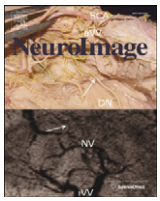
NeuroImage 60 (2012) 1538–1549

Contents lists available at SciVerse ScienceDirect

NeuroImage

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



Assessment of physiological noise modelling methods for functional imaging of the spinal cord

Yazhuo Kong *, Mark Jenkinson, Jesper Andersson, Irene Tracey, Jonathan C.W. Brooks

Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), University of Oxford, Oxford, UK

ARTICLE INFO

Article history:
Received 29 July 2011
Revised 26 October 2011
Accepted 25 November 2011
Available online 8 December 2011

Keywords:
fMRI
Physiological noise modelling
Spinal cord
Pain
Pre-whitening

ABSTRACT

The spinal cord is the main pathway for information between the central and the peripheral nervous systems. Non-invasive functional MRI offers the possibility of studying spinal cord function and central sensitisation processes. However, imaging neural activity in the spinal cord is more difficult than in the brain. A significant challenge when dealing with such data is the influence of physiological noise (primarily cardiac and respiratory), and currently there is no standard approach to account for these effects. We have previously studied the various sources of physiological noise for spinal cord fMRI at 1.5 T and proposed a physiological noise model (PNM) (Brooks et al., 2008). An alternative de-noising strategy, selective averaging filter (SAF), was proposed by Deckers et al. (2006). In this study we reviewed and implemented published physiological noise correction methods at higher field (3 T) and aimed to find the optimal models for gradient-echo-based BOLD acquisitions. Two general techniques were compared: physiological noise model (PNM) and selective averaging filter (SAF), along with regressors designed to account for specific signal compartments and physiological processes: cerebrospinal fluid (CSF), motion correction (MC) parameters, heart rate (HR), respiration volume per time (RVT), and the associated cardiac and respiratory response functions. Functional responses were recorded from the cervical spinal cord of 18 healthy subjects in response to noxious thermal and non-noxious punctate stimulation. The various combinations of models and regressors were compared in three ways: the model fit residuals, regression model F-tests and the number of activated voxels. The PNM was found to outperform SAF in all three tests. Furthermore, inclusion of the CSF regressor was crucial as it explained a significant amount of signal variance in the cord and increased the number of active cord voxels. Whilst HR, RVT and MC explained additional signal (noise) variance, they were also found (in particular HR and RVT) to have a negative impact on the parameter estimates (of interest) — as they may be correlated with task conditions e.g. noxious thermal stimuli. Convolution with previously published cardiac and respiratory impulse response functions was not found to be beneficial. The other novel aspect of current study is the investigation of the influence of pre-whitening together with PNM regressors on spinal fMRI data. Pre-whitening was found to reduce non-white noise, which was not accounted for by physiological noise correction, and decrease false positive detection rates.

Introduction

The spinal cord is the first relay site in the transmission of sensory information from the periphery to the brain (D'Mello and Dickenson, 2008; Willis and Coggeshall, 2003). To allow a more complete understanding of central nervous system processing in health and disease, a non-invasive technique for assessing the function of human spinal cord is desirable. To address this need, functional magnetic resonance imaging (fMRI) of the spinal cord has been developed. Since the publication of the first article on spinal cord fMRI in 1996 (Yoshizawa et al., 1996), there have been more than 50 papers on spinal cord fMRI performed in humans and animals. Spinal cord activation has been observed in cervical and lumbar spinal cord, for stimuli such as touch, vibration, thermal, brush and motor tasks (for review, Giove et al., 2004; Leitch et al., 2010; Stroman, 2005b).

Despite the increasing numbers of spinal cord fMRI studies in the literature, the characteristics of the functional signal is not fully understood (Bouwman et al., 2008; Cohen-Adad et al., 2010; Stroman, 2005a), and reported human data have been inconsistent. Possible reasons for this include: small cord area (~1 cm^2) producing intrinsically low signal to noise data; magnetic susceptibility differences (at bone/disc interfaces) causing signal loss and image distortion; physiological noise (primarily cardiac and respiratory related) obscuring functional signals; motion and flow artefacts due to cerebrospinal fluid (CSF) pulsation; variability of the results across repeated measurements and lack of a standard coordinate template for group analysis (Leitch et al., 2010; Stroman et al., 2008). One of the key challenges remaining to establish spinal cord fMRI as a routinely applied technique

* Corresponding author at: FMRIB Centre, Oxford University, John Radcliffe Hospital, Oxford, OX3 9DU, UK. Fax: +44 1865 222717.
E-mail address: ykong@fmrib.ox.ac.uk (Y. Kong).

1053-8119/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2011.11.077

is to adequately account for physiological noise effects when analysing data.

The blood oxygenation level dependent (BOLD) signal in fMRI is sensitive to a range of physiological variables. Physiological noise can induce changes in the BOLD signal and can therefore obscure the detection of neural activation using traditional experimental paradigms, and the relative contribution to measured signal could vary with increasing field strength, different acquisition parameters or multi-channel array coils (Bodurka et al., 2007; Hutton et al., 2011; Kruger and Glover, 2001; Triantafyllou et al., 2005, 2011). Physiological noise is a significant confound in connectivity studies, particularly estimation of the resting state (Birn et al., 2008a; Chang and Glover, 2009a,b; Jo et al., 2010; Lund, 2001; Shmueli et al., 2007). Several physiological noise sources have been studied in the brain fMRI literature, including the dominant low frequency fluctuations from the subject motion, cardiac and respiratory processes (Bhattacharyya and Lowe, 2004; Bianciardi et al., 2009; Birn et al., 2006; Chang et al., 2009; Gray et al., 2009a; Lund et al., 2006; Shmueli et al., 2007; Wise et al., 2004). Various methods of physiological noise correction for brain fMRI have been proposed e.g. k-space correction (Hu et al., 1995), independent component analysis (ICA) denoising (Kelly et al., 2010; Tohka et al., 2008) or image-based correction (Glover et al., 2000), known as RETROspective Image CORrection (RETROICOR). RETROICOR models cardiac and respiratory induced signal changes using a Fourier basis series defined by the relative phases in each cycle (i.e. cardiac and respiratory) at the time of image acquisition. More recently Deckers et al. (2006) proposed a selective averaging filter (SAF) method to account for the cardiac and respiratory effects by averaging imaging data based on their acquisition time relative to the cardiac and cycles. Heart rate (HR) (Chang et al., 2009; Shmueli et al., 2007) and respiratory related fluctuations (respiration volume per time — RVT, Birn et al., 2006) have been found to explain additional variance of the BOLD signal recorded in the brain.

Functional images of the brainstem and spinal cord are more severely affected by physiological noise (Brooks et al., 2008; Diedrichsen et al., 2010; Harvey et al., 2008) than the brain (Cohen-Adad et al., 2010). In particular, their proximity to the throat/trachea/lungs produces B0-induced susceptibility effects that can lead to signal change and image movement (Pattinson et al., 2009; Raj et al., 2001; Van de Moortele et al., 2002). The relative proximity of these structures (and relative sizes) to the major arteries lying near their surface and the CSF-filled spaces surrounding them (Anderson et al., 2009; Naidich et al., 2009), means that cardiac noise is particularly problematic (Dagli et al., 1999). The pulsatile nature of arterial blood and CSF flow around the cervical spinal cord, leads to large signal fluctuations near the boundaries of these structures (Piché et al., 2009). Indeed, by using hypercapnia as a positive control stimulus, Cohen-Adad et al. found that the highest responses at the spinal level were detected outside the cord and may reflect the contribution from larger veins or CSF pulsation effects (Cohen-Adad et al., 2010).

An improved RETROICOR method has been applied to brainstem fMRI, and gave a significant improvement in the detection of task-related activation (Harvey et al., 2008). The complex spatiotemporal structure of cardiac noise in the cervical spinal cord was recently examined, and ICA found to be useful when correcting for the cardiac effects (Piché et al., 2009). The use of principal component analysis (PCA) to decompose recorded cardiac parameters into their constituent signals has been proposed, and shown to account for their presence in spinal functional data (Figley and Stroman, 2009; Stroman, 2006). However this proposed PCA method has been optimised mainly for cardiac-related noise with spin-echo acquisition parameters, as in these images the respiratory effects are not as significant compared to data acquired with gradient-echo readouts. The SAF method was recently applied to a placebo analgesia spinal cord fMRI study (Eippert et al., 2009), but the performance of this correction scheme was not discussed. We have previously characterised various sources of physiological noise affecting spinal cord imaging and demonstrated the utility of a physiological noise model (PNM) for fMRI (Brooks et al., 2008). By including in the General Linear Model (GLM), basis functions derived from physiological measurements of cardiac and respiratory processes, and their interaction and a regressor defined by the measured CSF signal intensity time course, false-positive detection rates of pain-related activity recorded at 1.5 T were reduced and the expected location of activation was revealed.

The advantage of model-based techniques is that they are derived from actual physiological measurements from each subject, and the associated physiological noise can be automatically removed by including the model as nuisance regressors in the GLM. The noise structure in EPI time series could therefore be partially whitened by including these nuisance regressors into the GLM (Lund et al., 2006). Pre-whitening of fMRI data is advantageous for accurate statistical inference using the GLM, as temporal noise in the time series is assumed to be random (white); if it is not, then the statistical inference will be less accurate (Smith et al., 2007; Woolrich et al., 2001). A previous study compared the performance of the GLM when using RETROICOR regressors with and without an AR(1) model, and the AR(1) model was shown to be inferior to physiological noise modelling when used as a pre-whitening step (Lund et al., 2006).

In this study different model-based physiological noise correction techniques for spinal cord fMRI will be reviewed and implemented under the framework of the GLM. We evaluated the performance of different physiological noise models by comparing the residuals after model fitting, F-test regression model comparisons and the number of activated voxels under different types of stimulation. We aimed to find the optimal physiological noise model for spinal cord gradient-echo-based BOLD acquisitions, and explored whether pre-whitening performed using FMRIB's Improved Linear Model (FILM, Woolrich et al., 2001) is a necessary step to control false positive detection rates. Functional responses in the spinal cord of the rat have previously been studied using autoradiography with thermal pain stimuli (Coghill et al., 1991) and fMRI with electrical stimuli (Lilja et al., 2006). Thermal pain and sensory stimuli have been applied in human spinal cord fMRI studies, and segmental responses observed following stimulation of different dermatomes (Ghazni et al., 2010; Lawrence et al., 2008; Stroman, 2009). In the present study, thermal pain and punctate stimuli were used to investigate spinal cord responses. A summary of abbreviations used in this paper is given in Table 4.

