test if the GLM approach is capable of detecting significant human brain activation in a rapid ER-fNIRS design (replicating the findings of Izzetoglu et al., 2005) with a randomized stimulus order and, more crucially, if the GLM estimates of the amplitudes are valid regarding their relative height and correspond to varying stimulus intensities. A simple checkerboard paradigm was chosen because of its strong and robust activation effects in the primary visual cortex accessible with fNIRS (Plichta et al., 2006b). A parametric design with four levels of visual stimulus intensities is conducted. The visual stimuli vary regarding their luminance contrast levels (0%, 8%, 40% and 97%) and were chosen based on findings from fMRI studies (Advan et al., 2002; Boynton et al., 1999; Heeger and Ress, 2002; Tootell et al., 1998) which have demonstrated that an increase of stimulus contrast levels results in a monotonic increase of the primary visual cortex (V1) activity.

Methods

Subjects and stimulation protocol

A total of 15 subjects were examined (mean age = 25.3 ± 2.6 years). We included both males and females (7 males; 8 females) regardless of hair color or handedness. All subjects had normal or corrected to normal vision. No subject had a history of any neurological disorder. All subjects were informed about the nature of the experiment as well as the operating mode of the NIRS instrument, before giving written informed consent. A brief instruction to remain relaxed and to avoid any major body movement was given. The NIRS investigation of healthy participants was in accordance with the latest version of the Declaration of Helsinki and was approved by the Ethics Committee of the University of Wurzburg as well as the research conference of the Department of Psychiatry and Psychotherapy.

Visual stimulation

The visual stimulation was realized in an event-related paradigm by presenting a series of checkerboards, each for 2 s, reversing in contrast at 8 Hz (according to Ozus et al., 2001) followed by an ISI of an uniform gray color screen presentation of variable duration (ISI= 4–9 s, average: 6.5 s). Michelson contrast ratios of the checkerboards, which are given by:

\( \frac{(L_{\text{max}} - L_{\text{min}})}{(L_{\text{max}} + L_{\text{min}}) \times 100} \)  
(1)

(with \( L_{\text{max}} \) = luminance of the light element and \( L_{\text{min}} \) = luminance of the dark element) were varied as follows: 97% (high contrast), 40% (medium contrast) and 8% (low contrast). The ISI screen was set to 0% contrast. Additionally, the protocol contains 30 null-events with 0% contrast (duration and ISI identical to checkerboard presentation) enabling the calculation of event-related averages by subtracting the average of null-events from the average of real stimulation trials (Busse and Woldorf, 2003). This possibility allows further validation of the GLM results by comparing them to a model-free analysis approach.

Constant over all contrast levels, the check size was 0.50° (according to Miki et al., 2001) which corresponds to a spatial frequency of 1 cycle/degree (lowest frequency). The stimuli were presented on a 17” monitor with a resolution of 800 × 600/75 Hz, and the total field of stimulation was 24.4° × 31.9°. The mean luminance was held constant at 75 cd/m² across all stimuli and the ISI screen (measured with the luminance meter MAVO-MONITOR, Gossen, Germany). A central fixation cross was constantly presented during the whole experiment to avoid major eye and head movements.

Number of trials was set to \( n = 30 \) for each contrast level. The order of stimuli was random with the only constraint of not consecutively presenting more than two equal contrast grades in order to prevent adaptation.

Subjects were instructed to focus on the fixation cross. After each stimulus presentation, subjects were asked to report which contrast level they had perceived using a standard PC keyboard. This procedure enables to hold and control the subjects’ attention and to verify correct attribution of the contrast levels. Total measurement time was 16 min.

Functional near-infrared spectroscopy

The fundamentals of fNIRS are described in detail elsewhere (e.g. Hoshi, 2003; Obriag and Villringer, 2003). The continuous wave system (ETG-4000, Hitachi Medical Co., Japan) is described in detail in Plichta et al. (2006b). The interoptode distance was 30 mm and the sampling rate was set to 10 Hz.
Functional localizer and region of interest (fROI)

Initially, the probe set of 52-channel fNIRS (see Fig. 1) was placed on the scalp with its lowest-row center optode at the subjects’ inion. The validity of the probe set positioning in detecting visual activation was then verified by performing 3 cycles of a block-design checkerboard stimulation (97% contrast; 10 s stimulation; 20 s rest) prior to the main experiment. The activation patterns’ position was inspected and, whenever necessary, the probe set was re-localized and another 3 cycles of the block-design stimulation were performed until the activation pattern was located centrally (in vertical direction) and symmetrically.

