# Data-driven denoising in spinal cord fMRI with principal component analysis

Kimberly J. Hemmerling[1,2], Andrew D. Vigotsky[2,3], Charlotte Glanville[2], Robert L. Barry[4,5,6], Molly G. Bright[1,2]

[1] Department of Physical Therapy and Human Movement Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

[2] Department of Biomedical Engineering, McCormick School of Engineering, Northwestern University, Evanston, IL, USA

[3] Department of Statistics & Data Science, Weinberg College of Arts and Sciences, Northwestern University, Evanston, IL, USA

[4] Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA, USA

[5] Harvard Medical School, Boston, MA, USA

[6] Harvard-Massachusetts Institute of Technology Health Sciences & Technology, Cambridge, MA, USA

## ABSTRACT

Numerous approaches have been used to denoise spinal cord functional magnetic resonance imaging (fMRI) data. Principal component analysis (PCA)-based techniques, which derive regressors from a noise region of interest (ROI), have been used in both brain (e.g., CompCor) and spinal cord fMRI. However, spinal cord fMRI denoising methods have yet to be systematically evaluated. Here, we formalize and evaluate a PCA-based technique for deriving nuisance regressors for spinal cord fMRI analysis (SpinalCompCor). In this method, regressors are derived with PCA from a noise ROI, an area defined outside of the spinal cord and cerebrospinal fluid. A parallel analysis is used to systematically determine how many components to retain as regressors for modeling; this designated a median of 11 regressors across three fMRI datasets: motor task (n=26), breathing task (n=27), and resting state (n=10). First-level fMRI modeling demonstrated that principal component regressors did fit noise (e.g., physiological noise from blood vessels), particularly in the resting state fMRI dataset. However, group-level motor task activation maps themselves did not show a clear benefit from including SpinalCompCor regressors over our original denoising model. The potential for collinearity of principal component regressors with the task may be a concern, and this should be considered in future implementations for which task-correlated noise is anticipated.

Keywords: spinal cord, functional MRI, denoising, principal component analysis, nuisance regression, physiological noise

## 1 INTRODUCTION

Spinal cord functional magnetic resonance imaging (fMRI) can provide valuable insight into neural processes of the human spinal cord; however, well-known challenges include the influence of physiological noise and motion. Spinal cord voxels are highly susceptible to artifacts from cardiac and respiratory sources (Eippert et al., 2017). The nearby dilation of arteries after systole as well as the cardiac-linked pulsatile flow of cerebrospinal fluid (CSF) surrounding the cord confound the fMRI signal timeseries (Piché et al., 2009). Additionally, nearby inflation of the lungs and movement of the chest cavity during ventilation greatly impacts the homogeneity of the $B_0$ magnetic field and, thus, fMRI signals within the spinal cord itself (Raj et al., 2001).

Robust approaches to denoising are of utmost importance for spinal cord fMRI. One common technique for combatting physiological noise confounds is RETROICOR, a model-based physiological denoising technique

developed for brain fMRI (Glover et al., 2000), with adaptations to the spinal cord (Brooks et al., 2008; Kong et al., 2012). Fourier-based RETROICOR regressors are modeled from cardiac and respiratory signals externally recorded during scanning and then regressed out of the fMRI data. Data-driven nuisance regressors can also be generated in specific regions of interest (ROIs) (e.g., CSF (Brooks et al., 2008; Kong et al., 2012)) or via independent component analysis methods (Hu et al., 2018; Xie et al., 2012). Motion correction and the subsequent use of motion regressors in fMRI models are commonplace (see Eippert et al. 2017 for a detailed review of approaches to denoising in spinal cord fMRI (Eippert et al., 2017)).

Another approach is to use principal component analysis (PCA) within a subset of imaging voxels to derive multiple nuisance regressors for fMRI denoising. CompCor, a PCA-based denoising technique for brain fMRI, defines a noise ROI in which PCA is applied to obtain nuisance regressors (Behzadi et al., 2007). This noise ROI is defined either anatomically (i.e., in white matter and CSF) or based on voxel timeseries with the highest temporal standard deviation. Another method uses PCA in brain edge voxels to define regressors and improve removal of motion confounds in resting state brain fMRI data (Patriat et al., 2015).

CompCor methods for brain fMRI data have inspired related methods for spinal cord fMRI analysis. For example, in an implementation of PCA-based denoising for 7 T resting state spinal cord fMRI, a noise ROI was defined as anywhere external to the cord and its adjacent CSF; then, PCA of voxel timeseries within this noise ROI was used to produce nuisance regressors (Barry et al., 2014, 2016). In this technique, components were selected based on the percent of signal variance explained and were then regressed out from the data prior to motion correction to improve its efficacy. Although this method has been previously applied, it has yet to be rigorously evaluated for spinal cord fMRI more generally.

In this work, we extend on the method initially proposed for 7 T resting state spinal cord fMRI (Barry et al., 2014, 2016), and we will henceforth refer to PCA-based denoising for spinal cord fMRI as “SpinalCompCor”. For SpinalCompCor denoising, we assume that signal outside of the cord is non-neural and that this noise similarly affects the fMRI timeseries both within the spinal cord and in surrounding tissues. Under this assumption, principal components can be derived from the timeseries in these typically unused voxels and harnessed as a tool for denoising of spinal cord fMRI data.

We first define our method to calculate the principal components from a noise ROI in spinal cord fMRI data. It is important to identify meaningful principal components to include as regressors in the fMRI noise model. Thus, we employ a parallel analysis to determine, in a given dataset, how many components should be retained as nuisance regressors. This parallel analysis systematically determines how many components to retain by comparing principal component eigenvalues from real data to those from surrogate data, essentially identifying the number of components that explain greater variance than the surrogate components do by chance. The parallel analysis is evaluated in two 3 T task fMRI datasets and one 3 T resting state fMRI dataset. Subject-level fMRI analysis is subsequently performed in these three datasets to assess the benefit of including the selected SpinalCompCor regressors in fMRI modeling. Finally, we evaluate the effect of including data-driven SpinalCompCor regressors for group-level motor activation mapping in a hand-grasping task fMRI dataset, for which there is an anatomically defined *a priori* hypothesis of activation locale.

## 2 METHODS

Three previously collected spinal cord fMRI datasets were used in this study, including a hand-grasping task, a breath-holding task, and a resting state acquisition. The isometric unilateral hand-grasping task was expected to evoke motor activation localized to the ventral horns of the C7-C8 spinal cord segments within the field of view (Hemmerling et al., 2023a; Vanderah & Gould, 2016). The breath-holding task was designed to induce hypercapnia and evoke systemic vasodilation, from which spinal cord vascular reactivity could be estimated (Hemmerling et al., 2024a). The resting state acquisition was expected to be sensitive to intrinsic signal fluctuations and functional networks in the spinal cord (Barry et al., 2018). In all three datasets, we expect SpinalCompCor regressors to capture noise associated with physiology, non-physiological motion, and other artifacts. The hand-grasping and breath-holding tasks may additionally exhibit task-correlated motion as well as physiological changes (e.g., heart rate, chest position) related to the task, while the resting state may be expected to have more subtle motion and physiological fluctuations. Thus, we aim to explore any differences in the selection of principal component regressors and subject-level model performance with the new regressors across a few fMRI acquisitions.

### 2.1 Data collection

Data were collected in studies approved by either the Northwestern University Institutional Review Board (hand-grasping and breath-holding task datasets) or the Vanderbilt University Institutional Review Board (resting state data). Informed consent was obtained from all participants for being included in the studies.

