NeuroImage 43 (2008) 521–527

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Technical Note

An objective assessment of fetal and neonatal auditory evoked responses

R.B. Govindan [a,*], J.D. Wilson [b], H. Preissl [a,c], P. Murphy [a], C.L. Lowery [a], H. Eswaran [a]

[a] Department of Obstetrics and Gynecology, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA
[b] Graduate Institute of Technology, University of Arkansas at Little Rock, AR 72204, USA
[c] MEG-Center, University of Tuebingen, 72076 Tuebingen, Germany

ARTICLE INFO

Article history:
Received 1 April 2008
Revised 15 July 2008
Accepted 31 July 2008
Available online 8 August 2008

Keywords:
Auditory evoked response
Fetal magnetoencephalogram
Cross-correlation
Bootstrap

ABSTRACT

Objective: We propose to use cross-correlation function to determine significant fetal and neonatal evoked responses (ERs).
Methods: We quantify ERs by cross-correlation between the stimulus time series and the recorded brain signals. The statistical significance of the correlation is calculated by surrogate analysis. For validation of our approach we investigated a model which mimics the generation of ERs. The model assumes a fixed latency of the ER and contains two parameters, ε and λ. Whether or not the system responds to a given stimulus is controlled by ε. The amount to which the system is excited from the base line (background activity) is governed by λ. We demonstrate the technique by applying it to auditory evoked responses from four fetuses (21 records) between 27 and 39 weeks of gestational age and four neonates (eight records).
Results: The method correctly identified the ER and the latency incorporated in the model. A combined analysis of fetuses and neonates data resulted in a significant negative correlation between age and latency.
Conclusions: The analysis of ER, especially for fetal and newborn recordings, should be based on advanced data analysis including the assessment of the significance of responses. The negative correlation between age and latency indicates the neurological maturation.
Significance: The proposed method can be used to objectively assess the ER in fetuses and neonates.

Introduction

The maturation of sensory cortex in newborns and fetuses can be investigated by stimulus evoked responses (ERs) recorded with electroencephalography and magnetoencephalography (e.g. Huotti-lainen 2005, Eswaran et al., 2005; Wakai et al., 1996). The major maturation index is the latency of the ER. Because the ERs are elicited by a single stimulus and are much lower in amplitude compared to the background activity, it is difficult to identify them in the cortical signals. For this purpose the stimulus is repeatedly delivered and the ER is quantified by averaging the signals locked to the stimulus delivery time point. In the averaged data, if the post-stimulus amplitude is "higher" than the pre-stimulus amplitude, it is defined as ER. This definition is valid under the following assumptions: (i) the background spontaneous activity averages out to the noise level of (close to) zero amplitude in the presence of sufficiently large number of stimuli and (ii) there are no interfering (biological/measurement) signals. These conditions can be difficult to achieve on fetal and neonatal studies due to issues such as patient comfort, physical state of the fetuses/neonates, and interferences such as breathing and cardiac signals all of which can limit the number of applied stimuli. Hence we need a measure to objectively quantify the ER. Here we propose to use cross-correlation function (CCF) to quantify the ER and subsequently use a bootstrap method to assess its statistical significance. To validate this method we propose a simple stochastic model that mimics ER. Finally, we apply this approach to auditory evoked responses measured from four fetuses (21 records) between 27 and 39 weeks of gestational age (GA) and four neonates (eight records one each from right and left stimulations) obtained by a 151-sensor magnetoencephalography (MEG) device.

Methods

Quantification of ER

Traditionally ER is quantified by stimulus-locked averaging. Also there has been effort to quantify the fetal ER using wavelet transform, a time and frequency domain approach (Norton et al., 2004).

For our proposed approach, let the time instances of stimuli delivery be κ. Let us define a stimulus series η such that it has a value of one at the time points κ and zero at all other points. For analysis we extract series W and Z. W is related to η and Z represents the cortical recordings. Both time series contain one second of data post the time point κ. The selection of one second is arbitrary, however this should be adequate for evoked responses in fetuses and newborns. If there were N stimuli, and sf was the sampling frequency (in Hz), then W and

* Corresponding author. Fax: +1 501 603 1544.
E-mail address: rbgovindan@uams.edu (R.B. Govindan).