Following the above described procedure, the ROI definition was based on the results described in Plichta et al. (2006b) wherein the identical NIRS system, probe set dimension and a similar paradigm were applied. In this previous work, the second level results indicate that the visually evoked fNIRS activation pattern can be covered optimally by channels #15, #17, #25–#28, #36 and #38 if the pattern is located centrally (in vertical direction). Because NIRS lacks an anatomical image, the above described functional localization procedure ensures (as far as possible) a valid second level analysis with regard to the orientation of the activation pattern (but not its spatial extent).

Because spatial resolution of fNIRS is poor with a non-overlapping geometrical arrangement of optodes (no better than the emitter–detector distance of 30 mm) and no anatomical image was taken, the non-ROI was limited to channels outside the ROI minus the directly adjacent channels of the ROI (compare Fig. 1). The remaining channels are excluded from further analyses.

Functional data analysis

For analyzing the fNIRS time series GLM is applied. The GLM approach has been extensively described in fMRI literature (see Friston et al., 1995; Worsley and Friston, 1995; Bullmore et al., 1996). Briefly, the data matrix \( Y \) of order \( (T \times C) \) containing the functional NIRS time series \( T \) of each channel \( C \) is predicted by \( X \) consisting of a set of reasonable hemodynamic response functions (HRFs) which are convolved with the event sequence (the order of \( X \) is \( (T \times M) \) where \( M \) is the number of modeled effects—see below). The functional data can be modeled as:

\[
Y = X\beta + \epsilon
\]  

(1.1)

where \( X \) is the design matrix and \( \beta \) is the parameter matrix. In the simplest case, each column \( M \) of matrix \( X \) contains the predicted hemodynamic response for one experimental condition over time \( T \). To address inter-individual differences regarding the HRF’s latency and dispersion, the inclusion of the HRF’s first and second temporal derivative has been proposed (Friston et al., 1998). The inclusion of derivative terms results in an extension of the design matrix \( X \): for one experimental condition, \( X \) contains two (HRF+1. derivative) or three columns (HRF+1. derivative+2. derivative).

The ordinary least square (OLS) estimates of \( \beta \) are given by:

\[
\beta = (XX)^{-1}XY
\]  

(1.2)

The \( \beta \)-weights quantify the contribution of a predictor (e.g. HRF) for explaining the functional time series \( Y \) and serve as the parameter set for subsequent hypothesis testing. At the single subject level \( t \)-tests can be applied to an estimated beta weight (testing e.g. \( H_0: \beta = 0 \)) or the beta weights of a sample are collected and analyzed at the group level (e.g. with paired \( t \)-tests, analysis of variance etc.). Testing the beta weights (e.g. by one sample \( t \)-tests) gives an answer to the question whether a particular brain area is activated by the experimental condition. For models incorporating derivative term(s), the amplitudes are estimated from the non-derivative term only (Friston et al., 1998) based on the assumption that “error” variance caused by inter-individual latency or dispersion differences will be explained by the derivative terms (see Calhoun et al., 2004 for a critique of this assumption).

The unbiased estimates of the significance of \( \beta \)-weights (=\( t \)-values) are based on the assumption that the error term \( \epsilon \) (see formula (1.1)) is uncorrelated (representing an identity matrix \( I \)), independent and normally distributed:

\[
\epsilon \sim \text{i.i.d.}(0, \sigma^2 I)
\]  

(1.3)

Since several physiological processes (respiration, blood-pressure changes, heartbeat) are known to produce structured “noise” to the data (autocorrelation), a common strategy to deal with this problem is to calculate OLS estimate(s) first (see formula (1.1) and (1.2)) and fit an AR(\( p \)) model (autoregressive model of the order \( p \)) to the resulting residuals (Cochrane and Orcutt, 1949). This leads to a decomposition of the error term \( \epsilon \) into a systematic part as well as into the model conform error part. After this, the AR transformation coefficient is applied to both sides of the regression equation:

\[
Y_t - \rho Y_{t-1} = \beta_0 (1 - \rho) + (X_t - \rho X_{t-1})\beta + \eta_t
\]  