Materials and methods

Experimental design and data processing

Eighteen healthy subjects (7 female) aged between 22 and 40 years, were imaged with a 3 T Siemens Trio MR system (Siemens Medical Systems, Erlangen, Germany). Subjects were placed in the supine position, and images acquired with the standard 12-channel head coil, the 4-channel anterior-posterior neck array and upper

Table 1
Voxelwise F-test results showing the percentage of cord voxels in which variance was significantly reduced ($p<0.01$), by adding specific regressor(s), *above and beyond* the base model. Rightmost column gives the mean p-value for all voxels that passed the voxelwise significance threshold.

F-test (regressor(s): base model)	$p<0.01$	Mean p
SAF: Basic design	96%	0.0001
PNM: Basic design	99%	0.00004
CSF: Basic + PNM	64%	0.0005
MC: basic + PNM + CSF	64%	0.0006
HR: Basic + PNM + CSF	29%	0.0014
HRcrf: Basic + PNM + CSF	12%	0.0026
RVT: Basic + PNM + CSF	30%	0.0016
RVTrrf: Basic + PNM + CSF	14%	0.0023

Table 2
The number of activated voxels (relative to the number measured with the *Basic + PNM + CSF* model) of each experimental contrast passing an uncorrected statistical threshold $p<0.01$ for the different physiological noise models. RP for right punctate, RT for right thermal, LP for left punctate and LT for left thermal.

Model	LP	RP	LT	RT
Basic + PNM + CSF + MC	−79	23	10	20
Basic + PNM + CSF + HRcrf + MC	−70	−6	6	27
Basic + PNM + CSF + HRcrf	13	−20	−21	2
Basic + PNM + CSF + HRcrf + RVTrrf + MC	−61	−8	2	15
Basic + PNM + CSF	*0*	*0*	*0*	*0*
Basic + PNM + CSF + RVTrrf	0	15	−24	−1
Basic + PNM + CSF + HRcrf + RVTrrf	23	−19	−23	−1
Basic + PNM + CSF + HR	−5	42	−43	−73
Basic + PNM + CSF + RVT	−30	10	−37	−37
Basic + PNM + CSF + HR + RVT	−25	14	−53	−91
Basic + PNM + CSF + HR + MC	−83	38	−14	−40
Basic + PNM + CSF + HR + RVT + MC	−92	18	−43	−63
Basic + PNM	−131	−53	−67	−116
Basic + SAF	−136	−62	−81	−50
Basic design	−153	−64	−133	−56

element of the spine array. To monitor cardiac and respiratory processes subjects wore a pulse oximeter and respiratory bellows. The volume trigger from the scanner host computer was also recorded, and all data logged on Chart (ADInstruments, USA) at a sampling rate of 100 Hz. The experimental paradigm consisted of three second long moderately painful (average pain rating 3.6 out of 10) thermal stimuli and one second long non-painful punctate stimuli, which were applied to the medial aspect of the lower forearm (inner wrist) and base of the thumb, respectively. Stimuli were applied to the right and left arms, using pseudo-random ordering, and alternated between thermal and punctate modalities. We use the notation RP for right punctate, RT for right thermal, LP for left punctate and LT for left thermal throughout the rest of this paper. GRE-EPI was used to acquire 9 axial slices covering the expected locations for spinal activity (C4 to C8) with the following parameters: TE/TR = 39/1000 ms; flip angle = 68°; GRAPPA (factor = 2) parallel acceleration; phase encoding direction: posterior to anterior; image matrix 96 × 96. The field of view was 128 mm, which gave an in-plane resolution of 1.33 × 1.33 mm; slice thickness was 4 mm with variable gap (depending on each subjects' anatomy). In total, 80 stimuli were delivered in the course of the acquisition of the 2564 volumes (~40 min). During the scan, subjects were requested to remain as still as possible.

Data were pre-processed using the FSL software (FMRIB Software Library, www.fmrib.ox.ac.uk/fsl). Each slice was motion corrected in 2D (translation only) using FLIRT (part of FSL, Jenkinson et al., 2002). Subsequently, data were spatially smoothed (3 mm FWHM), high-pass temporal filtered (300 s cut-off), and activity within each slice assessed independently with slice-specific physiological noise regressors.

Table 4
Summary of abbreviations used in the paper.

Abbreviations		Abbreviations	
SAF	Selective averaging filter	RP	Right punctate
PNM	Physiological noise model	LP	Left punctate
CSF	Cerebrospinal fluid	RT	Right thermal
MC	Motion correction	LT	Left thermal
HR	Heart rate	RSS	Residual sum of squares
crf	Canonical HR response function	NRV	Normalised residual variance
HRcrf	HR convolved with crf	Ri	Relative index of parameter estimates
RVT	Respiration volume per time		
rrf	Respiration response function	B	Basic design only
RVTrrf	RVT convolved with rrf	BF	Basic design with FILM
GLM	General linear model	BP	Basic design with PNM + CSF
		BPF	Basic design + PNM + CSF with FILM
		G	Goodness metric of false-positive rate

Physiological noise modelling methods

Here we will give a brief introduction to the different physiological noise modelling methods implemented in this study. Physiological noise correction techniques rely critically on the timing of the image acquisition therefore should ideally be performed on a slice-by-slice basis (Birn et al., 2006; Jones et al., 2008). All the models used in this study utilise the acquisition timing for each slice, which was obtained by dividing the TR by the number of slices, time-locked to acquisition of each volume.

Physiological noise model (PNM)

RETROICOR is a retrospective image-based correction technique introduced by Glover et al. (2000). The cardiac and respiratory phases of each fMRI volume are derived from their acquisition time relative to a separate physiological recording, and modelled via a Fourier basis series with a combination of sine and cosine harmonics. The time series are regressed to the image data and can then be extracted from the fMRI signal to reduce the noise (Corfield et al., 1999). In our previous study (Brooks et al., 2008), various sources of physiological noise were explored and, in addition to independent effects of cardiac and respiratory noise, the interaction between these two processes

Table 3
Summary of physiological noise models.

Model		Description	Suggestion
Main model	SAF	Selective average filter — Divide each cardiac/respiratory cycle into a given number of bins	**Not recommended** Less effective model of physiological noise in the spinal cord, with potentially large loss of degrees of freedom (e.g. 140 regressors)
	PNM	Physiological noise model — Fourier series to model the cardiac, respiratory and interaction phases	**Recommended** Most effective physiological noise correction method tested
Additional model	CSF	Cerebrospinal Fluid — Averaged time course from voxels around the spinal cord showing the most variance	**Recommended** Significantly reduced residual variance in this dataset and increased statistics
	MC	Motion correction parameters — X and Y translations	**Recommended/careful** Produced significant reduction of variance in this dataset but can also affect statistics if correlated with stimuli
	HR	Heart rate — Smoothed heart beat rate (beats per minute)	**Careful** PRODUCED marginal reduction of variance in this dataset but can also affect statistics if correlated with stimuli
	RVT	Respiration volume per time — The difference between the maximum and minimum of each respiratory cycle divided by the period of the cycle	**Careful** Produced marginal reduction of variance in this dataset but can also affect statistics if correlated with stimuli

was demonstrated to be a significant source of noise. Therefore the PNM, a modification of conventional RETROICOR, was presented.

Smoothing and peak detection was conducted on the recorded pulse oximeter data to assign a cardiac phase (θ_C) to each acquired image. While respiratory phase (θ_R) was assigned using the histogram equalised transfer function proposed by Glover et al. (2000), which takes into account both the relative position in the respiratory cycle and depth of breathing. A summary illustration of cardiac and respiratory phase assignment is shown in Fig. 1(a). Cardiac and respiratory effects were modelled using the separate sine and cosine terms of the principal frequency (θ_C and θ_R) and harmonics. On the basis of empirical data, we have utilised regressors up to the third harmonic frequencies for both cardiac and respiratory terms (i.e. θ_C to $4\theta_C$, θ_R to $4\theta_R$) — giving a total of 16 regressors. An additional 16 multiplicative sine/cosine terms, which take into account the interaction between cardiac and respiratory effects in the physiological noise, were given by sin or cos ($A\theta_C \pm B\theta_R$), where A = 1, 2 and B = 1, 2. In total 32 nuisance regressors are included in the basic PNM.

Selective averaging filter (SAF)

The SAF models physiological signals by averaging imaging data based on their acquisition time relative to the nearest respiratory and cardiac event (e.g. end inspiration or systolic peak). Similar to the PNM, which uses a Fourier basis set, the SAF uses a basis set expansion of the phase values — in this case a finite impulse response (FIR) basis set. In the original implementation by Deckers et al. (2006), the period of each cycle (cardiac or respiratory) was evenly divided into a number of bins (equivalent to the FIR basis functions), and every fMRI *volume* assigned a unit entry in one of the cardiac and one of the respiratory bins. The resulting physiological noise estimates are subtracted from the fMRI time series. A summary illustration of this process is shown in Fig. 1(b). In this study we applied the SAF as a confound design matrix within the GLM. The period of each cycle (cardiac and respiratory) was evenly divided into N bins (e.g. N = 1,2,...20), and each bin modelled with a separate regressor. In our implementation the relative timing of *each slice* determines which bin it is allocated to, and thus they are assigned a unit entry in the corresponding (and separate) cardiac and respiratory regressors. An important parameter of this method is the number of bins. By using our FMRI data, we estimated the optimal number of bins for respiratory and cardiac cycles independently by calculating the reduction in variance obtained when systematically increasing the number of bins used for regressor generation.

Cerebrospinal fluid (CSF)

The spinal cord is surrounded by cerebrospinal fluid (CSF), which is pulsatile (Willis and Coggeshall, 2003). The effect of CSF pulsation is twofold: it induces cord movement and generates cardiac driven signal change across the imaging slice. The latter is driven by the relative flow sensitivity of EPI pulse sequences. In this study, the variance map for the raw signal time course (after motion correction) was computed within a region including both the spinal cord and CSF space, and thresholded such that only voxels whose variance lay in the top 20th percentile were included in the resultant CSF mask. The CSF mask was then applied to the motion corrected time course data to extract the mean time course for voxels exhibiting the largest variance. The obtained time course was orthogonalised against the experimental design matrix and included in the GLM as the CSF regressor. Voxels within the CSF mask were mainly located in the subarachnoid space, i.e. in the CSF adjacent to the cord. Fig. 2(a) shows the CSF masks, CSF time series and power spectrum from one subject as an example.