#### 2.1.1 Hand-grasping task fMRI

The hand-grasping task dataset was collected at the Northwestern University Center for Translational Imaging on a 3 T Siemens Prisma MRI system (Siemens Healthcare, Erlangen, Germany) with a 64-channel head/neck coil and SatPad™ cervical collar from 30 healthy participants (25.9 ± 4.5 years, 19 F) (Hemmerling et al., 2023b, 2023a). Four subjects were excluded because of excessive image artifacts, an incidental finding, and technical issues (included sample: n=26, 25.8 ± 4.5 years, 17 F).

A high-resolution T2-weighted sagittal anatomical MRI was acquired (repetition time (TR)/echo time (TE) = 1500/135 ms, sagittal slice thickness = 0.8 mm, in-plane resolution 0.39 $mm^2$, 64 slices, flip-angle (FA) = 140°, field-of-view (FOV) = 640 $mm^2$). The functional sequence was a gradient-echo echo-planar imaging (EPI) sequence with ZOOMit selective excitation (TR/TE = 2000/30 ms, 1 × 1 × 3 $mm^3$, 25 slices, FA = 90°, FOV = 128 × 44 $mm^2$). Cervical cord coverage was approximately the C4-C7 vertebral levels.

The 10-minute (300 volume) hand-grasping task run consisted of 24 trials of 15-second pseudo-randomized left and right grasps. Participants viewed real-time force feedback while squeezing custom MRI-safe handgrips, targeting 25% maximum-voluntary contraction (MVC). Physiological data (respiration, pulse, and exhaled $CO_2$) were collected during the session. The first of the two fMRI runs acquired in the original study was used in the parallel analysis and subject-level modeling; both runs were used for group-level activation mapping. Respiration belt data could not be collected for one subject.

#### 2.1.2 Breath-holding task fMRI

The breath-holding task dataset was collected from the same participants during the same sessions as the hand-grasping task (Hemmerling et al., 2024b, 2024a). Three subjects were excluded because of excessive image artifacts and an incidental finding (included sample: n=27, 26 ± 5 years, 18 F).

The 6.8-minute (205 volume) breath-holding task scan was acquired with the same parameters as the hand-grasping scan except for the number of volumes. The task run began with a 20-second rest, followed by 7 breath-

holding cycles, and ended with a 30-second rest. Each cycle had 24 seconds of paced breathing, an 18-second end-expiration breath-hold, a 2-second exhalation, and a 6-second recovery period. Physiological data (respiration, pulse, and exhaled $CO_2$) were collected during the session. The first of the two fMRI runs acquired in the original study was used here. Respiration belt data could not be collected for one subject.

#### 2.1.3 Resting state fMRI

The resting state dataset was collected at the Vanderbilt University Institute of Imaging Science on a Philips Achieva 3 T scanner (Best, The Netherlands) from 10 healthy participants (27±4.9, 5 F) (Barry et al., 2018).

A sagittal anatomical localizer scan was acquired. The functional sequence was a 3D gradient-echo multi-shot sequence (TR = 34 ms, TE = 8 ms, acquired voxel size = 1 × 1 × 5 $mm^3$, interpolated voxel size 0.59 x 0.59 x 5 $mm^3$, 12 slices, FA = 90°, FOV = 150 $mm^2$, echo train length=7, SENSE = 2.0 (left-right), volume acquisition time (VAT) = 2.76 s). The cervical cord coverage was from approximately the C2-C5 vertebral level.

A 20-minute (434 volume) resting state run was acquired. In the original work, these data were referred to as the “full-k” data in the “scan length” study (Barry et al., 2018). The other functional data also collected in that study were not used here. Physiological data (respiration, pulse) were collected during the session. Pulse data were not used for one subject due to poor quality; both respiration and pulse data were not available for one other subject.

### 2.2 Data analysis

The FMRIB Software Library (FSL; version 6.0.3), Spinal Cord Toolbox (SCT; version 6.1), Neptune Toolbox (version 1.211227) (Deshpande & Barry, 2022), AFNI (version 22.0.05), MATLAB (MathWorks, Natick, MA, R2019b), and R (version 4.3.1) were used to perform data analyses (unless otherwise noted).

#### 2.2.1 Motion correction

A mask was manually drawn around the spinal cord and CSF region, which is used to compute a Gaussian weighting mask for motion correction. Note, when using the term “CSF” throughout this paper, this refers to the CSF in the spinal subarachnoid space, unless otherwise stated. Slice-wise motion correction was performed with the Neptune Toolbox and AFNI as previously described (Hemmerling et al., 2023a). Note, one difference is that temporal median filtering was de-selected for the task datasets but retained for the resting state dataset because larger movements may be sporadic during rest (Barry et al., 2016).

#### 2.2.2 Noise ROI definition and SpinalCompCor regressor generation

Regressors were derived from noise ROI voxels external to the spinal cord and CSF region. These voxels are expected to contain signal fluctuations from noise sources that similarly affect the timeseries within the cord. Note that noise associated with CSF signal fluctuations was addressed by generating a separate, CSF-specific regressor (see section 2.2.4), as this is a common practice in the literature **(Table S1)**. To create the noise ROI, the mask drawn for motion correction around the spinal cord and CSF was dilated by 18 mm, rounded to the nearest whole voxel (*sct_maths*). The spinal cord/CSF mask was subtracted from the dilated region, and any voxels at the edge of the slice field-of-view (up to 3 voxels) were removed from the final noise ROI mask (*fslmaths*). Examples of this noise ROI are shown in **Fig. S1**. All voxel timeseries from within the noise ROI were output (*fslmeants*) and arranged in a matrix (rows: time points, columns: voxels). Slice-wise PCA by singular value decomposition was performed on this matrix with MATLAB (*pca*). These steps are depicted schematically **(Fig. 1A-C)**. Components were then selected from the ordered principal components to be SpinalCompCor regressors.