(1.4)

Fig. 1. Schematic representation of the 52-channel NIRS probe set. White squares are light emitters; black squares represent detectors; and numbers represent the measurement channels. The lowest-row center optode is indicated by an asterisk. The ROI is indicated by the dark gray shaded areas in the middle of the probe set (channels: 15, 17, 25–28, 36 and 38). The non-ROI channels are shown within light gray shaded areas (left and right side of the probe set). The aim of the functional localization procedure (see Methods section) was to get the individual activation spots congruent to the predefined ROI which was chosen based on the findings described in Plichta et al. (2006b).
Fig. 2. Panel A shows the second-level overall effect of the checkerboard stimulation (average beta weight of 8%, 40% and 97%) overlaid on a standard brain (the overlay is an approximation and is not based on a co-registration) for O2Hb and HHb, respectively. Panel B shows the second-level results evoked by the three different stimulation conditions (contrast level: 8%, 40% and 97%). Confirming the adequacy of the choice of channels comprising the ROI (see Methods section), the measured activation within the ROI is significantly higher than in the non-ROI (O2Hb: $t=5.99; df=13; p<.001$; HHb: $t=-4.42; df=13; p<.001$, see panel C for O2Hb and panel D for HHb). Furthermore, the activation measured in the ROI is significantly different from zero in both NIRS parameters (O2Hb: $t=6.54; df=13; p<.001$; HHb: $t=-4.32; df=13; p<.001$). In contrast to the absence of a significant activation within the non-ROI in HHb ($t=-1.36; df=13; n.s.$), activation was detectable within the non-ROI in O2Hb data ($t=2.24; df=13; p<.05$). Corresponding to the activation maps shown in panel B, panels E and F represent the differential activation effect due to location (REG) and stimulation intensity (CON). Note that the scale of HHb was inverted to simplify matters and that the activation maps (A and B) are based on interpolations from single channels.
\[ Y^* = Y_t - \rho Y_{t-1} \]

\[ X^* = X_t - \rho X_{t-1} \]

\[ Y^* = \beta_0 + \beta_1 Y_{t-1} + \epsilon_t \] (1.5)

As a result, the serial correlation is reduced.

Prior to the GLM analysis, the functional data were pre-processed by applying a low-pass filter (cut-off frequency of 0.7 Hz). Thereafter, GLM is applied by using a gauss function as HRF (peak time = 6.0 s; Full Width Half Maximum (FWHM) = 5.89) and its first and second temporal derivative to modulate the onset and the dispersion of the HRF.

The HRF was used for both NIRS parameters (O$_2$Hb and HHb). Thus significant positive beta weights indicate activation in the O$_2$Hb data while significant negative beta weights indicate activation in the HHb data. An autoregressive process of order 1 is applied by default.

After performing the GLM analysis, two separate 4 × 2 repeated analyses of variance (ANOVA) are applied to the resulting beta weights of [O$_2$Hb] and [HHb] at the second level. The applied ANOVAs comprise two within factors: parametric factor contrast (CON) with four levels (highest, moderate, lowest contrast, no stimulation) and dichotomous factor region (REG) with two levels: ROI and non-ROI.

Post hoc tests were performed by one-tailed t-tests (alpha 0.05; Bonferroni–Holm corrected) whenever directed a-priori hypotheses about the differences are available (rank order comparisons in the predefined ROI: highest contrast > moderate contrast > lowest contrast > no stimulation). Two-tailed t-tests and an uncorrected alpha level of 5% were chosen whenever no directed hypotheses are available a-priori (i.e. rank order differences in the non-ROI).