Heart rate (HR)

Heart rate is regulated by processes at the level of the central nervous system and alters blood volume and blood flow through hemodynamic mechanisms (de Munck et al., 2008). In a recent study it was reported that subjects' heart rate explained an additional 1% of BOLD signal variance beyond respiratory (RVT) and eight RETROICOR regressors (Shmueli et al., 2007). In agreement with Shmueli et al., a canonical HR response function (crf) was proposed (Chang et al., 2009) and found to significantly reduce the variance of resting BOLD time series when applied together with a respiratory variation function in a linear model. In this study, HR was calculated by translating the beat-to-beat interval (in seconds) from the pulse oximeter trace into beats per minute (i.e. 60/interval BPM). The BPM trace was smoothed using a moving average filter (kernel size = 10 adjacent points in the BPM trace). The HR regressor is simply the value of the smoothed BPM trace at the acquisition time for each slice. In addition the HR regressor was convolved with the crf (adapted from Chang et al., 2009) to give a new time course of HRcrf — an alternative

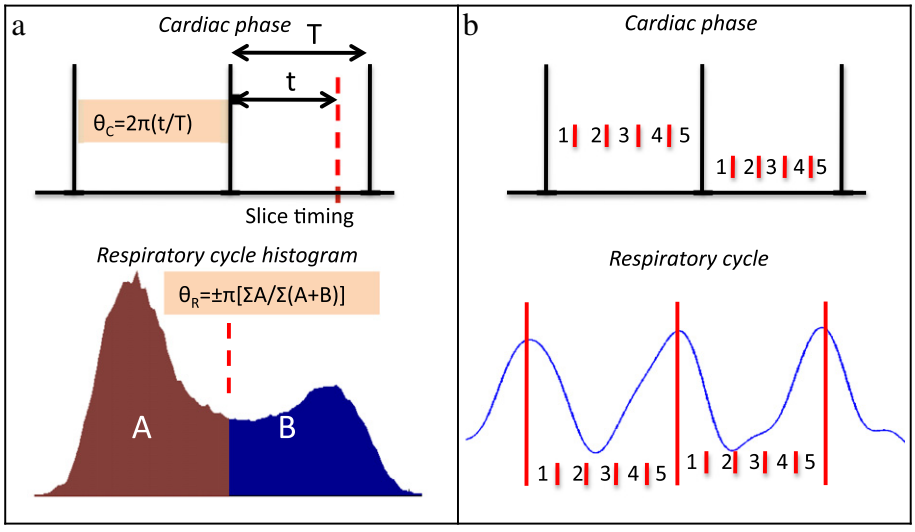


Fig. 1. Schematic illustrations of the two physiological noise modelling methods. Information about cardiac cycle, respiratory cycle and fMRI slice timing is obtained from physiological monitoring data. a) The PNM models cardiac and respiratory noise using a Fourier-series with a combination of sine and cosine harmonics and interaction terms. Cardiac and respiratory phases are determined as shown in the figure. b) The SAF models physiological signals by averaging imaging data based on their acquisition time relative to the nearest respiratory and cardiac event. The period of cardiac or respiratory cycle is evenly divided into bins. FMRI volume assigned to the same bins will be averaged.

regressor to further investigate the heart rate influence in the spinal cord. Fig. 2(b) shows the instantaneous HR plot (blue), the smoothed HR (red), the HRcrf (green) and the corresponding power spectra from one subject as an example.

For pain-related studies, clinicians often use HR as a complementary objective measure of pain. HR should increase as pain increases, which is mostly likely related to an increase in sympathetic activity. A significant correlation between HR and pain intensity has previously been observed in healthy subjects (Tousignant-Laflamme et al., 2005). Therefore, HR might have an impact on parameter estimation of the thermal painful stimuli applied in this study if they are correlated.

Respiration volume per time (RVT)

Previously observed in the brain, low frequency fluctuations in end tidal blood CO2 concentration (EtCO2) are significantly correlated with the BOLD signal and are an important source of signal variance (Wise et al., 2004). These low frequency CO2 fluctuations are hypothesised to be due to changes in the rate and depth of breathing, which are not corrected by the basic Fourier series terms based on respiratory phase. The respiration volume per unit time (RVT) was proposed to model the variations in breathing depth and rate (Birn et al., 2006). Furthermore, it was shown that convolving the RVT time series with the derived respiration response function (rrf) has a significant correlation to resting state brain fMRI time series across widespread regions of grey matter, potentially obscuring "real" task activation and altering correlations between brain regions (Birn et al., 2008b; Chang and Glover, 2009b).

In this study, RVT was estimated according to Birn et al. (2006) as the difference between the maximum and minimum of each respiratory cycle (i.e. the amplitude difference between the peaks/troughs of the respiration waveform) divided by the period of the cycle. The RVT values were interpolated to obtain the fMRI RVT regressor. In addition the RVT regressor was convolved with the rrf (adapted from Birn et al., 2008b) to give a new time course: RVTrrf. Fig. 2(c) shows the instantaneous respiratory cycle plot (blue), the RVT (red), the RVTrrf (green) and the associated power spectra from one subject as an example.

Motion correction parameters (MC)

Even after motion correction, various effects like interpolation errors, changes in local magnetic susceptibility due to movement of the body, spin history effects, etc. can still induce motion-correlated intensity changes in BOLD data (Johnstone et al., 2006). Motion parameters are frequently included in the design matrix to remove motion-correlated signal, which can reduce the number of type I errors (false positives) at the expense of increasing type II errors (false negatives) (Friston et al., 1996; Johnstone et al., 2006). However, if motion is correlated with task then this will reduce statistical significance. In this study, motion correction parameters were included in the model test. Each slice was motion corrected in 2D (translation only) and the x and y translation parameters were used as motion correction regressors (Fig. 2d).

Model evaluation

Different models (*SAF*, *PNM*) and combinations of regressors (*CSF*, *HR*, *HRcrf*, *RVT*, *RVTrrf*, and *MC*) were compared to the basic paradigm design matrix (referred to as *Basic* below) used in a GLM model in MATLAB (R2010a, Natick, Massachusetts; The Mathworks Inc). Each model set was first evaluated by computing the residual sum of squares (RSS) after model fitting. To provide an indicator of the relative improvement (i.e. variance reduction) obtained with each model, the RSS was divided by the number of degrees of freedom (dof) and then normalised by the RSS/dof for the model with the basic design only (i.e. just the convolved stimulus time courses), in order to give a measure of normalised residual variance (abbreviated to *NRV*). A mask containing only the spinal cord area was manually drawn in FSLView using the mean functional image from each run. All analysis was based on the 6455 voxels lying within the spinal cord masks combined across all 18 subjects. Next, the significance of sets of

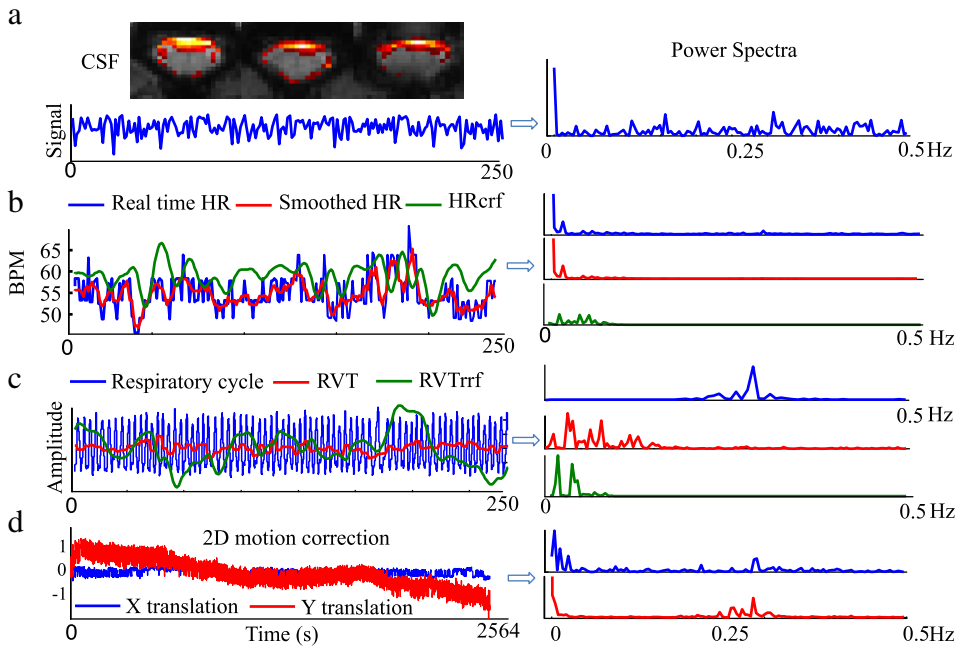


Fig. 2. Additional regressors used to model low frequency signals (examples taken from a single subject). Left: time courses of measured and derived parameters, and right: associated power spectra. a) Top 20th percentile variance map used for CSF time series extraction. Three CSF masks and one time course are shown as example. b) Real time heart rate (blue), smoothed heart rate time series used as HR regressor (red) and HR convolved with cardiac response function ("HRcrf", green). c) Respiratory waveform (blue), computed RVT time series (red) and RVT convolved with respiratory response function ("RVTrrf", green). d) 2D motion correction regressors.

regressors was tested using the F-test regression model comparison, which identified voxels where the test model explained a significant amount of noise above the base model. Then, the model set was evaluated by determining the number of voxels marked "active" by each model.

The correlations between all regressors were examined to assess the impact of non-zero correlation between nuisance regressors and the experimental regressors (i.e. RT, LT, RP, LP). The magnitude of these correlations, the relative sizes of the parameter estimates (β) and their combined effect on the estimated statistics was explored. The relative index (Ri) was defined as the ratio of parameter estimates (β for nuisance regressor divided by β for experimental regressor) multiplied by the correlation coefficient (cc), as $Ri = (\beta_{regressor}/\beta_{Experimental}) \cdot cc_{regressor,Experimental}$. β parameters reflect the amount of variance accounted for by each regressor in the GLM. The relative index (Ri) reflects the relative importance of the tested nuisance regressor, modulated by the correlation between the tested nuisance regressor and the experimental regressor. The relationship between Ri and experimental T statistics after including separate nuisance regressors (HR, RVT, x or y) was also investigated.

