Journal of Neuroscience Methods 223 (2014) 30–34

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Computational Neuroscience

Manual selection of spontaneous activity maps derived from independent component analysis: Criteria and inter-rater reliability study

Daniel R. Roquet[a,*], Bich-Tuy Pham[b], Jack R. Foucher[a,c]

[a] ICube, UMR 7357, University of Strasbourg, CNRS, Fédération de médecine translationnelle, Strasbourg, France
[b] INSERM U1114, Fédération de Médecine Translationnelle de Strasbourg, Strasbourg, France
[c] HUS, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

HIGHLIGHTS

- Criteria to reliably select spontaneous activity maps resulting of ICA of fMRI data.
- Excellent inter-rater agreement of manual selection of all spontaneous activity maps including idiosyncratic ones.
- This spontaneous activity maps selection allows to conduct reproducible experiments.

ARTICLE INFO

Article history:
Received 22 July 2013
Accepted 17 November 2013

Keywords:
FMRI
Independent component analysis
Default mode network
Spontaneous activity
Manual selection
Inter-rater reliability

ABSTRACT

Background: During the last years, many investigations focused on spontaneously active cerebral networks such as the default-mode network. A data-driven technique, the independent component analysis, allows segregating such spontaneous (co-)activity maps (SAM) from noise in functional magnetic resonance imaging (fMRI) time series. The inter-rater reliability of manual selection of not only the default-mode network but all SAMs remained to be assessed.
New method: The current study was performed on 20 min (400 volumes) fMRI time series of 30 healthy participants. SAMs' selection criteria were first established on past experience and from the literature. The inter-rater reliability of SAMs vs non-SAMs manual selection was then investigated from 250 independent components per participant.
Results: Inter-rater Kappa coefficient was of 0.89 ± 0.01 on whole analysis, and 0.88 ± 0.09 on participant per participant analysis.
Comparison with existing methods: Without focusing on specific and predetermined SAMs only, our criteria allow a reliable selection of all SAMs including the idiosyncratic networks.
Conclusions: The proposed SAM's selection criteria are reliable enough to allow scientific exploration of all SAMs at the single subject level.

1. Introduction

Studies using functional magnetic resonance imaging (fMRI) generally compare the blood oxygenation level-dependent (BOLD) signals between an experimental and a control task. This model-based approach as used by General Linear Model captures only part of the global brain functioning as (1) it needs strong a priori knowledge of the expected time course of task-related signal's fluctuations, (2) this massively univariate method needs drastic thresholds to remove false-positives, leading on the other side to an increase of false-negatives, and (3) it does not take into account the background of spontaneous activities. In contrast, as an exploratory approach (i.e., model-free or data-driven methods), spatial independent component analysis (sICA) is the most frequently used multivariate method. It considers that each voxel time course is a mixture from several contributions of independent sources such as artifacts, but also BOLD signals (McKeown et al., 1998). Components can then be separated into individual spatial maps as long as sources affect many voxels in a similar way. Such decomposition method highlights artifacts such as head motion, physiological events or machine artifacts (Kelly et al., 2010), but it can also detect cerebral networks since sICA natively assesses the functional connectivity between voxels. Beside artifact and paradigm dependant components, sICA gives some components that look like cognitive

---

* Corresponding author at: ICube – IPB, Faculté de Médecine, 4 rue Kirschleger, 67085 Strasbourg Cedex, France. Tel.: +33 03 68 85 40 71.
E-mail addresses: daniel.roquet@unistra.fr (D.R. Roquet), bich.thuy.pham@unistra.fr (B.-T. Pham), jack.foucher@laposte.net (J.R. Foucher).

0165-0270/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jneumeth.2013.11.014

networks unrelated to the task, which we will further referred to as spontaneous activity maps or SAMs. The most frequently observed SAM matches the default mode network (DMN) (Anticevic et al., 2012; Buckner et al., 2008; Greicius et al., 2004; Raichle et al., 2001). But other SAMs have been described (Damoiseaux et al., 2006) and many others remain to be studied in a normal population.

As physicians, our main interest in SAMs is that some of them can be altered by pathological processes, e.g., changes in the DMN-like SAM have been described in Alzheimer's disease (Greicius et al., 2004). Other SAMs might be a functional marker of specific symptoms, e.g., sensorial hallucinations in schizophrenia (Jardri et al., 2009). Despite that ICA studies are generally handled at a group level, there are strong clinical and therapeutic interests in better understanding the cerebral activity of a given patient.