Comparison of model-free and model-based results

To evaluate the adequacy of our applied GLM for analyzing the functional NIRS data, two strategies were followed. First, the results of the model-free averaging procedure are descriptively compared with the GLM results. Therefore, the amplitudes of the event-related averages are extracted at second-level peak time and compared with paired t-tests regarding ROI and non-ROI (as in the GLM post hoc analyses with the only exception of not analyzing the null-events since they were subtracted from the individual averages to obtain an overlap-reduced average). Second, the original model estimates are compared with two alternative GLM

\[ \beta X^* + \epsilon_t \]

Table 1

<table>
<thead>
<tr>
<th>Contrast</th>
<th>O$_2$Hb</th>
<th>HHb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_{d=13}$</td>
<td>$p$</td>
</tr>
<tr>
<td>97%&gt;40%</td>
<td>1.83</td>
<td>.05</td>
</tr>
<tr>
<td>97%&gt;8%</td>
<td>4.08</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>40%&gt;8%</td>
<td>2.38</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>97%&gt;0%</td>
<td>7.46</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>40%&gt;0%</td>
<td>7.23</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>8%&gt;0%</td>
<td>5.43</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

variants: HRF without derivatives (variant I) and HRF with its first temporal derivative to modulate the onset of the hemodynamic response (variant II). Model evaluation was based on two criteria: (1) correlation between the model-based amplitudes and the model-free amplitudes resulting from the second-level analyses of the overall effect of stimulation (quantitative criterion) and (2) inspection of the sensitivity and false alarm rate of the three model variants in detecting significant activation inside or outside the ROI (qualitative criterion). Therefore, the correlation coefficients were z-transformed and tested against each other (dependent samples; alpha=0.05 (uncorrected); cf. Sachs, 1999).

Results

Behavioral data

All presented checkerboard stimuli were followed by button presses in each subject, indicating that the subjects remained attentive during the whole experiment. Subjective ratings by the participants indicate that the chosen contrast levels were clearly distinguishable intra- and inter-individually: the entire set of checkerboards was correctly classified by 11 subjects, one misclassification (error rate = 0.6%) occurred in one subject and two misclassifications (error rate = 1.2%) were made by two subjects.

The averaged time of the button presses was 623 ms (SD = 180 ms) after the stimulus disappeared and did not significantly differ across the different stimulation levels (from 8% to 100%, 620 ms (SD = 190 ms), 609 ms (SD = 176 ms) and 642 ms (SD = 194 ms); F(2,28) = 2.27; n.s.).

O$_2$Hb

Analysis by repeated measurement ANOVA revealed a highly significant main effect of factor CON ($F(1,13)=30.91; p<.001$; eta$^2=.70$). Regardless of the contrast level, channels comprising the ROI show significantly higher concentration changes than channels of the non-ROI (compare Fig. 2C). The main effect of factor CON was also significant ($F(3,39)=22.77; p<.001$; eta$^2=.64$). Testing the studies’ crucial hypothesis, the analysis of CON × REG revealed a highly significant interaction effect ($F(3,39)=35.17; p<.001$; eta$^2=.73$).

Post hoc analysis of the mean differences by paired t-tests showed that all four CON levels are significantly different from each other within the ROI ($p<.05$, corrected). Furthermore, these significant differences mirror the stimulus input function: highest

---

1 One subject had to be excluded from analysis because of problems with localising the visual cortex.
Fig. 3. Event-related grand averages and model-based parameter estimations within the ROI. Panel A shows the O2Hb data and panel B shows HHb data. On the left side of each panel, the event-related average time course is shown along with the beta-weighted HRF plus its first and second (weighted) derivatives. On the right side of each panel, similarity of the model-free and the model-based amplitudes is shown with scatter-plots and determination coefficients (the amplitudes were z-transformed). ROI channels are labeled—the remaining channels belong to the non-ROI plus directly adjacent channels (compare Fig. 1). The scales for HHb are consistently inverted to simplify matters.
contrast>moderate contrast>lowest contrast>no stimulation (see Table 1 and Fig. 2E).

Regarding non-ROI, none of the three “real” stimulation CON levels was significantly different from each other (all p-values > .05, uncorrected). However, the real stimulation levels differed significantly from the no stim (0% contrast) condition (see Table 2 and Fig. 2E).

\[ \text{HHb} \]

Repeated measure ANOVA revealed significant mean effects for factor REG \((F(1,13)=24.43; \ p<.001; \ \text{eta}=.65)\) as well as for factor CON \((F(3,39)=18.38; \ p<.001; \ \text{eta}=.58)\). Mirroring the O2Hb results, the ROI exhibits significantly more activation than the non-ROI (compare Fig. 2D). Furthermore, the interaction of CON x REG is highly significant \((F(3,39)=21.51; \ p<.001; \ \text{eta}=.62)\).