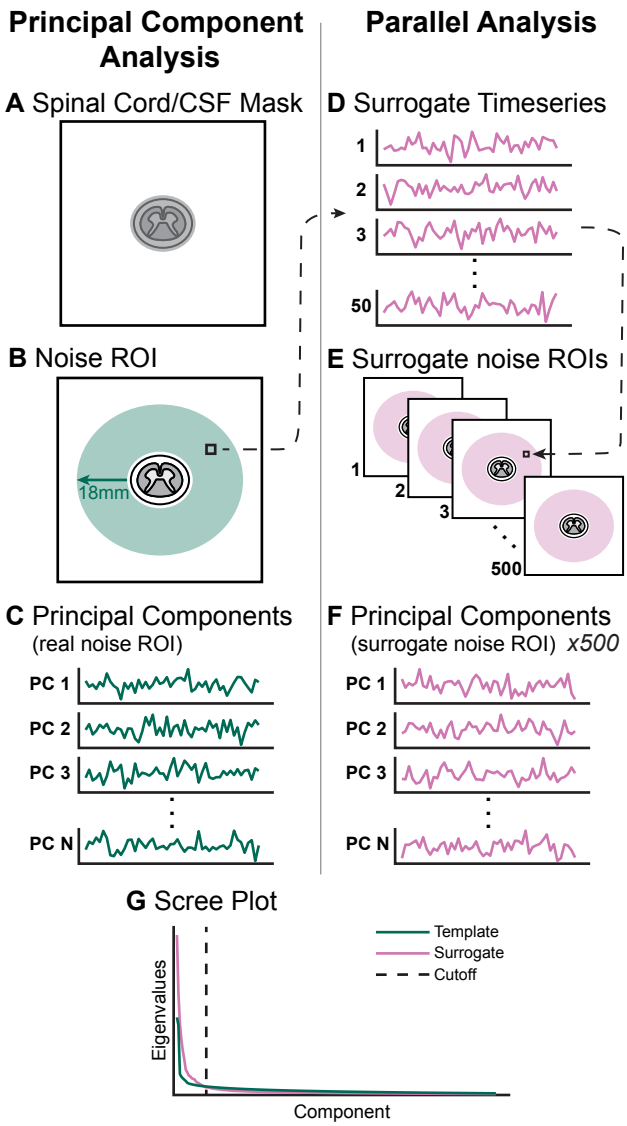


***Fig 1. Schematic overview of (left) principal component analysis and (right) parallel analysis methods for each slice. (A)*** *The manually drawn spinal cord/CSF contour.* ***(B)*** *The noise ROI.* ***(C)*** *Slice-wise principal component analysis is performed in the noise ROI, generating N components (N=# Timepoints-1) per imaging slice.* ***(D)*** *Surrogate timeseries (50 per voxel) generated from the real voxel timeseries.* ***(E)*** *Surrogate slice noise ROI timeseries matrices (500) generated by randomly sampling the surrogate voxel timeseries.* ***(F)*** *Principal component analysis is performed in the surrogate slice noise ROI, generating N components. This is repeated for each of the 500 surrogate noise ROIs.* ***(G)*** *Scree plot with mean surrogate and template curves; the indicated cutoff is at the intersection of these two curves. Note, the timeseries shown in panels C, D, and F are for illustrative purposes and are not actual timeseries used in this work.*

### 2.2.3 Surrogate data generation and parallel analysis

To systematically determine how many principal components may be reasonable to retain, a parallel analysis was performed comparing real principal component eigenvalues to those generated from simulated data. An

Iterative Amplitude Adjusted Fourier Transform (IAAFT) method was used to generate surrogate timeseries data[1]. Surrogate timeseries share some statistical properties with the original timeseries; in this case, the IAAFT algorithm preserves amplitude and power spectrum traits of the original, but with random phases (Venema et al., 2006). Since the phases are randomly shifted, these surrogate timeseries are, on average, uncorrelated with the original timeseries on which they are based. As such, surrogate timeseries can be used to generate "null" principal components that properly account for temporal dynamics.

For a given template (real) voxel timeseries, an initial surrogate timeseries was first created by shuffling datapoints in the template voxel timeseries. The IAAFT algorithm then iteratively performs spectral adjustment and amplitude adjustment until convergence, or the maximum number of iterations (500) was reached (Venema et al., 2006). This was repeated for each template voxel timeseries in each slice of the noise ROI such that 50 surrogate timeseries were generated from each template (i.e., real) timeseries.

500 surrogate noise ROI timeseries matrices were generated by randomly sampling from the 50 surrogate timeseries for each voxel. Slice-wise PCA was run on the template and each of the 500 surrogate noise ROIs. A scree plot was generated as follows: (1) the template Eigenvalues and mean surrogate Eigenvalues were plotted against the component number; (2) the component number at which these curves intersected was selected as the cutoff number of principal components for that slice. This cutoff represents the last template component that explains greater variance than the surrogate components do by chance. These steps are depicted schematically **(Fig. 1D-G)**. (Note, MATLAB version r2022b was used for the parallel analysis because of software availability on the high-performance computing cluster.)

The parallel analysis could be applied to any given dataset to determine a cutoff number of components. The code necessary to generate SpinalCompCor regressors and perform the parallel analysis with IAAFT is provided alongside this publication. Alternatively, the Neptune toolbox for spinal cord fMRI could also be used to perform PCA-based denoising (Deshpande & Barry, 2022).

We implemented the parallel analysis in task and rest fMRI datasets to evaluate whether a different number of components would be designated to retain as SpinalCompCor regressors across these different fMRI acquisitions. Timeseries were first truncated to 205 timepoints to match the breath-holding task data (the dataset with the fewest timepoints) so that the number of timepoints would not influence the number of principal components in the following parallel analysis. The parallel analysis was then implemented as described above. For each fMRI run, the median cutoff number of components across slices was selected.

### 2.2.4 Subject-level models with SpinalCompCor regressors

The downstream effect of using SpinalCompCor regressors in subject-level fMRI modeling was tested in the breath-holding task, hand-grasping task, and resting state dataset. The following additional preprocessing was performed (these steps are more thoroughly described in previous work from our lab (Hemmerling et al., 2023a)). First, additional nuisance regressors were generated. Physiological data were preprocessed via a bespoke MATLAB script to detect peaks in the respiratory belt, pulse, and $CO_2$ traces. An end-tidal $CO_2$ ($P_{ET}CO_2$) regressor (if applicable) and slice-wise RETROICOR (Brooks et al., 2008; Glover et al., 2000) regressors (8 respiratory and 8 cardiac) were generated from the preprocessed physiological data. A slice-wise CSF regressor was generated from CSF voxels with the top 20% temporal variance (Brooks et al., 2008; Kong et al., 2012). Left and right "%MVC" hand-grasp task regressors were also generated by normalizing the force trace to the participant's MVC

[1] The IAAFT algorithm described here was based on a webpage that is no longer accessible. The cited work is from the same author; the IAAFT description can be found in Venema et al. 2006 (section 2).

(if applicable). Then, a manual spinal cord mask was drawn on the motion-corrected mean functional image. The high-resolution anatomical image was segmented and registered to the PAM50 template (De Leener et al., 2018; Gros et al., 2019). Then the functional data were also registered with the PAM50 template using the spinal cord mask and initial warp from the anatomical registration.

The subject-level analysis was performed with FSL's FILM general linear model using the FEAT tool (Jenkinson et al., 2012) (FILM prewhitening; high-pass filter cutoff = 100 s). The hand-grasp task models included 2 task regressors (L, R) and 20+ nuisance regressors (X and Y motion (MOCO), CSF, $P_{ET}CO_2$, 8 cardiac and 8 respiratory RETROICOR (RETRO)). A Right Grasp > 0 and Left Grasp > 0 contrast were defined. Two hand-grasp task models were computed:

(1) "Base" model with no SpinalCompCor components (20 nuisance regressors)

$$Y_{Base} \sim X_R + X_L + X_{MOCO} + X_{CSF} + X_{PETCO2} + X_{RETRO}$$

(2) "SpinalCompCor" model with additional PCA-derived regressors (>20 nuisance regressors)

$$Y_{SpinalCompCor} \sim X_R + X_L + X_{MOCO} + X_{CSF} + X_{PETCO2} + X_{RETRO} + X_{PCA}$$

The breath-hold model was computed similarly, but with $P_{ET}CO_2$ as the task regressor and without hand-grasp regressors; a $P_{ET}CO_2 > 0$ contrast was defined. The resting state model did not include hand-grasp or $P_{ET}CO_2$ regressors.

The median number of principal components chosen by the parallel analysis across slices was selected for each dataset to avoid creating models with different numbers of regressors in each slice. Subject-level models were computed on the full timeseries, so the parallel analysis was first also run on the full timeseries before selecting the number of component regressors for modeling. Models were only computed for voxels within a spinal cord mask to minimize computational intensity. For a few fMRI runs, the GLM was rerun without a mask (i.e., in the entire field of view) so that spatial information about the principal component regressors could be examined outside of the cord.