Studying SAMs suppose that they can be reliably selected, i.e., distinguished from many other artifactual components in each subject. Several approaches has been developed either to remove artifactual components (Perlbarg et al., 2007; Sui et al., 2009; Tohka et al., 2008) or to automatically select SAMs, using spatial information (templates are used as reference for goodness-of-fit or regression analyses) (Calhoun et al., 2008; Greicius et al., 2004; Van de Ven et al., 2004), temporal information (Thomas et al., 2002) or both (De Martino et al., 2007; McKeown, 2000; Storti et al., 2013). However, since these methods need a priori information on spatial patterns and/or do not present in sensibility of 100 percent, they are not relevant neither at the subject level due to inter-individual differences (Franco et al., 2009) nor in exploratory approaches of pathological processes involving idiosyncratic networks. Last, methods involving a priori on task-related time-course are not appropriated for resting-state analysis. In contrast, our criteria are constructed to select manually all resting-state networks at an individual level, including the ones specific of a given subject.

The aim of this study was to ascertain the inter-rater reliability of SAMs vs non-SAMs manual selection at the subject level by raters trained according to expert defined criteria.

2. Materials and methods

2.1. Participants

After giving written informed consent, 38 right-handed healthy participants (age 37.3 ± 7.9 years; 20 males/10 females) with no history of neurological or psychiatric disorders underwent a resting-state fMRI session. This study was part of a protocol approved by the local Ethics Committee.

2.2. Data acquisition

Four hundred and five whole-brain T2*-weighted echo planar images were acquired interleaved on a 2 T Bruker scanner (Ettlingen, Germany) (session parameters: TR = 3 s; flip angle = 90°; TE = 43 ms; FOV = 256 mm × 256 mm × 128 mm; Imaging matrix = 64 × 64 × 32; 4 mm isotropic voxels, with fat saturation preparation) during 20 min (and 15 s). Participants were instructed to lie down with their eyes closed without falling asleep.

2.3. Preprocessing

After conversion to Analyze format, images were preprocessed using Statistical Parametric Mapping toolbox v99 (Welcome Department of Cognitive Neurology, London, UK) working on Matlab R2009b (The MathWorks, Inc., Sherborn, MA, USA).

For each participant, the first 5 images were removed to account for T1 partial saturation. The 400 remaining images were then motion corrected, and all the volumes were realigned on the 200th volume (*sinc* interpolation).

2.4. Statistical analysis of fMRI data

For each participant, sICA was performed using FMRLAB toolbox 2.3 (Swartz Center for Computational Neuroscience, University of San Diego, CA, USA) with an implementation of INFOMAX algorithm (Bell and Sejnowski, 1995). Since we planned to capture even small spontaneous activities for medical application, the original structure of the data should be preserved and so an excessive reduction of the dimensionality may be avoided (Abou-Elseoud et al., 2010; Green and Cordes, 2002; McKeown et al., 1998; Van de Ven et al., 2004). Moreover, Allen et al. (2012) reported that estimation quality of components do not decrease when the model order was greater than the true dimensionality. In this way, the dimension of the data was only reduced from 400 to 250 using a principal component analysis. This procedure implemented in FMRLAB allowed maintaining the computational time for the algorithm to converge in acceptable limits while conserving a maximum of variance. For display purpose, the components were superimposed on the EPI mean image at a threshold of ±1.5 standard deviation (SD).

2.5. Selection of SAM

The experimenter specialized in SAM selection (JF) elaborated a list of criteria with two raters (DR, BTP) on a set of independent data (Box 1). The raters were trained to select the SAMs on a subset of 8 datasets out of 38 (4 datasets on two training sessions). For each dataset, the two independent raters visually examined each component, and selected the SAMs. After each training session, raters compared their results and discussed with the expert to optimize the criteria. Once trained, the raters selected the SAMs of the remaining 30 datasets using three categories of scoring: SAM, non-SAM and doubtful-SAM.

2.6. Statistical analysis

Inter-rater agreement was evaluated by a Cohen's Kappa coefficient on both the global set (7500 items classified) and a participant per participant basis. This allows to evaluate the range of reliability achieved for more than 95% of the participants and to look for possible outliers. The Cohen's Kappa is a statistical tool for the assessment of the inter-rater agreement for qualitative items. It is more accurate than the percentage of agreement since it takes into account the agreement occurring by chance (Cohen, 1960).

3. Results

When considering doubtful-SAMs as non-SAMs (Table 1a), the two raters judged 402 and 456 components as SAMs. In all, the raters similarly marked 384 (14.3 components per participant ±5.2 SD) as SAMs and 7026 as non-SAMs (noise, artifacts or doubtful-SAMs). Thus, among 7500 components, 7410 were similarly classified, leading to 98.8% of agreement. Evaluation of inter-rater agreement lead to a global Cohen's Kappa coefficient of $\kappa = 0.89 \pm 0.01$.

It is noteworthy to mention that few components were considered as doubtful-SAM, due to a scattered spatial distribution generally coupled with a low level of activation, or with possible artifactual voxels. This group comprised between 18 and 91 components, depending on the rater, i.e., 11.13 ± 10% of all the SAMs. When considering doubtful-SAMs as SAMs components, the Cohen's Kappa coefficient was calculated on the basis of the selection results, as shown in Table 1b, leading to a global coefficient of $\kappa = 0.87 \pm 0.1$.