Confirming the studies’ main hypothesis, post hoc t-tests of the HHb data demonstrate that all CON levels were significantly different from each other in the predefined ROI \((p<.05,\ corrected)\)—except the contrast 97% vs 40%. Similar to the O2Hb results, the significant differences mirror the stimulus input function: highest contrast>moderate contrast>lowest contrast>no stimulation (see Table 1 and Fig. 2F).

No significant differences between the contrast levels (97%, 40%, 8%) were detectable in the non-ROI (all p-values > .05, uncorrected—see Table 2). Two of the contrast levels (97% and 8%) differed significantly from the “no stim” condition (see Table 2 and Fig. 2F).

Comparison of GLM results and event-related averages

Figs. 3a and 3b show the model-free averages in the ROI for O2Hb and HHb along with the GLM results. Peak times of the 97%, 40% and 8% stimulation were 6.6s, 6.5s and 6.5s (O2Hb) and 6.7s, 6.8s, and 6.7s (HHb) and do not considerably differ across the conditions. The latter result is important to rule out that the model fit varies due to different latencies rather than different amplitudes evoked by different stimulation intensities (see Calhoun et al., 2004) – however, compared to the event-related averages the model peaks slightly but consistently too early (about 0.5s). The amplitude rank order of the evoked responses (highest contrast > moderate contrast>lowest contrast) is identical to the GLM amplitude estimates. Furthermore, the absence of a significant difference between 97% and 40% in HHb as indicated by the GLM results is mirrored by the event related averages (compare Fig. 3B).

Table 3 and 4 show the amplitude comparisons of the model-free approach. In line with the GLM approach, the model-free analyses of the O2Hb data show a significant difference for the 97% vs 40% and the 97% vs 8% ROI contrast \((p<.05,\ corrected)\). However, no difference was detectable in the 40% vs 8% contrast.

Table 3

<table>
<thead>
<tr>
<th>Contrast</th>
<th>O2Hb (T_{df=13})</th>
<th>p</th>
<th>d</th>
<th>HHb (T_{df=13})</th>
<th>p</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>97%&gt;40%</td>
<td>2.06 &lt; .05 0.78</td>
<td></td>
<td></td>
<td>−0.69 n.s. 0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>97%&gt;8%</td>
<td>3.01 &lt; .01 1.14</td>
<td></td>
<td></td>
<td>−3.39 &lt; .01 1.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40%&gt;8%</td>
<td>0.18 n.s. 0.08</td>
<td></td>
<td></td>
<td>−2.52 &lt; .05 0.95</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the non-ROI, the 40% stimulation was significantly higher than the 8% stimulation \((p<.05,\ uncorrected)\). The other contrasts were not significantly different.

Regarding HHb, the 97% vs 40% difference is not significant whereas the other contrasts (97% vs 8% and 40% vs 8%) yield significant results, which is in line with the GLM results. The effect size parameters of the GLM and the model-free approach are of comparable height. Regarding the non-ROI, there are considerable differences between the GLM and the model-free results. None of the contrasts was significantly different from each other.

Post hoc evaluation of model adequacy

Table 5 shows the correlation coefficients of the model-free grand average amplitudes with (a) the amplitude estimates resulting from the original model (including two derivative terms); (b) a model variant without derivative terms and (c) a model variant with only the first derivative (quantitative criterion). Significant differences were apparent between variant I and variant II \((t=−3.22, \ p<.01)\) but not between variant I and the original model \((t=−1.15; \text{n.s.})\) or variant II and the original model \((t=1.06; \text{n.s.})\).