### 2.2.5. Evaluation of nested models

Next, we addressed whether adding principal component noise regressors derived from outside the spinal cord successfully explained variance inside the spinal cord. A voxel-wise *F*-test was performed between the full (SpinalCompCor) and reduced (Base) models. PAM50 template gray matter and white matter masks were warped to the native space for each run, thresholded (at 0.5), and binarized to provide subject-specific tissue masks. Median *F*-statistics and the proportion of statistically significant voxels, as determined by the *F*-test ($p<0.05$, not corrected for multiple comparisons), were computed in these masks.

Considering the potential spinal cord segment-dependent effect observed in the *F*-test results, we wanted to evaluate whether there was in fact a segment-dependent benefit to using the SpinalCompCor model compared to the Base model. We used a mixed effects generalized additive model with a quasibinomial family and logit link function, which was fit using restricted maximum likelihood (REML). The dependent variable, proportion of significant voxels, was regressed on spinal cord level, task, and tissue type as fixed effects, and subject was included as a random effect, saving the parameter estimate associated with spinal cord level as the test statistic. Spinal cord level was coded as −1, 0, and +1 for C6, C7, and C8, respectively, for hand-grasp and breath-hold datasets, or as –1, 0, and +1 for C4, C5, and C6, respectively, for the resting state dataset. To build a null

distribution, we relied on permutation testing (5000 permutations). Within each permutation, each subject's spinal cord levels were randomly "flipped" (rostrocaudal vs. caudorostral) via sign flipping (*e.g.*, −1, 0, and +1 *or* +1, 0, and −1 for C6, C7, and C8, respectively), which would preserve within-participant autocorrelation across adjacent levels. The spinal cord level parameter estimate (test statistic) from each permutation was saved, and the observed test statistic was compared to this null distribution to calculate a P-value.

#### 2.2.6. Group modeling and comparisons

Only the hand-grasp task dataset was carried through to the group-level analysis because it had the simplest *a priori* hypothesis for localization of hand-grasping activation. For this analysis, each participant's second fMRI run was also considered (i.e., the parallel analysis and subject-level modeling described above were also performed on these runs). Parameter estimate maps associated with the hand-grasping task for each run were averaged for each subject and then concatenated across subjects. For group-level analysis, a non-parametric one-sample *t*-test with threshold-free cluster enhancement was performed with 5000 permutations (*randomise*) (Nichols & Holmes, 2002; Smith & Nichols, 2009; Winkler et al., 2014). This was performed for each model (Base, SpinalCompCor) and contrast (Right Grasp > 0, Left Grasp > 0). The proportion of voxels that were "Now active" and "No longer active" with the SpinalCompCor model compared to the Base model in ROI masks were extracted.

## 3 RESULTS

### 3.1 Parallel analysis

A parallel analysis was run to determine the number of principal components that should be considered as regressors for fMRI modeling. This analysis generated a scree plot for each slice of every fMRI run, from which the cutoff number of principal components was recorded (see **Fig. 1G** as an example). The median cutoff number of principal components across slices is shown in **Fig. 2A** for all three datasets. The distribution across slices shows that the cutoff number of components is relatively consistent across most subject/slices **(Fig. 2B)**. However, there are areas of higher variability, e.g., in the inferior-most slices for the hand-grasp task dataset. It is pertinent to note that the task fMRI data (breath-hold, hand-grasp) were acquired on a Siemens scanner using a sequence with selective excitation in an inner FOV with a TR of 2s. Selective excitation should lessen susceptibility artifacts in areas near air cavities (Rieseberg et al., 2002), such as the spinal cord. The resting state fMRI data were acquired on a Philips scanner, did not use a sequence with selective excitation, and had a VAT of 2.76s. While all data were acquired in the cervical cord, the resting state imaging volume was also centered more rostrally (approximately vertebral levels C2-C5) than the other acquisitions (approximately vertebral levels C4-C7), which may impact the specific noise characteristics captured by PCA.

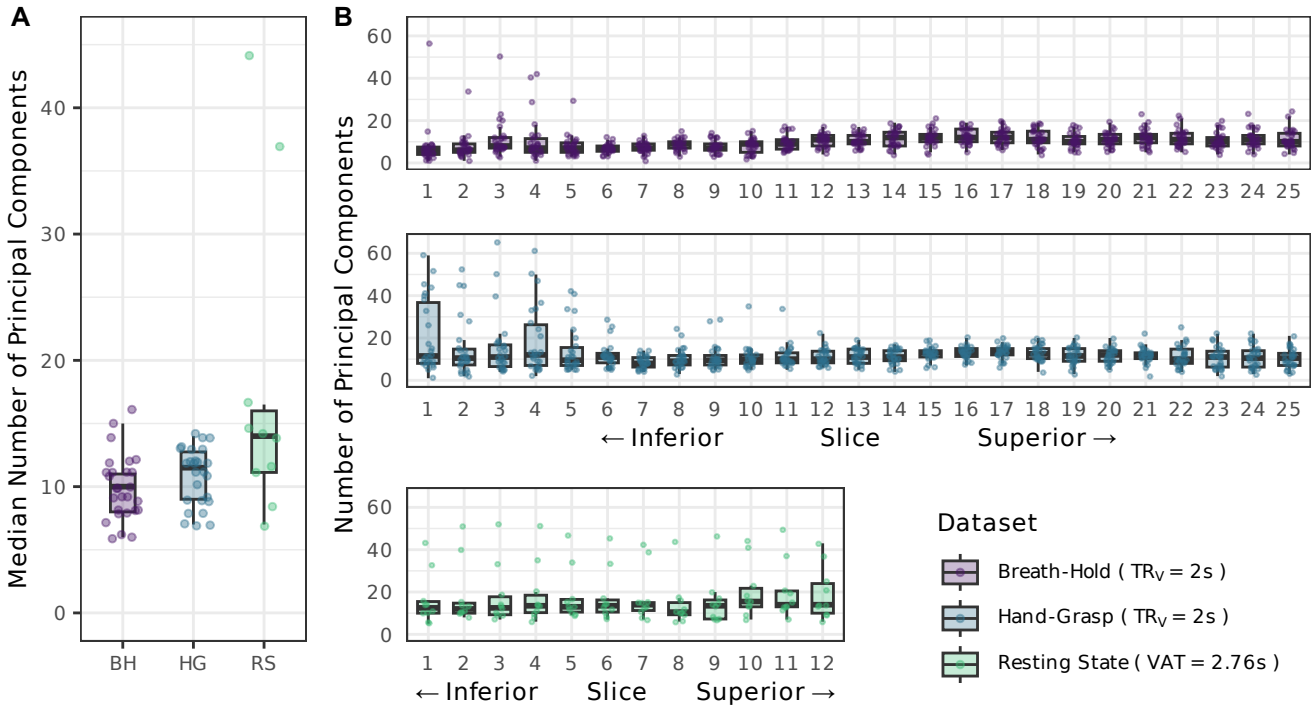


***Fig. 2. Parallel analysis summary across datasets. (A)*** *Median cutoff number of principal components for each fMRI run (each datapoint represents the parallel analysis result across slices for one subject's fMRI run).* ***(B)*** *The cutoff number of principal components in each subject and slice. Note, the parallel analysis was performed on data restricted to the same number of timepoints (i.e., fMRI data that were originally longer than 205 timepoints were truncated).*

Longer fMRI runs were truncated to 205 timepoints for this parallel analysis, so the PCA performed here was not biased by the scan length. Having more timepoints does, however, influence the cutoff number of principal components selected in the parallel analysis. For example, the hand-grasp task dataset's original length is 300 timepoints, and truncating the dataset to 205 timepoints decreases the median number of cutoff components for most subjects' scans **(Fig. S2)**.