Pre-whitening tests

Finally, pre-whitening tests were performed using the temporal autocorrelation estimation and pre-whitening tool, FILM, in FSL. It is known that fMRI time series typically show temporal autocorrelation, even after a high-pass filter has been applied to remove slow varying trends. Autocorrelation estimation by FILM is performed locally at each voxel on the residuals after the initial GLM fit. Importantly, estimation is further improved by using nonlinear spatial smoothing (SUSAN, part of FSL) within tissue type. A filter is constructed in order to pre-whiten each voxel's time series, by including the pre-whitening matrix in the GLM. This step gives improved estimation efficiency compared with methods that do not pre-whiten (Woolrich et al., 2001). Autocorrelation in fMRI time series may be due to various factors, only one of which is physiological noise. The effectiveness of reducing residual variance in fMRI time series by using physiological noise models alone is hampered by the fact that the noise spectrum cannot generally be characterised by a few discrete frequencies (Toga and Mazziotta, 2002). We investigated whether pre-whitening could reduce non-white noise, which, if not already compensated for by the applied physiological noise model, might additionally decrease false positive detection rates (through improved modelling).

We only choose the basic experimental design and PNM + CSF for testing, as this was the optimal model in a previous report (Brooks et al., 2008). Four cases were examined: basic design only (B), basic design with FILM (BF), basic design with PNM + CSF (BP) and basic design + PNM + CSF with FILM (BPF). To assess the impact of pre-whitening on FMRI time series, normalised residual variance (NRV, normalised against basic design only) for each applied model was inspected. Subsequently, spatial patterns of activated voxels in the spinal cord and CSF were studied to determine if pre-whitening (in addition to physiological noise modelling) altered false-positive detection rates. To achieve this we defined a new metric (G, Goodness) which is the ratio of the normalised voxel counts (by total voxel number) in the spinal cord and CSF. A CSF mask was manually drawn around the previously defined spinal cord mask for each subject, and a one-voxel gap left between cord and CSF masks to avoid partial volume voxels on the edge of the spinal cord (which might include draining veins). Active voxels in the cord or CSF were normalised to the total voxel count in each specific area. Therefore, the model that reduced putative false positives in the CSF space, but increased the number of active cord voxels (i.e. greater G value) was judged to be superior. To facilitate comparison, all G values were normalised to the G values obtained using the basic design without pre-whitening.

Results

Determining the optimal number of bins for the SAF model

Plots of NRV as a function of the number of bins (1 to 70) in the SAF model for cardiac and respiratory separately are shown in Fig. 3. The mean and standard error of each bin number for all subjects over all 6455 spinal cord voxels are shown. It can be seen from the figure, the largest reduction in residuals, i.e. the lowest NRV, was found for the model with the most bins (70), which was used for further analyses.

Model evaluation

A plot of NRV (mean and standard error across all voxels) for different models is shown in Fig. 4. NRV was lower for the PNM (with 32 regressors) than for the SAF (with 140 regressors/bins). Therefore the PNM was used as the basis for our physiological noise model, and the other low frequency confound regressors — CSF, HR, HRcrf, RVT, RVTrrf and MC, included in further tests. It can be seen from Fig. 4 that the CSF and MC regressors clearly explain considerable variance in the cord as NRV dropped substantially with their inclusion. NRV was also reduced by including HR and RVT regressors, but not by the same amount as for the CSF and MC regressors. Note that the residual variance was not reduced by including either respiratory (RVTrrf) or cardiac (HRcrf) response function.

Voxelwise F-tests were performed by comparing a base model and the base model *plus* specific regressor(s), and results were pooled over all spinal cord voxels. Table 1 shows the percentage of all cord voxels in which a given test model accounted for significantly more variance than the base model (threshold $p<0.01$), and the mean p-value of the thresholded voxels. SAF and PNM were tested separately against the basic experimental design first. Including either SAF or PNM reduced the signal variance significantly. Almost all voxels (99%) showed that PNM gave a significantly (mean p-value of 0.00004) better fit to the data than the basic design. PNM also outperformed SAF in this model F-test comparing the number of thresholded voxels and the mean p-value. Including CSF was significant for over 64% of all voxels. On top of PNM and CSF, 64% of all voxels also significantly reduced in variance when including MC. By comparison, HR and RVT regressors were significantly beneficial for only about 30% of all voxels. It should be noted that the HRcrf and RVTrrf regressors were not as effective as the original HR and RVT across the various F-tests performed.

Finally, we assessed the amount of activity from statistical parametric maps obtained using different testing models. The number of voxels

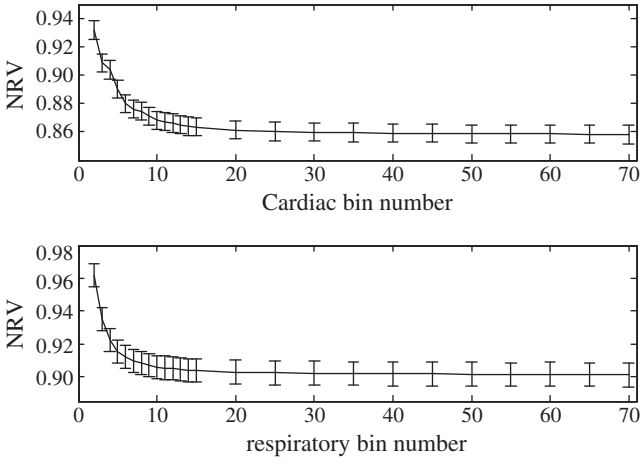


Fig. 3. Plot of normalised residual variance (NRV) as a function of the number of bins in the SAF model. As expected, the largest reduction in residuals, i.e. the lowest NRV, was found for the model with highest number of bins. The NRV would not be expected to increase because of the huge number (2564) of time points in the data.

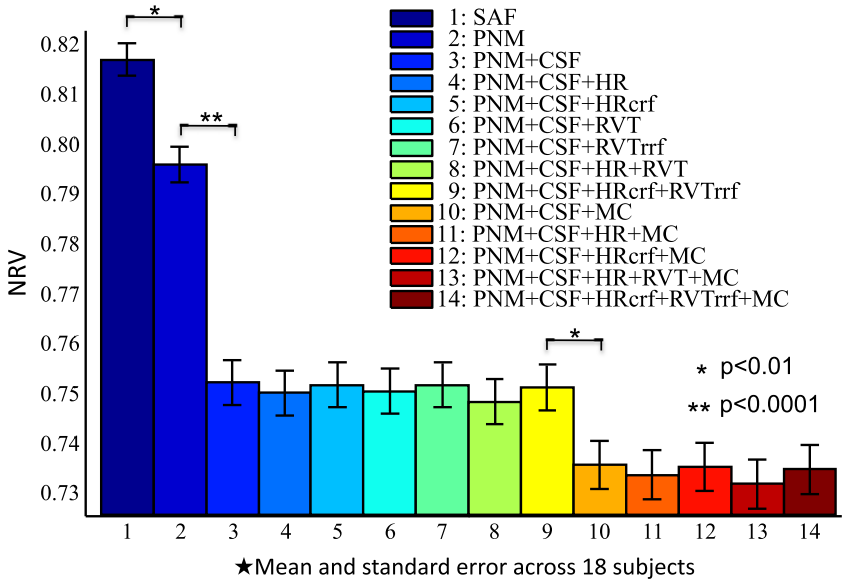


Fig. 4. Quantifying the influence of different physiological noise models on residual variance in spinal cord FMRI data. Note that NRV was clearly lower when modelling physiological noise with a PNM compared to modelling it with SAF. The CSF signal clearly explains variance *in the cord* as NRV dropped substantially with inclusion of this regressor. NRV was also lower after adding HR, RVT and MC regressors. Note that residual variance within the spinal cord did not show any reduction when including either respiratory (RVTrrf) or cardiac (HRcrf) response function.

passing a statistical threshold (uncorrected $p<0.01$) for each of the four stimuli was compared between the models. Basic + PNM + CSF, previously shown to be a highly robust model (Brooks et al., 2008), was treated as a base model to show the relative changes in the number of activated voxels for various different models, as detailed in Table 2. The models listed in Table 2 are ordered according to the number of experimental contrasts (i.e. different stimuli or combinations of stimuli) that showed increases in the number of active voxels. For the purposes of this ordering, each experimental contrast scored +1 if the number of active voxels increased and scored −1 otherwise. Compared to the basic model alone, more voxels were identified as active with both PNM and SAF. Inclusion of CSF and MC regressors (except for LP) produced a noticeable increase in both the number and significance of activated voxels. In general, a decrease in the number of voxels passing statistical threshold was observed when the model included HR or RVT regressors. Including HRcrf and RVTrrf was rarely beneficial.

Correlation between nuisance regressors and the basic design

The correlations between all slice basis regressors (except SAF) were examined first, including the experimental design (8 EVs and corresponding temporal derivatives), PNM (32 terms), HR, CSF and MC (x and y translations). HR correlated modestly with RVT ($R=-0.05$ to 0.38). Interestingly CSF mainly correlated to the first harmonics of the cardiac term from the PNM regressors (sine term $R=-0.49$ to 0.84; cosine term $R=-0.77$ to 0.55). This indicates that cardiac function has a significant influence on CSF signal recorded from around the spinal cord and agrees with previous studies (Dagli et al., 1999; Friese et al., 2004; Piché et al., 2009). Furthermore, RVT correlated with the second harmonics of the respiratory *cosine* term of the PNM ($R=0$ to 0.36).

The y translation motion parameter correlated most strongly with the first harmonics of respiratory cosine term of PNM ($R=-0.16$ to 0.85), which indicates that respiratory changes have a major impact on estimates of spinal cord motion, primarily in the dorsal-ventral direction (i.e. in the phase encoding direction). Respiratory processes produce changes in the main magnetic fields (B_0) that lead to phase offsets across the imaging slice that are physiological in origin. These respiratory influences produce apparent movement in the phase-encode direction in spinal cord imaging data. Indeed, this effect can be visualised from inspection of the power spectrum of y-translations (Fig. 2d), which reveals a respiratory influence on this parameter (y) at ~0.3 Hz (the typical respiratory rate).