1053-8119/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2008.07.054

Z will contain $N' = N \times sf$ data points. Now we define the cross-correlation between the series W and Z as follows:

$$C(\tau) = \frac{\sum_{i=1}^{N'-m\tau} [W(i)-\overline{W}] \cdot [Z(i+\tau)-\overline{Z}]}{\sqrt{\sum_{i=1}^{N'-m\tau} [W(i)-\overline{W}]^2} \sqrt{\sum_{i=1}^{N'-m\tau} [Z(i)-\overline{Z}]^2}}.$$

Since we are looking for the response elicited by the external stimulus, it is reasonable to shift the MEG data Z backwards in time to identify the ER. Hence we compute $C(\tau)$ only for $\tau = 0$ to sf. Shifting the stimulus signal W backwards in time and computing $C(\tau)$ would correspond to quantifying the impact of the MEG on the stimulus which does not have any physiological relevance. However, in the conventional stimulus-locked average, this scenario corresponds to the amplitude in the pre-stimulus duration which used to serve as a base line to gauge the signal intensity in the post-trigger duration. In this work we assess the significance of $C(\tau)$ by the bootstrap method and hence we did not study the correlation between the two signals in the pre-trigger regions. Further, cross-correlation is dependent on the length of the data. For each τ value $C(\tau)$ will be estimated for different number of data points and hence there will be an element of bias in this definition. In order to avoid this bias, for each shift (τ) we discard the same number of data points (i.e.) $N' - m\tau$, where $m\tau$ is the maximum value of τ which is one second in our example.

Assessing the statistical significance of CCF

We use bootstrap approach to assess the statistical significance of CCF. If one of the time series is uncorrelated, then assuming that the two time series have Gaussian distribution, we can use $\pm N^{-1/2}$ as the confidence limit for $C(\tau)$ (Theiler et al., 1992, Timmer et al., 1998), where N is the length of the data. However, in our case neither the trigger series nor the MEG is uncorrelated and hence we need to use a non-parametric approach to assess the statistical significance of CCF. For this purpose the surrogate signals are generated by randomly shuffling the disjoint one-second blocks of the MEG data (Govindan et al., 2006). Shuffling of the entire (trigger) time series has been considered in earlier works (Lv et al., 2007, McCubbin et al., 2008). The advantage of block shuffling over shuffling the entire time series is that the former preserves the (two point) correlation properties of the original signal whereas the later does not. This property is necessary to correctly address a certain type of null hypothesis that is related to biological signals as presented here. For example, if the neuronal data have certain type of correlations and produce spurious $C(\tau)$ (i.e. show a high correlation though there is no actual interaction between the two signals), then random shuffling of the entire signal (which destroys the auto-correlations in the signals) would falsify the null hypothesis testing while the surrogates obtained by block shuffling method would correctly address the null hypothesis. Further, we would like to mention that different bootstrap techniques to compute the confidence intervals of the evoked response are known (Kruglikov and Schiff, 2003, Fujioka et al., 2004). While the bootstrap approach proposed by Kruglikov and Schiff is designed for studies involving multiple triggers, the approach proposed by Fujioka et al., assumes a stationary background activity. Hence these methods cannot be applied to the fetal and neonatal data reliably.