There are also two notable outliers in the median cutoff number of principal components for the resting state dataset. To explore this, the median number of principal components was also compared to the in-plane framewise displacement **(Fig. S3A)**. We see a significant negative trend between number of components and overall motion across all datasets ($r$=-0.76, $p$=0.01) and observe that the two outlier scans with the highest median number of principal components also had some of the lowest overall motion. Motion may also be related to the magnitude of the eigenvalue associated with the first principal component. **Fig. S3B** shows that these outliers have the lowest first eigenvalue, however the positive correlation between these elements is non-significant ($r$=0.32, $p$=0.37).

### 3.2 Spatial information from component regressors

Maps of the fitted parameter estimates show how variance associated with each SpinalCompCor regressor was fit across the entire cord and neck region. For the datasets in which the subject-level model was rerun on the

entire field of view, stronger parameter estimates were observed in a few common areas: neck muscle, blood vessels, and CSF. Examples from hand-grasp scans are shown in **Fig. 3** (A: blood vessel/CSF; B: blood vessel/CSF; C: no predominant area; D: neck muscle) and **Fig. S4**, from a breath-hold scan in **Fig. S5**, and from a resting state scan in **Fig. S6**.

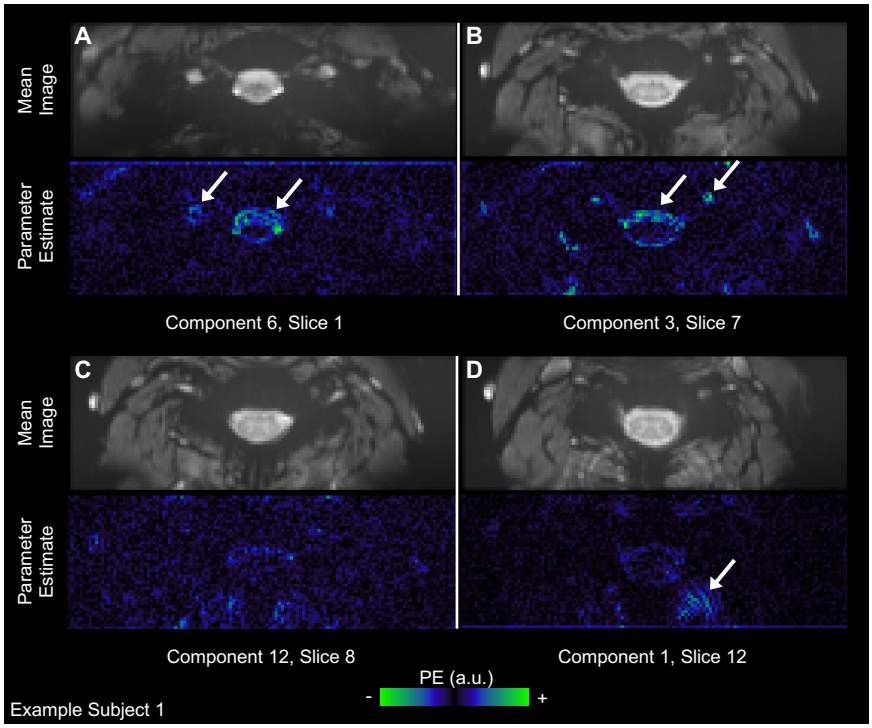


***Fig. 3. Parameter estimate fits for SpinalCompCor regressors in an example subject from the hand-grasp task dataset.*** *The motion corrected mean image and corresponding principal component parameter estimate maps are shown for each example component/slice. Arrows point out stronger parameter estimates in areas with neck muscle, blood vessels, and CSF. There are 25 slices in this acquisition, numbered 0-24. (PE=parameter estimate, a.u.=arbitrary units)*

### 3.3 Evaluation of nested models

Two models were calculated for each breath-hold, hand-grasp, and resting state fMRI run to explore the effect of adding SpinalCompCor regressors. Differences between nested Base and SpinalCompCor models (i.e., the reduced and full models, respectively) were evaluated with a voxel-wise *F*-test. Summary metrics were extracted from gray matter and white matter ROIs in spinal cord segments in the native subject space (within tissue masks warped from the PAM50 template). Specifically, the median *F*-statistic and the proportion of statistically significant voxels (*F*-test) were calculated in ROIs for each fMRI run **(Fig. 4)**.