Notably both x ($R=-0.85$ to 0.93) and y ($R=-0.91$ to 0.94) translation parameters highly correlated with the CSF regressor. This is expected as rapidly moving CSF is a major source of motion artefacts. HR ($R=-0.25$ to 0.25), RVT ($R=-0.14$ to 0.38), x ($R=-0.11$ to 0.07) and y ($R=-0.07$ to 0.1) translation parameters were found to be slightly correlated with the experimental design terms. The impact of this non-zero correlation between nuisance and experimental regressors was explored further due to the potentially large effects it can have on the final statistics. Fig. 5 plots the changes in T-statistics for the four experimental regressors (RT, LT, RP and LP) against the relative index Ri after adding *one* specific regressor (either HR, RVT, x or y) to the *Basic + PNM + CSF* model used to account for physiological noise: the change in T-statistic is defined as the T-statistic of the base model plus extra regressor minus the T-statistic of the base model (base model = Basic + PNM + CSF). A negative linear trend was found for the correlated nuisance regressors (HR, RVT and motion parameters), demonstrating a reduction in T-statistics due to shared variance between regressors. Where the β value of the nuisance regressor is large relative to the experimental regressor, even a small positive correlation between the two translated to a large reduction in T-statistics. However, negative values of Ri were associated with increased T-statistics. Without ground truth it is impossible to determine whether this effect was due to an increase of either true or false positives. It is interesting to note that this effect appears to be stronger (steeper slope) during painful thermal stimuli, particularly for HR and RVT. Among these four tested regressors, motion parameters appear to have a smaller impact on model T-statistics.

Pre-whitening tests

A plot of NRV (mean and standard error across all voxels) for different models is shown in Fig. 6. Including the basic design (B) with FILM prewhitening (BF) reduced the variance of the residuals in our spinal cord data, as indicated by the reduction in NRV. However, a larger reduction was found when the model included just the PNM plus CSF terms (BP). The greatest reduction in variance was obtained when using prewhitening together with PNM plus CSF regressor (BPF). Note that the

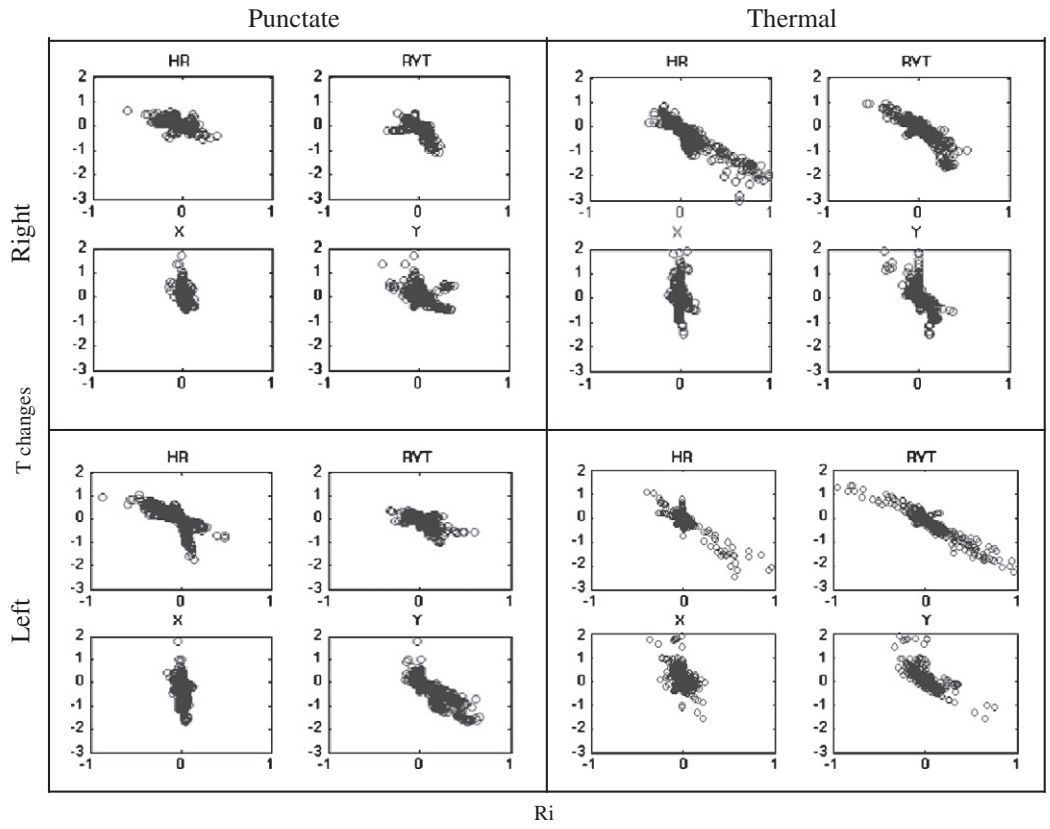


Fig. 5. Impact for T statistics. This figure plots the changes in T-statistics for the four experimental regressors (RT, LT, RP and LP) against the relative index Ri after adding one specific regressor (either HR, RVT, x or y) to the Basic + PNM + CSF model. Negative linear trend was found between T statistics changes and Ri. Note that 1) all regressors have more impact for painful thermal stimuli than non-painful punctate stimuli. 2) Negative linear trend is more apparent and consistent for HR and RVT.

variance obtained when modelling with the basic model alone (B) is shown for reference.

Fig. 7 shows the model comparison results for all four stimuli from first level analysis of 18 subjects. The upper panels in the figure show the normalised activated voxel numbers (mean ± standard error, $p<0.01$ uncorrected) in cord (solid) and CSF (dashed) areas for different models and stimuli (as labelled underneath). The PNM increased the number of activated voxels both in the cord and CSF areas mainly because the PNM models significant amounts of physiological noise.

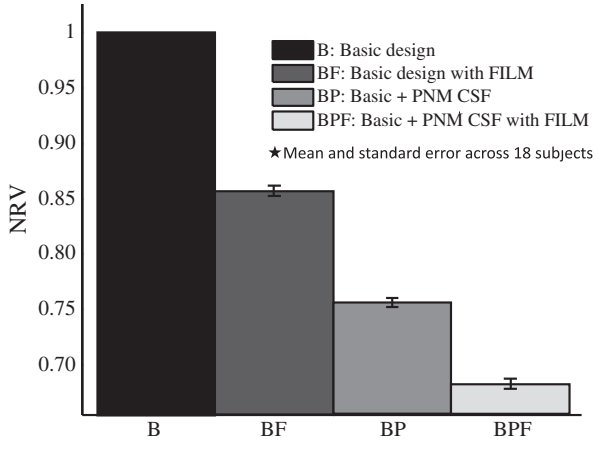


Fig. 6. A plot of NRV for models with or without pre-whitening. Note that the variance obtained when modelling with only the basic model alone (B) is shown for reference. Including the basic design (B) with FILM (BF) pre-whitening reduced variance of the residuals. A larger reduction was found when the model included just the PNM plus CSF terms (BP). The greatest reduction in variance was obtained when using pre-whitening together with PNM plus CSF regressor (BPF).

However, the number of activated voxels, both in the cord and CSF areas, drops when using pre-whitening — which may relate to a reduction in false positive detection rates. Lower panels in the figure show the normalised G values (mean ± standard error) for all models. G increased for the RP and LP stimuli when using FILM (pre-whitening) with the basic design, but slightly dropped when used with the RT and LT stimuli. Adding on PNM and CSF without pre-whitening helped to increase G for all four stimuli, but the greatest G values were achieved for each stimulus when applying PNM and CSF together with pre-whitening. Therefore, pre-whitening appeared to produce a greater reduction in false positive detection rates in the CSF than in the cord.

Discussion

Human spinal cord fMRI is feasible, but care should be taken when removing the physiological noise from the data. Most previous studies did not adequately account for physiological noise, and instead have used low-pass or band-pass filters to remove high frequency noise (Agosta et al., 2008; Maieron et al., 2007; Summers et al., 2010). Such filtering typically removes a large proportion of the variance associated with the actual data measurement (Friston et al., 2000). To adjust residuals for the associated loss of degrees of freedom and avoid artificial inflation of t-scores, a matrix describing the applied pre-filtering, or a noise correlation matrix, should be incorporated within the GLM. An alternative approach is to implement a physiological noise model (based on measurements taken from each subject) within the framework of GLM, which corrects for the physiological noise and implicitly accounts for loss of degrees of freedom. In this study we have demonstrated the impact of physiological noise on our ability to detect activity recorded using fMRI in the human cervical spinal cord. To achieve this we evaluated several different approaches to modelling

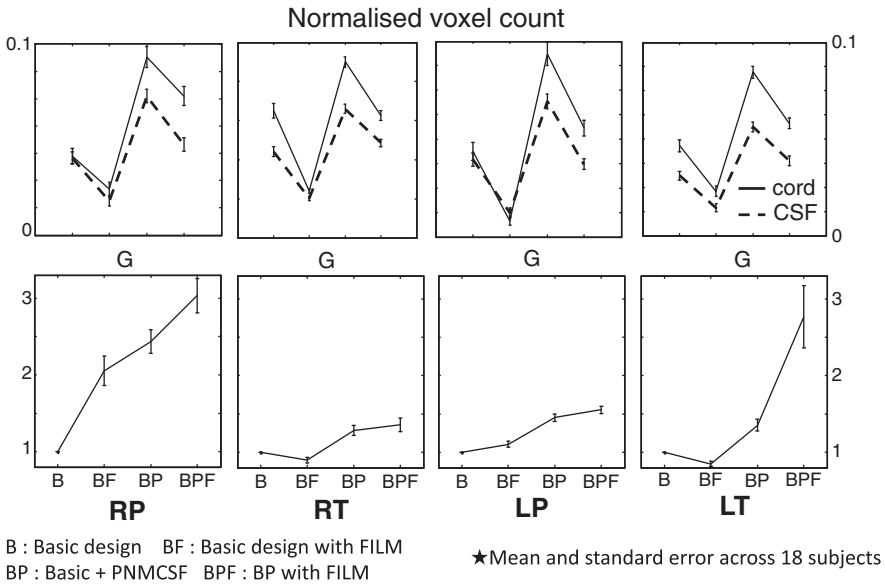


Fig. 7. How pre-whitening affects first level statistics. Upper panels: plots (with error bars) of activated voxel numbers ($p<0.01$ uncorrected) in cord (solid) and csf (dashed) areas for four different models (normalised by total voxel number in the area of each subject). Lower panels: Goodness of model (G) — ratio of activated voxel numbers in cord and csf, normalised to the values found using the basic design.

physiological noise within the framework of the general linear model. The major findings are discussed below.