Furthermore, Table 5 shows the percentage of significantly activated channels inside and outside the ROI plus the directly adjacent channels (qualitative criterion). While none of the compared models leads to false positives in the non-ROI, the amount of significant channels is largest in the original model \((n=11)\) compared to variant I \((n=5)\) and variant II \((n=9)\) if O2Hb is considered – but a similar amount of significant channels in the

Table 5

<table>
<thead>
<tr>
<th>NIRS parameter</th>
<th>Quantitative</th>
<th>Qualitative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original model</td>
<td>Oxy .981</td>
<td>75.0%/31.2%/0%</td>
</tr>
<tr>
<td></td>
<td>Deoxy .926</td>
<td>50.0%/6.2%/0%</td>
</tr>
<tr>
<td>Variant I</td>
<td>Oxy .975</td>
<td>50.0%/6.2%/0%</td>
</tr>
<tr>
<td></td>
<td>Deoxy .922</td>
<td>50.0%/12.5%/0%</td>
</tr>
<tr>
<td>Variant II</td>
<td>Oxy .985</td>
<td>62.5%/25.0%/0%</td>
</tr>
<tr>
<td></td>
<td>Deoxy .926</td>
<td>50%/12.5%/0%</td>
</tr>
<tr>
<td>ER-average</td>
<td>Oxy –</td>
<td>100%/93.7%/46.5%</td>
</tr>
<tr>
<td></td>
<td>Deoxy –</td>
<td>100%/75.0%/35.5%</td>
</tr>
</tbody>
</table>

\[ a \] Based on the bivariate correlation of the GLM estimates and the model-free peak amplitudes (across 52 channels).

\[ b \] First value represents the percentage of channels defined as ROI which reach significance in the particular model; second value represents the percentage value of channels directly adjacent to the ROI (compare Fig. 1, white area) which reach significance; third value represents the percentage value of significant channels in the non-ROI. The alpha level was set to 0.05 (Bonferroni-corrected).
The last two rows of Table 5 show the qualitative analysis of the event-related average approach. If a Bonferroni-corrected alpha level of 5% is applied, the model-free approach leads to a substantial amount of false positives (i.e. activation in the non-ROI). 46.5% (O2Hb) and 35.5% (HHb) of the activated channels were located in the non-ROI, whereas no significant activation was seen if the GLM approach is applied.

Discussion

The aim of the present study was to examine if a model-based time series approach as performed in fMRI ER data analysis delivers estimates of fNIRS amplitudes which mirror a known rank order of differently graded stimulation intensities. Therefore, a rapid ER approach according to the GLM is capable of detecting event-related human brain activity recorded with fNIRS in the occipital cortex with ISIs of 4–9 s (6.5 s average). This can be regarded as a replication of the findings in Schroeter et al. (2004a) who showed fNIRS activation detection during a Stroop task with ISIs of 12, 6, 4 and 2 s.

More importantly, the GLM estimates of the amplitudes mirror the stimulus input function (highest contrast>moderate contrast>lowest contrast>no stimulation) in both NIRS parameters. Limited to the visual system, this clear-cut rank order indicates that higher (absolute) beta values correspond to a higher cortical activation within the same brain structure. The credibility of the findings is underpinned by the complementary results of both NIRS parameters. Furthermore, the findings are highly consistent with former fMRI studies (Avidan et al., 2002; Boynton et al., 1999; Heeger and Ress, 2002; Tootell et al., 1998) which contributes to the validity of fNIRS.

Beside the confirmation of the study’s main hypothesis, two other important findings have to be highlighted: (1) by using a multi-channel fNIRS device, it was possible to observe an extensive cortical area which includes brain structures outside the visual cortex. This enables to verify if the pattern of the presumed amplitude rank order is limited to the visual cortex. Indeed, outside the ROI no systematic activation was detectable (compare Figs. 1E and F) indicating the absence of a global (regionally unspecific) activation effect; (2) by including null-events and comparing the “evoked” activation to the “real” contrast levels, it was possible to further validate the GLM amplitude estimations. For both O2Hb and HHb, the “no stim” condition leads to the smallest beta values (see Figs. 2E and F). Furthermore, by comparing “no stim” activation with real stimulation activation outside the ROI (non-ROI), our results show that signal changes occur whenever stimuli of any kind are presented (compares Figs. 2C and D), which may indicate the presence of systemic effects. Herein, our results of a significant O2Hb non-ROI (see Fig. 2C) effect but non-significant HHb non-ROI effect (see Fig. 2D) are in line with the finding that O2Hb may be more altered by systemic changes than HHb (Franceschini et al., 2003). Another possibility is that the motor response differently affected channels nearby left and right motor cortices because different fingers were moved for the three different responses (motor activation would correspond to an activation in channel #1, #11–#12 and #10, #20–#21—see Plichta et al., 2006a). However, we re-calculated the analyses and excluded these channels with no substantial change of the results. Since no anatomical image is available, a final answer to this question is not possible.