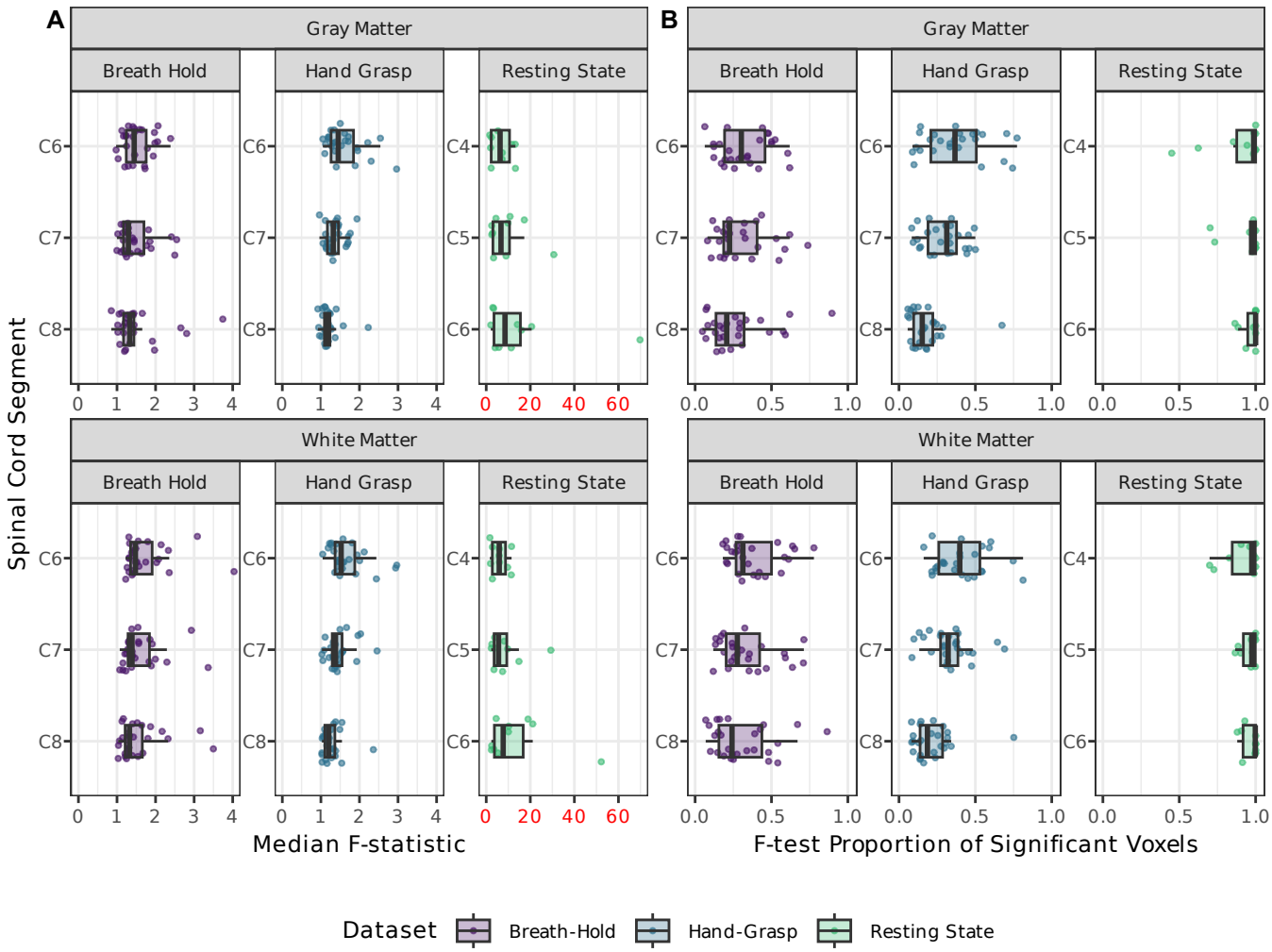


***Fig. 4. Summary of F-test comparing the nested Base and SpinalCompCor models.*** *Summary statistics are shown in ROIs of gray matter and white matter across spinal cord segments C6-C8 or C4-C6.* ***(A)*** *Median F-statistic in each of the spinal cord ROIs.* ***(B)*** *The proportion of voxels in the ROI for which the SpinalCompCor model significantly fit additional variance compared to the Base model (p<0.05, not corrected for multiple comparisons). Each datapoint represents one fMRI run. Resting state median F-statistic axis is colored red to draw attention to the scaling.*

For both the breath-hold and hand-grasp task datasets, the median *F*-statistic appears to decrease slightly from superior to inferior (C6 to C8) in both gray and white matter. The proportion of statistically significant voxels in the *F*-test similarly appears to decrease from superior to inferior in both gray and white matter. In the resting state dataset, the percent of significant voxels (*F*-test) clusters near 100% and may also differ across levels. Assessing this using a mixed effects model and flipping the spinal cord segments, we can reject the null that all spinal cord segments benefit similarly ($p < 0.0002$).

### 3.4 Group analysis

A group analysis was performed on the hand-grasp task data to test the effect of SpinalCompCor regressors on group-level motor activation, comparing results from the Base and SpinalCompCor models. Example axial slices of the hand-grasp activation map are shown for the Left>0 and Right>0 contrasts for both models **(Fig. 5A,B)**. To directly compare the spatial distribution of the difference in activation between the two models, the voxels that changed between conditions were mapped and summarized across the level and tissue ROIs **(Fig. 5C,D)**. These panels show voxels that were not active for the Base model but are now active for the SpinalCompCor model (blue), voxels which were active for the Base model but are no longer active for the SpinalCompCor model (red), and voxels that were active for both (gray). A higher proportion of voxels in the dorsal gray matter are now active with the SpinalCompCor model, while the opposite is seen for the ventral gray matter. Often, the voxels that change between conditions are at the edge of a cluster of active voxels.

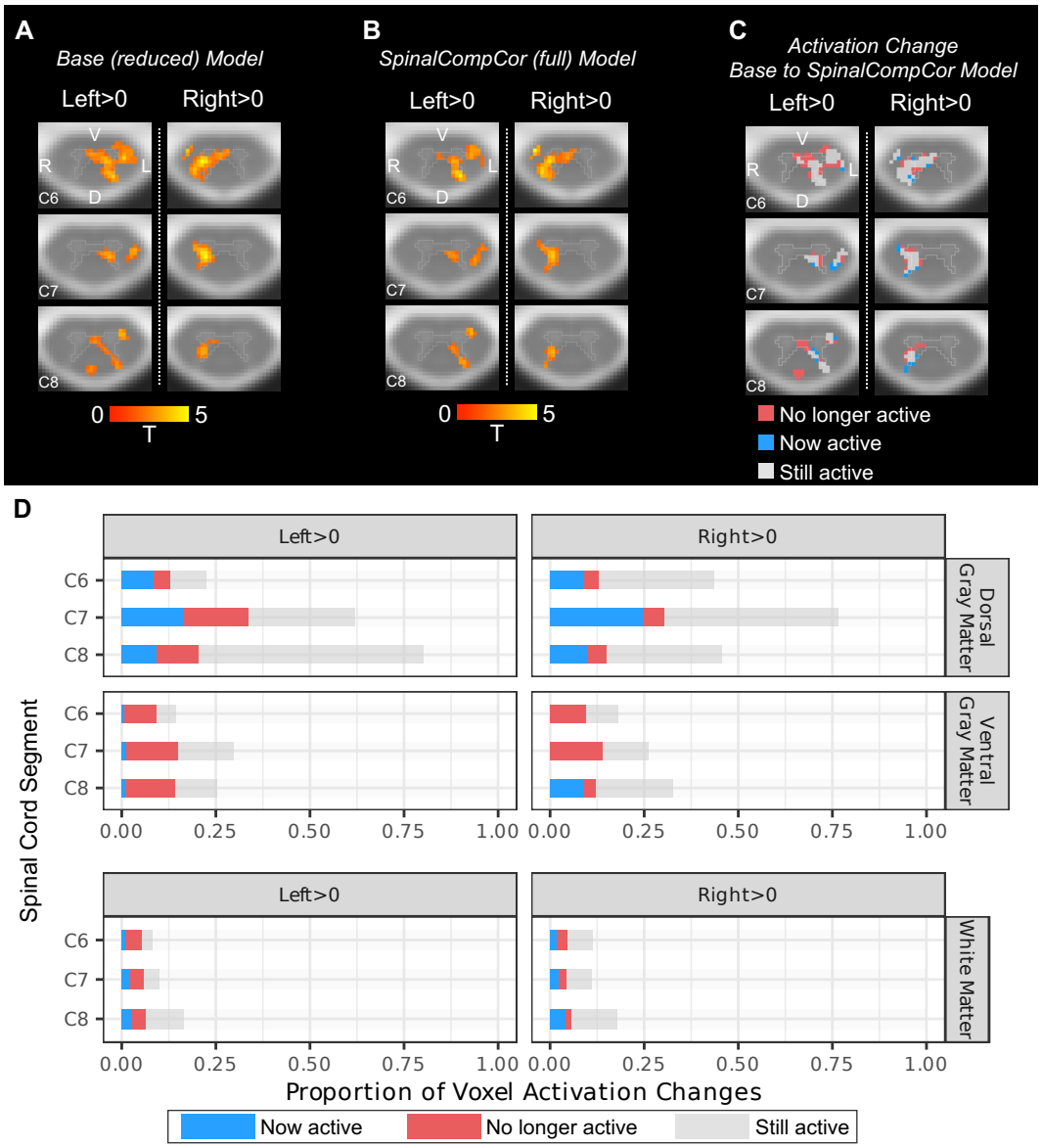


***Fig. 5. Comparison of Base vs. SpinalCompCor model group analysis for hand-grasping task.*** *Example activation map axial slices are shown for Left>0 and Right>0, for the* ***(A)*** *Base model and* ***(B)*** *SpinalCompCor*