PNM outperformed SAF

We were able to demonstrate the improvement in removing physiological noise from fMRI time series with a model-based technique (the PNM — a modified RETROICOR model, Brooks et al., 2008; Glover et al., 2000) compared to a technique based on selectively averaging data points acquired at similar points in the cardiac and respiratory cycle (Deckers et al., 2006). The improvement in noise modelling was assessed from the reduction in residual variance after filtering with the PNM compared to the SAF (Fig. 4), the F-test regression model comparison (Table 1) and the increased voxel counts detected within the spinal cord (Table 2).

The SAF uses a selective averaging method (FIR basis functions) to model the noise estimates according to discrete intervals (bins) in each cardiac and respiratory cycle (Deckers et al., 2006). The PNM assumes that physiological effects will be periodic, and uses a series of sine and cosine (Fourier) basis functions to best fit the data. In principle, the main difference between these two approaches is the set of basis functions used. When using low bin numbers with the SAF, it is more likely that the PNM would be a better representation of the induced physiological signal, as a Fourier basis always relates small changes in phase with smooth changes in the signal. For large bin numbers, there is likely to be little difference between the techniques, as in this situation FIR functions can be quite smooth. However large bin numbers incur a significant loss in degrees of freedom. Concerning estimation of respiratory phase, the PNM uses a histogram equalised transfer function that accounts for both the rate and *depth* of the respiratory cycle — and thus is a better model of B_0-related artefacts (Raj et al., 2001). The improved performance of the PNM over SAF/FIR-based approaches may be due to this crucial difference.

In the original paper that describes the SAF technique, in the brain, 40 and 21 bins were chosen to account for cardiac and respiratory-related signals, respectively. These numbers were derived from simulation tests on voxels with significant physiological artefacts, and were based on an identical metric (NRV). Whilst we were not able to exhaustively examine the effect of increasing bin numbers for the SAF approach, we chose a relatively high cut off of 70 bins for the cardiac and respiratory cycles. For bin numbers greater than 20 the reduction in NRV (see Fig. 4) is relatively modest. Whilst this difference between the optimal bin numbers might be due to increased physiological noise in spinal cord data compared to the brain, we believe it is more likely that when modelling complex structured noise with this technique there is no single recommendation for the optimal number of bins. The reasoning behind this is as follows: the selection of the optimal bin number is driven by both the tendency to require more bins to account for small differences in complex noise structures, and the penalty for loss of degrees of freedom (dof) – thus short time series data will encounter a dof penalty earlier than long time-series data – but the noise components may be no better modelled in either scenario. The final SAF model required 140 independent regressors, significantly higher than the PNM, which produced lower residual variance with only 32 regressors and therefore incurs a smaller penalty through loss of degrees of freedom than for the SAF model.

HR and RVT

Besides the cardiac-cycle dependent terms in the PNM and low frequency CSF regressor, a regressor describing variation in heart rate was also tested in this study. When including the HR term on top of the PNM and CSF regressors, around 29% of all voxels were found to have significantly reduced variance (Table 1). The PNM, CSF and HR regressor together reduce the NRV to 75% of its original value (Fig. 4). However, a decrease in the number of voxels identified as active was found when using a model including the HR regressor (see Table 2), in particular for painful thermal stimulation. One possible reason for this is the slight correlation ($R=-0.25$ to 0.25) between HR and the experimental design, which was confirmed by tests based on stimulus parameter estimation (beta), correlation and T-statistic changes (see Fig. 6). Three outcomes resulting from non-zero correlation between HR and the design are possible: false positives are removed (lower type I error); "true" active voxels sharing variance with HR may be removed (higher type II error); and voxels in slices where there is a negative correlation between the design and HR will have increased T-statistics. Without ground truth it is impossible to determine the relative impact of these processes, however, as always one should be careful when including a nuisance regressor that potentially correlates with your design — particularly HR in pain-related experiments.

The observation of correlation between HR and painful stimuli has important clinical neuroimaging implications. It agrees with the findings of Tousignant-Laflamme et al. (2005) that there was a link between heart rate and pain perception. However, the correlation was only found for men and not for women in that study. We did not find a gender effect in the current study. Heart rate variation has been shown to have substantial correlations to fMRI signal and strong correlations to alpha power in resting state (de Munck et al., 2008), indicating that heart rate variations may be reflected in/or respond to resting state networks. More recently an fMRI study using synchronous or delayed electrical skin stimuli showed that sensory processing in brainstem, limbic and cortical regions depends on when stimuli are experienced in relation to heart beat timing (Gray et al., 2009b). It is reasonable to expect that there might be a similar response in the spinal cord — the first site to process incoming sensory/nociceptive information.

RVT, like the HR regressor in this study, was found to remove additional variance from our fMRI data, but slightly decrease the number of voxels marked as active. In the original Birn et al. (2006) study, the regression of RVT was repeated for 51 temporal offsets (−10s to +15s) in order to account for the unknown and likely variable delay between respiration changes and BOLD responses in the brain. It should be noted that in the current study we did not shift the RVT regressor in this way. RVT correlated with HR ($R = 0.15 \pm 0.03$) and the second harmonic respiratory *cosine* term in the PNM ($R = 0.18 \pm 0.1$). Tests based on stimulus parameter estimation and correlation (see Fig. 5) showed that RVT had a greater effect on T-statistics for thermal stimuli, which could be due to alteration in breathing patterns during experiments utilising painful stimulation. As with the HR regressor, one should be careful when including an RVT regressor in studies utilising painful stimuli, in the same way as care needs to be taken when including correlated motion parameters as regressors. This is discussed more fully in the latter section on *Limitations*.

Concerning the nature of the HR induced signal change, a specific response function associated with the HR has been proposed, and when used in conjunction with a respiratory response function (rrf, Birn et al., 2008b) resulted in significant reductions of variance in resting BOLD timeseries (Chang et al., 2009). Compared to the unconvolved HR and RVT, the convolved response functions (HRcrf and RVTrrf) did not remove additional signal variance from the data. F-test results showed that HRcrf and RVTrrf accounted for unmodelled signal variance in less than 15% voxels (Table 1). When used in combination, HRcrf and RVTrrf increased active voxel counts for punctate stimulation of the left hand, and reduced counts for the other 3 stimuli (Table 2). One possible reason for this observation is that convolving with the low-frequency kernel response function decreases the correlations between HR/RVT and the experimental design.

Motion correction parameters

Commonly used physiological noise correction techniques rely critically on the timing of the image acquisition relative to the cardiac and respiratory cycle, but do not account for the direct effects of subject motion. Motion-induced fluctuations in the magnetic resonance signal will decrease the reliability of the GLM parameter estimates and therefore decrease statistical sensitivity. Motion parameters can be included in the design matrix as nuisance variables to regress out residual effects following motion correction (Friston et al., 1996; Lund et al., 2005). Considering cord motion, Summers et al. and Mikulis et al. found that it depended on position in the cardiac cycle and was principally in the rostro-caudal (S/I) direction (Mikulis et al., 1994; Summers et al., 2006). However Figley and Stroman (2007) characterised cord motion in three dimensions, and found that movement is largest in the dorso-ventral (A/P) direction. In this study, the y translation motion parameter (in the dorso-ventral direction) correlated with the first harmonic respiratory *cosine* term in the PNM ($R = -0.16$ to 0.85), which confirms that respiration has a clear impact on the spinal cord movement and mostly on dorsal-ventral direction (in the phase encoding direction of these axially acquired slices). Motion parameters significantly reduced the variance of the data (Fig. 4) and were shown to significantly account for additional signal variance for 64% of all voxels (assessed by F-test, Table 1). Moreover, they increased the number of detected active voxels for 3 out of the 4 experimental contrasts (Table 2). There was slightly correlation (−0.11 to 0.1) between the MC regressors and the experimental design, but the impact for parameter estimation and statistical results was not as great as for HR and RVT.

Motion-corrected RETROICOR was recently proposed to take slice-timing errors introduced by registration into account (Jones et al., 2008). A new set of Fourier regressors specific to each voxel were scaled by a slice motion contribution factor, and the modified regressors found to explain additional variance in the data. However, such correction schemes assume linear relationships between the detected motion and the induced signal changes, and are also difficult to implement as they require voxel wise regressors. Certainly the best way to correct for motion is to limit it in the first place — via appropriate head padding and subject preparation. In addition, prospective motion correction techniques can also be used, which correct for motion during the acquisition of an fMRI time series on a slice-by-slice basis by continuously updating the imaging volume position to follow the motion of the head (Speck et al., 2006). Unfortunately, the applicability of these techniques to the neck is currently not known.

Pre-whitening

GLM analysis in fMRI assumes that temporal noise in the time series is random (white); if it is not, then the statistical inference will be less accurate (Smith et al., 2007; Woolrich et al., 2001). In this study pre-whitening (FILM in FSL) was investigated to determine if it is beneficial, in addition to physiological noise modelling, for the analysis. Model fitting residuals revealed that pre-whitening with FILM was useful for removing extra non-white noise after physiological noise correction (Fig. 6). The activated voxel numbers in cord (solid) and CSF (dashed) areas for different models and stimuli were then investigated (Fig. 7). The PNM increases the number of activated voxels both in the cord and CSF areas. The finding that "activation" may be observed both within the cord and in the CSF space surrounding the cord is in agreement with previous animal (Zhao et al., 2008) and human data (Cohen-Adad et al., 2010). The relative size and proximity of the neuronal tissue of interest to large draining veins on the surface of the cord, makes such observations probable. Whilst we are unable to provide a "ground-truth" estimate of the correct number of true positives in our data, we can point to the relative increases in voxels marked as active in the cord and CSF. A new metric G, the ratio of the normalised activated voxel counts in the spinal cord and CSF, was defined. The G-index attempts to provide a measure of confidence in the activation data, and demonstrate the relative impact of both the PNM and pre-whitening. Fig. 7 demonstrates that although the PNM does increase the number of positives in the both the cord and CSF, the relative increases are always greatest in the cord — as indicated by positive G values, and this ratio increased when we included the PNM, and more so when including both PNM and FILM pre-whitening.