To assess the statistical significance of CCF, we make a null hypothesis that the $C(\tau)$ computed between $\eta(t)$ and the MEG data is due to spurious correlations. To test the null hypothesis we synthesize the block shuffled MEG data as discussed above. We compute CCF of the surrogate data as $C(\tau)_{sur}$. Based on an earlier work (Schreiber and Schmitz, 2000), to reject the null hypothesis at α-level, we synthesize, $N = 2 \cdot K/\alpha - 1$ number of surrogates, where 2 in this equation denotes the two sided testing to include the positive and negative values (magnitudes) of $C(\tau)$ and K is a positive integer. In our calculations we used 50 realizations of surrogate which corresponds to $\alpha \cong 0.04$. The null hypothesis can be rejected if the distribution of $C(\tau)_{sur}$ computed from the surrogate realizations is different from $C(\tau)$ obtained from original time series (i.e.) if the values of $C(\tau)$ are well separated from the values of $C(\tau)_{sur}$ (from all the realizations). Conventionally $C(\tau)_{sur}$ (or any statistic computed for the original data and the surrogate data as well) is assumed to have Gaussian distribution and the confidence limits are defined by computing the number of standard deviations $C(\tau)$ that lie outside the distribution of $C(\tau)_{sur}$. However, in practice, the assumption of Gaussianity may not hold for the chosen measure and hence we might not be able to correctly address the null hypothesis. Instead, we use rank approach to address the null hypothesis (Theiler et al., 1992). We compute the K-th largest value and K-th smallest value of $C(\tau)_{sur}$ for each realization of surrogate. In our calculations we set $K = 1$ and hence we compute the maximum and minimum of $C(\tau)_{sur}$ for each realization of surrogate. Since we use 50 realizations of surrogates, we will have 50 different maxima and minima. To this end we use the maximum of all the maxima and minimum of all the minima as the upper and lower confidence level of $C(\tau)$. The values of τ at which $C(\tau)$ exceeds the upper/lower confidence level is considered as the statistically significant response. In this case, the null hypothesis that $C(\tau)$ is obtained due to spurious correlations can be rejected at the α level of confidence. Further, the number false positives or specificity of the approach can be improved by decreasing the α value. This can be accomplished by increasing the number of surrogates. However, in this work 50 realizations are chosen to keep the false positives and computational time minimal. Further, one can also increase K which again would result in computational effort and hence we set its value to one in our analysis.

Modeling aspects

One of the fundamental assumptions involved in quantifying ER (based on stimulus-locked averaging) is that the spontaneous brain activity is a zero mean random process. However, in practice, this limit will be reached only when the average is performed over a large number of stimuli. Hence to model the background brain activity $Y(t)$ we use zero mean Gaussian distributed random numbers. Further, we assume the sample frequency to be 312.5 Hz (see below) and generate $Y(t)$ for six minute duration. In an ER study, the externally applied stimuli will usually elicit a response (i.e.) modify the base level activity of the brain, at certain time point later from the onset of stimuli. This time delay is called latency and in our model it is denoted by δ. In an ER study, stimulus is delivered in a pseudo periodic fashion. In our model we use $\eta(t)$ to represent stimulus series. This series is populated with ones and zeros. The stimulus will be delivered only at time point when $\eta(t)$ has a value of one. In order to decide the time points (sample numbers) to deliver the stimuli, we generate a sequence κ. In this simulation the time interval between two successive stimuli is assumed to be 2 s (i.e.) 625 data points (sample frequency based on the real recordings 312.5 Hz). Thus, the sequence κ is generated using Poisson distributed random numbers with mean value of 625. We then compute the cumulative sum of this sequence which gives the continuous time representation of the stimulus points and use it as sequence κ. Thus $\eta(t)$ will have a value of one at time points dictated by the sequence κ and zero at all other time points. Thus, if t takes on a value in the sequence κ (i.e. where $\eta(t) = 1$), the evoked response elicited, $X(t+\delta)$, is modeled by modifying (adding a constant λ to) the value of the spontaneous brain activity $Y(t)$ at a latency δ (i.e.) $Y(t+\delta)$. λ can be thought of as the parameter which controls the signal to noise ratio of $X(t)$ and is varied between zero and one. For time t that is not in the sequence κ (i.e. where $\eta(t) = 0$), $X(t)$ will assume the values of background spontaneous activity $Y(t)$.

The developing brain of the fetus and neonate may not respond to all the externally delivered stimuli with an identical ER. In general it