*model. Significant t-statistics are shown (p<0.05, Family-wise error rate corrected).* ***(C)*** *Example axial slices of the activation change for the Base vs. SpinalCompCor models. Axial slices are shown in radiological view/orientation and have an outline of the PAM50 gray matter mask overlaid. The same three slices are shown for panels A, B, and C.* ***(D)*** *Voxel-wise activation change for the Base vs. SpinalCompCor models in ROIs. The proportion of voxels with the SpinalCompCor model that are now active (blue), are no longer active (red), or are still active (gray) are shown. (V=ventral, D=dorsal, R=right, L=left).*

## 4 DISCUSSION

### 4.1 Data-driven denoising in spinal cord fMRI

In this study, we formalize a framework for performing PCA of noise ROI timeseries in spinal cord fMRI data to derive nuisance regressors, using a parallel analysis to automatically determine the appropriate number of component-based nuisance regressors to avoid overfitting. Our work builds on prior research methods developed to denoise both brain and spinal cord fMRI timeseries. In these prior studies, PCA-based nuisance regression in spinal cord fMRI has been implemented but not specifically evaluated, and we lack a standard approach for defining the number of nuisance regressors to be incorporated into the timeseries model. Thus, the current work specifically assesses PCA-based denoising in the spinal cord (i.e., SpinalCompCor), providing a standard, automated approach for deciding the number of PCA-based nuisance regressors to include, and considers the impact of such denoising on the mapping of task activation.

To begin, there are key methodological similarities and differences between the implementation of PCA-based denoising of spinal cord fMRI data in past literature and the SpinalCompCor implementation in the present work that should be discussed. First, Barry et al. (2014) and Barry et al. (2016) define a noise ROI by logically inverting a mask drawn around the entire spinal cord and CSF region. Based on these studies, the Neptune toolbox for spinal cord fMRI analysis contains methods for 'non-cord' PCA-based noise regression (Deshpande & Barry, 2022). Similarly, another study derives regressors with PCA from CSF and 'not-spine' ROIs and then regresses them from the resting state fMRI timeseries along with other noise regressors (RETROICOR and motion parameters) (Combes et al., 2023). One other resting state fMRI study also includes PCA-derived CSF regressors in their noise model (Landelle et al., 2023). Here, we opt to limit PCA to a noise ROI that does not include the spinal canal CSF, instead addressing pulsatile CSF artifacts with a dedicated regressor, acknowledging that whether to include spinal canal CSF in the noise ROI is a nuanced choice that could be addressed in future work.

Furthermore, in the Barry et al. 2016 method, components were selected for each slice based on their associated variance. Specifically, the components retained represented up to 80% of the cumulative variance, or where the difference in variance explained between successive components was less than 5% (Barry et al., 2016), resulting in a different number of components for each slice. In the current work, principal components are retained based on a parallel analysis, which compares template (real data) eigenvalues to surrogate (simulated data) eigenvalues to determine a cutoff number of components for each slice, typically around 11 components in the data considered here (see next section for further details) **(Fig. 2)**. Different thresholds may be needed for different fMRI datasets, and the parallel analysis automatically accounts for this variability by finding the number of components that explain greater variance than the surrogates do by chance. The median across slices is then selected as the number of components to retain as SpinalCompCor regressors. This final step avoids a difference across slices in the number of regressors and model degrees of freedom, simplifying the analysis but potentially over or underestimating the number of regressors that "should" be used for each slice.

Finally, the Barry et al. method regresses the PCA-derived nuisance regressors from the fMRI timeseries prior to motion correction to improve the efficacy of volume registration. In the current work, however, we incorporated

415 the SpinalCompCor regressors into the subject-level GLM, opting instead to minimize the number of regression
416 steps in the analysis pipeline. Although this eliminates the potential benefits to motion correction described in
417 the Barry et al. method, this choice should simplify preprocessing and avoids the possible reintroduction of
418 artifacts back into the data that can occur with sequential regression steps (Lindquist et al., 2019).

419 **4.2 Selecting principal components as nuisance regressors**

420 A parallel analysis with surrogate timeseries data was used to systematically determine how many principal
421 components to retain, to minimize overfitting BOLD signal variance. This analysis outputs a cutoff number of
422 principal components for each slice of each run, from which the median across slices was chosen. The median
423 number of principal components ranged from 6–17 across each of the three datasets evaluated here, save for a
424 couple of large outliers in the resting state dataset **(Fig. 2A)**. When the slice-wise parallel analysis output was
425 considered, particularly in the hand-grasp task data, higher variability across subjects was observed in the
426 inferior slices, which may be associated with lower signal quality in these regions (Hemmerling et al., 2023a), or
427 different noise characteristics for this task. This variability is averaged out when selecting the median number of
428 principal components across slices for each fMRI dataset.

429 The parallel analysis is computationally intensive, taking us well over an hour just to define the number of
430 regressors to include. (Note, this analysis was done on a high-performance computing cluster and may not
431 reflect timing for all computer systems/datasets.) However, the consistency of the cutoff number of principal
432 components across the three datasets also implies that an appropriate number of components for a study could
433 be reasonably extrapolated from a parallel analysis performed on a randomly selected subset of subjects, rather
434 than on every run in the dataset, thus reducing the overall computational burden. Additionally, there is variability
435 in the cutoff number of principal components across individual datasets, with low repeatability between Scan 1
436 and Scan 2 for the hand-grasping task dataset (intraclass correlation (ICC)=0.316) **(Fig. S7)**. However, the median
437 across subjects is constant, so performing the parallel analysis on a randomly selected subset of subjects may
438 still be a reasonable option to reduce computational intensity. It is worth pointing out here that the number of
439 components indicated by the parallel analysis appears to be affected by the number of timepoints (see **Fig. S2**)
440 and may relate to the overall motion present in the data (see **Fig. S3**), so the results presented in **Fig. 2** should not
441 be extrapolated for other datasets.

442 The two resting state scans that had a large cutoff number of principal components also had some of the
443 lowest motion in the resting state dataset. This may initially raise concerns, as we are suggesting that
444 SpinalCompCor should include the greatest number of nuisance regressors when there is the smallest amount of
445 anticipated motion in the data. However, in this case the noise may be arising from more diverse sources with
446 more spatially varying temporal signatures, spreading it across a greater number of components. If a subject has
447 a large movement artifact (i.e., bulk motion affecting the entire slice or volume), it can supersede these more
448 subtle and local physiological fluctuations across the entire field of view. Indeed, the positive relationship
449 between the first eigenvalue and motion supports this explanation, whereby data with larger motion-related
450 artifacts have more variance explained by the first component.

451 **4.3 Modeling with PCA-derived SpinalCompCor regressors**

452 In this work, we made the intentional choice to add the SpinalCompCor noise regressors to our existing
453 denoising model that already contains motion and physiological noise regressors, rather than directly comparing
454 the two as separate models. This allows us to investigate if adding these extra data-driven regressors is useful
455 *beyond* the denoising we are already doing as is standard in the field (**Table S1**; (Braaß et al., 2023; Dabbagh et
456 al., 2024; Kinany et al., 2023; Kowalczyk et al., 2024; Landelle et al., 2023; Seifert et al., 2024)). These models

typically include RETROICOR, motion, and CSF regressors, albeit with variation in the number and the specific methodology to calculate the regressors. Future work evaluating a wider variety of models may reduce the overall number of nuisance regressors, thus maximizing the degrees of freedom in the GLM. However, typical spinal cord fMRI datasets are long, with several hundred timepoints, suggesting that the addition of 6–17 SpinalCompCor regressors may not be overly problematic.