A novel finding from the G values is that pre-whitening is important for controlling false-positive rate in spinal cord fMRI analysis. In a recent study of the effects of pre-whitening (Smith et al., 2007), false-positive voxels in the brain were obtained by modelling a "pseudo-design" from resting state data. Their results suggested that pre-whitening is required for single-subject GLM analysis and fixed-effects group analysis of fMRI data. Similarly, we found that pre-whitening was beneficial for single-subject analyses by controlling for putative false positive activations in the CSF.

Limitations

A number of limitations should be considered in the current study. First, fMRI data utilising thermal and punctate stimuli, and not resting data, were used for physiological noise model evaluation. Resting state data might provide a cleaner baseline, without any complications arising from correlations with the experimental design. Second, we did not directly measure changes in CO2, and instead used a surrogate marker: RVT. Furthermore, we did not investigate temporal shifting of RVT in order to account for the likely variable delay between respiration changes and the BOLD response. Third, the cardiac (crf) and respiratory (rrf) response functions that were used were taken from brain-based studies, using the same equations and parameters, which might explain why these response functions (crf and rrf) were not found to be beneficial for this spinal cord data. Further study is needed to determine how best to model the response functions in the human spinal cord.

Finally, the decision whether to include in the GLM covariates that might potentially be correlated to the stimuli of interest analysis is still hotly debated, and well known in the context of motion parameters and stimulus correlated motion. Field et al. (2000) found that motion parameters could have a deleterious effect if correlations approached $R = 0.5$. A recent study concluded that subject motion, whether correlated to the design model or not, can have a significant impact on the sensitivity of group-level fMRI data analysis (Johnstone et al., 2006). In our study, HR ($R = -0.25$ to 0.25) and RVT ($R = -0.14$ to 0.38) correlated with the experimental design and had an impact on parameter estimation and statistical results — particularly for noxious painful stimuli. Including stimulus-correlated regressor(s) will inevitably alter estimated effect sizes, as shared variance will be split across the design regressors. The impact of this will depend on the relative effect sizes and correlations between regressors and can lead to either increases or decreases in the statistics, which might be due to changes in either sensitivity or specificity. In the absence of ground truth it is not possible to test this definitively. It should be noted that there is still no consensus on these issues for GLM analysis.

ICA-based denoising is an alternative approach for removing physiological noise effects. However, ICA requires a robust method for labeling independent components as representing artefacts to be removed or neural signals of interest to be spared. Typically, visual inspection of ICA maps would not be sufficient for unambiguous identification of noise components in the spinal cord. Both Brooks et al. (2008) and Piché et al. (2009) successfully explored the patterns of physiological noise using ICA, however, both studies used very short TR acquisitions to critically sample the physiological processes, and so could identify the noise components easily from corresponding power spectra. In general, it will be extremely difficult to distinguish physiological noise from signals of interest with ICA — where typically used TRs will lead to aliasing of the signals of interest.

Conclusions

Extending functional MRI beyond the brain to the spinal cord and brainstem is of great potential value in both basic research and clinical settings. Reducing physiological noise in sub-cortical central nervous system structures will aid interpretation of fMRI data. In this study, we explored the utility of different model-based physiological noise correction approaches via the general linear model (GLM), incorporating direct physiological measurements taken from the subject. Both physiological noise model (PNM) and selective averaging filter (SAF) significantly reduced cardiac and respiratory noise compared to not modelling physiological noise. However, PNM outperformed SAF in terms of the residual signal variation and activation statistics. Inclusion of a cerebrospinal fluid (CSF) regressor was found to be extremely beneficial as it explained a significant amount of unmodelled signal variance and increased activation statistics. Including motion correction parameters as nuisance regressors was found to have a mostly positive impact, increasing the number of activated voxels in the cord. However, including nuisance regressors that account for variation in heart rate (HR) and the rate and depth of breathing (RVT) reduced the number of activated voxels in the cord, due to the non-zero correlation with the experimental design. We present a summary of the different physiological noise models, with recommendations, in Table 3. Finally, we found that using pre-whitening in addition to physiological noise correction was advantageous in order to remove additional non-white noise from the data and control false-positive detection rates.

Acknowledgment

The authors would like to acknowledge the financial support of MRC (YK and JCWB), BBSRC David Philips Fellowship (MJ), NIH (JA) and Welcome Trust (IT). The authors would also like to thank Dr Michael Lee, Dr Catherine Warnaby and Dr Vishvarani Wanigasekera for their help in data collection.

References

Agosta, F., Valsasina, P., Rocca, M.A., Caputo, D., Sala, S., Judica, E., Stroman, P., Filippi, M., 2008. Evidence for enhanced functional activity of cervical cord in relapsing multiple sclerosis. Magn. Reson. Med. 59, 1035–1042.
Anderson, C.R., Ashwell, K.W.S., Collewijn, H., Conta, A., Harvey, A., Heise, C., Hodgetts, S., Holstege, G., Kayalioglu, G., Keast, J.R., McHanwell, S., McLachlan, E.M., Paxinos, G., Plant, G., Scremin, O., Sidhu, A., Stelzner, D., Watson, C., Charles, W., George, P., Gulgun, K., 2009. The Spinal Cord: A Christopher and Dana Reeve Foundation Text and Atlas. The Spinal Cord. Academic Press, San Diego (pp. 1–7).
Bhattacharyya, P.K., Lowe, M.J., 2004. Cardiac-induced physiologic noise in tissue is a direct observation of cardiac-induced fluctuations. Magn. Reson. Imaging 22, 9–13.
Bianciardi, M., Fukunaga, M., van Gelderen, P., Horovitz, S.G., de Zwart, J.A., Shmueli, K., Duyn, J.H., 2009. Sources of functional magnetic resonance imaging signal fluctuations in the human brain at rest: a 7 T study. Magn. Reson. Imaging 27, 1019–1029.
Birn, R.M., Diamond, J.B., Smith, M.A., Bandettini, P.A., 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. Neuroimage 31, 1536–1548.
Birn, R.M., Murphy, K., Bandettini, P.A., 2008a. The effect of respiration variations on independent component analysis results of resting state functional connectivity. Hum. Brain Mapp. 29, 740–750.
Birn, R.M., Smith, M.A., Jones, T.B., Bandettini, P.A., 2008b. The respiration response function: the temporal dynamics of fMRI signal fluctuations related to changes in respiration. Neuroimage 40, 644–654.
Bodurka, J., Ye, F., Petridou, N., Murphy, K., Bandettini, P.A., 2007. Mapping the MRI voxel volume in which thermal noise matches physiological noise — implications for fMRI. Neuroimage 34, 542–549.
Bouwman, C.J., Wilmink, J.T., Mess, W.H., Backes, W.H., 2008. Spinal cord functional MRI at 3 T: gradient echo echo-planar imaging versus turbo spin echo. Neuroimage 43, 288–296.
Brooks, J.C., Beckmann, C.F., Miller, K.L., Wise, R.G., Porro, C.A., Tracey, I., Jenkinson, M., 2008. Physiological noise modelling for spinal functional magnetic resonance imaging studies. Neuroimage 39, 680–692.
Chang, C., Glover, G.H., 2009a. Effects of model-based physiological noise correction on default mode network anti-correlations and correlations. Neuroimage 47, 1448–1459.
Chang, C., Glover, G.H., 2009b. Relationship between respiration, end-tidal CO2, and BOLD signals in resting-state fMRI. Neuroimage 47, 1381–1393.
Chang, C., Cunningham, J.P., Glover, G.H., 2009. Influence of heart rate on the BOLD signal: the cardiac response function. Neuroimage 44, 857–869.
Coghill, R.C., Price, D.D., Hayes, R.L., Mayer, D.J., 1991. Spatial distribution of nociceptive processing in the rat spinal cord. J. Neurophysiol. 65, 133–140.
Cohen-Adad, J., Gauthier, C.J., Brooks, J.C., Slessarev, M., Han, J., Fisher, J.A., Rossignol, S., Hoge, R., 2010. BOLD signal responses to controlled hypercapnia in human spinal cord. Neuroimage 50, 1074–1084.
Corfield, D.R., Murphy, K., Josephs, O., Fink, G.R., Frackowiak, R.S., Guz, A., Adams, L., Turner, R., 1999. Cortical and subcortical control of tongue movement in humans: a functional neuroimaging study using fMRI. J. Appl. Physiol. 86, 1468–1477.
Dagli, M.S., Ingeholm, J.E., Haxby, J.V., 1999. Localization of cardiac-induced signal change in fMRI. Neuroimage 9, 407–415.
de Munck, J.C., Goncalves, S.I., Faes, T.J., Kuijer, J.P., Pouwels, P.J., Heethaar, R.M., Lopes da Silva, F.H., 2008. A study of the brain's resting state based on alpha band power, heart rate and fMRI. Neuroimage 42, 112–121.
Deckers, R.H., Van Gelderen, P., Ries, M., Barret, O., Duyn, J., Ikonomidou, V.N., Fukunaga, M., Glover, G.H., De Zwart, J., 2006. An adaptive filter for suppression of cardiac and respiratory noise in MRI time series data. Neuroimage 33, 1072–1081.
Diedrichsen, J., Verstynen, T., Schlerf, J., Wiestler, T., 2010. Advances in functional imaging of the human cerebellum. Curr. Opin. Neurol. 23, 382–387.
D'Mello, R., Dickenson, A.H., 2008. Spinal cord mechanisms of pain. Br. J. Anaesth. 101, 8–16.
Eippert, F., Finsterbusch, J., Bingel, U., Büchel, C., 2009. Direct evidence for spinal cord involvement in placebo analgesia. Science 326, 404.