Limitations of the present study

In the present study, we used objectively different contrast levels which were not individually fitted to the sensoric potencies of the particular subject. That approach can be discussed critically based on the findings of Ress and Heeger (2003) who demonstrated that V1 activity evoked by a visual stimulus corresponds to the subjects’ percept rather than on the objective appearance. However, due to the results of the collected behavioral data, it is highly likely that the chosen contrast levels led to clearly distinguishable percepts across all subjects. Furthermore, the scope of the present study was to hold the objective stimulation intensity constant across all subjects to examine the inter-individual concordance regarding the rank order of activation amplitudes.

From a methodological point of view, the present study starts with an a-priori assumed HRF, where the parameters are taken from existing evidence. The shape, peak time and latency are not extracted from the actual sample nor are different parameters systematically tested because the present study is not designed for an extraction of overlap-unbiased grand averages. Even though it has been shown that random ISIs plus null-events (as applied in the present study) prevent a bias due to overlapping responses, the amount of bias in the calculated grand averages remains unknown. The grand averages may be affected by an omitted stimulus response (Busse and Woldorff, 2003) and thus cannot be used as a “gold standard” for the model-based analysis approach. Therefore, our comparison of different models should be considered with caution because it is a post hoc view and thus afflicted with the risk of choosing a model variant which fits to idiosyncratic sample features. Furthermore, the model-based amplitude estimations and the grand average results are in part incommensurable since the model-based approach includes a serial correlation correction whereas the grand averages remained uncorrected in our analyses. No systematic study exists wherein different orders of AR processes are tested against each other.

These are limiting factors for a full numerical comparison of both analysis strategies. However, a full numerical comparison was not the scope of the present study. Future studies which compare the model-based and the model-free approach are needed. These studies should include overlap-unbiased grand averages, which implies (1) the use of long ISIs (15–20 s or even longer, if the NIRS undershoot phenomenon (Schroeter et al., 2006) is also taken into account) without producing the disadvantageous effects (see introduction), (2) a systematic variation of the model parameters; different parameters for oxy- and deoxy-hemoglobin have to be tested (see Huppert et al., 2006; Jasdzewski et al., 2003) and (3) a systematic comparison of different autoregressive model orders and the comparison to unbiased grand averages.

Conclusion

The present study shows that the GLM framework of statistical analyses as applied in the fMRI domain can be expanded to the fNIRS domain. The GLM approach delivers valid fNIRS amplitude estimations and enables the analysis of rapid event-related fNIRS data series, which is highly relevant in particular for cognitive fNIRS studies. Moreover, the effective application of a
GLM based analysis in the present study may facilitate a straightforward and intuitive understanding of fNIRS results for researchers and practitioners who are familiar with fMRI interpretations. If future studies can demonstrate further indications for the validity of the model-based GLM approach, it may help to accelerate establishing GLM as proposed by Schroeter et al. (2004b) as a useful additional tool in analyzing fNIRS data, in particular resulting from rapid event-related designs.

Acknowledgments

The authors would like to thank: Hitachi Medical Corporation for the ETG-4000 equipment; Thomas Meigen for letting us use his MAVO-MONITOR®. Thomas Kammer for his helping comments regarding the visual stimulation; Antje B.M. Gerdes for the fruitful discussion of the experimental design; Andreas J. Bartsch, Vínce D. Calhoun and the staff of the Bender Institute of Neuroimaging (B.I.O.N.), University of Giessen (director: Prof. Dr. Dieter Vaitl) for generously sharing their analysis expertise.

This study was supported by the Deutsche Forschungsgemeinschaft (KFO 125/1-1).

References