Field, A.S., Yen, Y.F., Burdette, J.H., Elster, A.D., 2000. False cerebral activation on BOLD functional MR images: study of low-amplitude motion weakly correlated to stimulus. AJNR Am. J. Neuroradiol. 21, 1388–1396.
Figley, C.R., Stroman, P., 2007. Investigation of human cervical and upper thoracic spinal cord motion: implications for imaging spinal cord structure and function. Magn. Reson. Med. 58, 185–189.
Figley, C.R., Stroman, P., 2009. Development and validation of retrospective spinal cord motion time-course estimates (RESPITE) for spin-echo spinal fMRI: Improved sensitivity and specificity by means of a motion-compensating general linear model analysis. Neuroimage 44, 421–427.
Friese, S., Hamhaber, U., Erb, M., Kueker, W., Klose, U., 2004. The influence of pulse and respiration on spinal cerebrospinal fluid pulsation. Invest. Radiol. 39, 120–130.
Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S., Turner, R., 1996. Movement-related effects in fMRI time-series. Magn. Reson. Med. 35, 346–355.
Friston, K.J., Josephs, O., Zarahn, E., Holmes, A.P., Rouquette, S., Poline, J., 2000. To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis. Neuroimage 12, 196–208.
Ghazni, N.F., Cahill, C.M., Stroman, P., 2010. Tactile sensory and pain networks in the human spinal cord and brain stem mapped by means of functional MR imaging. AJNR Am. J. Neuroradiol. 31, 661–667.
Giove, F., Garreffa, G., Giulietti, G., Mangia, S., Colonnese, C., Maraviglia, B., 2004. Issues about the fMRI of the human spinal cord. Magn. Reson. Imaging 22, 1505–1516.
Glover, G.H., Li, T.Q., Ress, D., 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magn. Reson. Med. 44, 162–167.
Gray, M.A., Minati, L., Harrison, N.A., Gianaros, P.J., Napadow, V., Critchley, H.D., 2009a. Physiological recordings: basic concepts and implementation during functional magnetic resonance imaging. Neuroimage 47, 1105–1115.
Gray, M.A., Rylander, K., Harrison, N.A., Wallin, B.G., Critchley, H.D., 2009b. Following one's heart: cardiac rhythms gate central initiation of sympathetic reflexes. J. Neurosci. 29, 1817–1825.
Harvey, A.K., Pattinson, K.T., Brooks, J.C., Mayhew, S.D., Jenkinson, M., Wise, R.G., 2008. Brainstem functional magnetic resonance imaging: disentangling signal from physiological noise. J. Magn. Reson. Imaging 28, 1337–1344.
Hu, X., Le, T.H., Parrish, T., Erhard, P., 1995. Retrospective estimation and correction of physiological fluctuation in functional MRI. Magn. Reson. Med. 34, 201–212.
Hutton, C., Josephs, O., Stadler, J., Featherstone, E., Reid, A., Speck, O., Bernarding, J., Weiskopf, N., 2011. The impact of physiological noise correction on fMRI at 7 T. Neuroimage 57, 101–112.
Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17, 825–841.
Jo, H.J., Saad, Z.S., Simmons, W.K., Milbury, L.A., Cox, R.W., 2010. Mapping sources of correlation in resting state FMRI, with artifact detection and removal. Neuroimage 52, 571–582.
Johnstone, T., Ores Walsh, K.S., Greischar, L.L., Alexander, A.L., Fox, A.S., Davidson, R.J., Oakes, T.R., 2006. Motion correction and the use of motion covariates in multiple-subject fMRI analysis. Hum. Brain Mapp. 27, 779–788.
Jones, T.B., Bandettini, P.A., Birn, R.M., 2008. Integration of motion correction and physiological noise regression in fMRI. Neuroimage 42, 582–590.
Kelly Jr., R.E., Alexopoulos, G.S., Wang, Z., Gunning, F.M., Murphy, C.F., Morimoto, S.S., Kanellopoulos, D., Jia, Z., Lim, K.O., Hoptman, M.J., 2010. Visual inspection of independent components: defining a procedure for artifact removal from fMRI data. J. Neurosci. Methods 189, 233–245.
Kruger, G., Glover, G.H., 2001. Physiological noise in oxygenation-sensitive magnetic resonance imaging. Magn. Reson. Med. 46, 631–637.
Lawrence, J., Stroman, P., Kollias, S., 2008. Functional magnetic resonance imaging of the human spinal cord during vibration stimulation of different dermatomes. Neuroradiology 50, 273–280.
Leitch, J.K., Figley, C.R., Stroman, P., 2010. Applying functional MRI to the spinal cord and brainstem. Magn. Reson. Imaging 28, 1225–1233.
Lilja, J., Endo, T., Hofstetter, C., Westman, E., Young, J., Olson, L., Spenger, C., 2006. Blood oxygenation level-dependent visualization of synaptic relay stations of sensory pathways along the neuroaxis in response to graded sensory stimulation of a limb. J. Neurosci. 26, 6330–6336.
Lund, T.E., 2001. fcMRI — mapping functional connectivity or correlating cardiac-induced noise? Magn. Reson. Med. 46, 628–629.
Lund, T.E., Norgaard, M.D., Rostrup, E., Rowe, J.B., Paulson, O.B., 2005. Motion or activity: their role in intra- and inter-subject variation in fMRI. Neuroimage 26, 960–964.
Lund, T.E., Madsen, K.H., Sidaros, K., Luo, W.L., Nichols, T.E., 2006. Non-white noise in fMRI: does modelling have an impact? Neuroimage 29, 54–66.
Maieron, M., Iannetti, G.D., Bodurka, J., Tracey, I., Bandettini, P.A., Porro, C.A., 2007. Functional responses in the human spinal cord during willed motor actions: evidence for side- and rate-dependent activity. J. Neurosci. 27, 4182–4190.
Mikulis, D.J., Wood, M.L., Zerdoner, O.A., Poncelet, B.P., 1994. Oscillatory motion of the normal cervical spinal cord. Radiology 192, 117–121.
Naidich, T.P., Duvernoy, H.M., Delman, B.N., Sorensen, A.G., Kollias, S.S., Haacke, E.M., 2009. Duvernoy's Atlas of the Human Brain Stem and Cerebellum. Springer.
Pattinson, K.T., Mitsis, G.D., Harvey, A.K., Jbabdi, S., Dirckx, S., Mayhew, S.D., Rogers, R., Tracey, I., Wise, R.G., 2009. Determination of the human brainstem respiratory control network and its cortical connections in vivo using functional and structural imaging. Neuroimage 44, 295–305.
Piché, M., Cohen-Adad, J., Nejad, M.K., Perlbarg, V., Xie, G., Beaudoin, G., Benali, H., Rainville, P., 2009. Characterization of cardiac-related noise in fMRI of the cervical spinal cord. Magn. Reson. Imaging 27, 300–310.
Raj, D., Anderson, A.W., Gore, J.C., 2001. Respiratory effects in human functional magnetic resonance imaging due to bulk susceptibility changes. Phys. Med. Biol. 46, 3331–3340.
Shmueli, K., van Gelderen, P., de Zwart, J.A., Horovitz, S.G., Fukunaga, M., Jansma, J.M., Duyn, J.H., 2007. Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI BOLD signal. Neuroimage 38, 306–320.
Smith, A.T., Singh, K.D., Balsters, J.H., 2007. A comment on the severity of the effects of non-white noise in fMRI time-series. Neuroimage 36, 282–288.
Speck, O., Hennig, J., Zaitsev, M., 2006. Prospective real-time slice-by-slice motion correction for fMRI in freely moving subjects. MAGMA 19, 55–61.
Stroman, P., 2005a. Magnetic resonance imaging of neuronal function in the spinal cord: spinal FMRI. Clin. Med. Res. 3, 146–156.
Stroman, P.W., 2005b. Magnetic resonance imaging of neuronal function in the spinal cord: spinal FMRI. Clin. Med. Res. 3, 146–156.
Stroman, P., 2006. Discrimination of errors from neuronal activity in functional MRI of the human spinal cord by means of general linear model analysis. Magn. Reson. Med. 56, 452–456.
Stroman, P.W., 2009. Spinal fMRI investigation of human spinal cord function over a range of innocuous thermal sensory stimuli and study-related emotional influences. Magn. Reson. Imaging 27, 1333–1346.
Stroman, P., Figley, C.R., Cahill, C.M., 2008. Spatial normalization, bulk motion correction and coregistration for functional magnetic resonance imaging of the human cervical spinal cord and brainstem. Magn. Reson. Imaging 26, 809–814.
Summers, P., Staempfli, P., Jaermann, T., Kwiecinski, S., Kollias, S., 2006. A preliminary study of the effects of trigger timing on diffusion tensor imaging of the human spinal cord. AJNR Am. J. Neuroradiol. 27, 1952–1961.
Summers, P.E., Ferraro, D., Duzzi, D., Lui, F., Iannetti, G.D., Porro, C.A., 2010. A quantitative comparison of BOLD fMRI responses to noxious and innocuous stimuli in the human spinal cord. Neuroimage 50, 1408–1415.
Toga, A.W., Mazziotta, J., 2002. Brain Mapping: The Methods. Academic Press, Elsevier Science.
Tohka, J., Foerde, K., Aron, A.R., Tom, S.M., Toga, A.W., Poldrack, R.A., 2008. Automatic independent component labeling for artifact removal in fMRI. Neuroimage 39, 1227–1245.
Tousignant-Laflamme, Y., Rainville, P., Marchand, S., 2005. Establishing a link between heart rate and pain in healthy subjects: a gender effect. J. Pain 6, 341–347.
Triantafyllou, C., Hoge, R.D., Krueger, G., Wiggins, C.J., Potthast, A., Wiggins, G.C., Wald, L.L., 2005. Comparison of physiological noise at 1.5 T, 3 T and 7 T and optimization of fMRI acquisition parameters. Neuroimage 26, 243–250.
Triantafyllou, C., Polimeni, J.R., Wald, L.L., 2011. Physiological noise and signal-to-noise ratio in fMRI with multi-channel array coils. Neuroimage 55, 597–606.
Van de Moortele, P.F., Pfeuffer, J., Glover, G.H., Ugurbil, K., Hu, X., 2002. Respiration-induced B0 fluctuations and their spatial distribution in the human brain at 7 Tesla. Magn. Reson. Med. 47, 888–895.
Willis Jr., W.D., Coggeshall, Richard E., 2003. Sensory Mechanisms of the Spinal Cord.
Wise, R.G., Ide, K., Poulin, M.J., Tracey, I., 2004. Resting fluctuations in arterial carbon dioxide induce significant low frequency variations in BOLD signal. Neuroimage 21, 1652–1664.
Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M., 2001. Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage 14, 1370–1386.
Yoshizawa, T., Nose, T., Moore, G.J., Sillerud, L.O., 1996. Functional magnetic resonance imaging of motor activation in the human cervical spinal cord. Neuroimage 4, 174–182.
Zhao, F., Williams, M., Meng, X., Welsh, D.C., Coimbra, A., Crown, E.D., Cook, J.J., Urban, M.O., Hargreaves, R., Williams, D.S., 2008. BOLD and blood volume-weighted fMRI of rat lumbar spinal cord during non-noxious and noxious electrical hindpaw stimulation. Neuroimage 40, 133–147.