#### 4.3.1 Spatial variation of SpinalCompCor regressor parameter estimates

Before evaluating model performance, we probed what the origin of the SpinalCompCor regressors might be. The SpinalCompCor method relies on the assumption that structured noise sources outside the spinal cord impart or concur with time-varying noise within spinal cord tissue, such that this variance within the cord can be explained by regressors derived from the noise ROI outside the cord. The anatomical structures around the cervical spinal cord are primarily CSF in the subarachnoid space, blood vessels, the trachea, vertebrae, and neck muscles (Brooks et al., 2008). Noise would primarily be expected to be imparted from cardiac (e.g., CSF, vessels) and respiratory (e.g., trachea) sources (Eippert et al., 2017), with quasi-periodic time-varying cycles, or from motion.

The parameter estimate maps associated with the SpinalCompCor regressors tell us where these regressors are fitting noise *outside* of the spinal cord region, alluding to what noise sources the components might be representing. Parameter estimate maps from fMRI runs for which the GLM was rerun for the entire field of view, show interesting spatial information about the principal components. (As there are many principal component parameter estimate maps, we selected illustrative examples of the anatomical structure in the maps but also acknowledge that these examples are not fully representative.)

Greater parameter estimates are frequently observed in CSF and blood vessels – structures affected by the cardiac cycle **(e.g., Fig. 3A,B)**. As noted previously, there is already a separate slice-wise CSF regressor in the model, which is the average timeseries of the CSF voxels with the top 20% variance (Brooks et al., 2008; Kong et al., 2012). This established approach for extracting a nuisance regressor from CSF voxels may mask pulsatile effects that vary (i.e., “travel”) along the length of the cord (Hemmerling et al., 2022), which SpinalCompCor regressors can potentially address. Another area showing greater parameter estimates is in the neck muscles **(e.g., Fig. 3D)**. This may be related to respiration, as Brooks et al. (2008) observed respiratory effects in these tissues, but may also reflect any number of other physiological fluctuations occurring during scanning, such as swallowing, which we are unable to specifically observe in these data. Elevated parameter estimates at the edges of structures across the field of view may represent cases for which the SpinalCompCor regressor is explaining bulk motion **(e.g., Fig. S4D)**. Most maps appear to have some structure; however, a more unstructured example is seen in **Fig. 3C**.

A wider FOV in the resting state data commonly shows higher parameter estimates in voxels that might be influenced by swallowing or contain air-filled respiratory structures **(Fig. S6)**. While the noise ROI does not actually cover this more ventral part of the image, the parameter estimate maps suggest that signal fluctuation in this region may be influential enough to affect voxels in the nearby noise ROI. This highlights the utility of acquisitions that limit the impact of noise confounds from outside of the imaged region, such as a reduced FOV acquisition with outer volume suppression (OVS) or inner FOV imaging with selective excitation (Kinany et al., 2022). The latter of these was used for the breath-hold and hand-grasp task datasets in this study (specifically, ZOOMit from Siemens). Kinany et al. (2022) specifically compares OVS and ZOOMit for detection of task and resting state fMRI activity, demonstrating strengths and weaknesses of the approaches and their ability to image the spinal cord.

#### 4.3.2 Nested model evaluation

Next, we aimed to test whether SpinalCompCor regressors, reflecting structured noise from outside of the spinal cord, explained significant signal variance *inside* of the spinal cord. We performed a voxel-wise *F*-test to evaluate whether the SpinalCompCor (i.e., full) model fit the hand-grasp task fMRI data significantly better than the Base (i.e., reduced) model. All spinal cord segments do not benefit similarly, as supported by the mixed modeling approach looking at the proportion of statistically significant voxels (*F*-test) across spinal cord segments **(Fig. 4)**. It is possible there are differences in the noise profile rostrocaudally along the cord contributing to the differential benefit from modeling with SpinalCompCor regressors. It would be interesting to determine whether this variability in SpinalCompCor benefit would persist for fMRI data in other regions such as the lumbar cord.

Most breath-hold and hand-grasp fMRI runs saw SpinalCompCor explain significantly more variance in 10–60% of voxels in both white and gray matter regions (based on the *F*-test). However, there is substantial variability in this percentage, suggesting SpinalCompCor may have a variable impact on modeling across subjects and runs. Extracting these tissue-specific metrics requires warping individual functional datasets to the PAM50 template, so these percentage metrics will inevitably capture some partial volume and misregistration effects that may contribute to this variability. Although the masks were visually checked during preprocessing, the imprecision in tissue type delineation at this field strength and spatial resolution is important to recognize. Still, these results indicate that the SpinalCompCor regressors, derived from a noise ROI outside of the spinal cord, do indeed explain significant additional variance within the spinal cord tissues, beyond variance explained by a standard set of motion and physiological noise regressors in our fMRI data.

Notably, most resting state fMRI runs saw SpinalCompCor explain significant additional variance in 90-100% of voxels across regions. This suggests that the SpinalCompCor method may provide an appreciably greater and more consistent benefit to resting state fMRI data as compared to task fMRI data. Resting state spinal cord fMRI is also primarily where methods similar to SpinalCompCor (i.e., PCA-based methods for deriving nuisance regressors) have already been applied (Barry et al., 2014, 2016; Combes et al., 2023; Landelle et al., 2023).

#### 4.3.3 Effect on downstream group-level activation results

We evaluated the utility of SpinalCompCor in improving sensitivity and specificity to motor neural activity using the hand-grasping task fMRI dataset because it has the simplest *a priori* hypothesis of the three datasets. Hand-grasping motor activation is primarily expected to localize (within the imaging field of view) to gray matter voxels in the C7-C8 spinal cord segments, ipsilateral to the side of the task (Kuypers, 1982; Leinberry & Wehbé, 2004). (See Hemmerling et al. (2023a) for a more detailed discussion of expected activity localization.) The task was modeled with hand-grasping force trace regressors. The Base model included a base set of nuisance regressors (see section 2.2.4) and the SpinalCompCor model additionally included the number of component regressors indicated by the parallel analysis for the full 300 timepoint run, ranging from 7–19 across the 52 datasets studied (two runs/subject).

Group-level modeling of task activation with the Base and SpinalCompCor models produced mixed results. For our lateralized hand-grasping task, we would primarily expect ipsilateral gray matter activation associated with motor activity (i.e., ventral horns) as well as some proprioceptive or sensory activity related to the task (i.e., dorsal horns). We would expect less activation in white matter voxels, given functional neurovascular anatomy and BOLD sensitivity at 3 T. Simplistically, when comparing the Base model to the SpinalCompCor model, a

reduction in statistically significant activation in white matter voxels might be expected, whilst we might anticipate gray matter activation to stay the same or increase.

In white matter voxels, the proportion of active voxel changes to now active or no longer active is inconsistent between left and right grasping, suggesting that the benefit indicated at the individual dataset level **(Fig. 4)** does not translate to a consistent change in (potentially spurious) task activation in white matter regions. However, we note that white matter activation may not be entirely spurious; there is increasing evidence of meaningful neural BOLD signals in white matter in the brain (Gore et al., 2019), as well as in recent work in the primate spinal cord (Sengupta et al., 2023). Therefore, we cannot discount the white matter BOLD signal as containing true activation.

For right hand-grasping, in ventral gray matter in the C8 spinal cord segment– part of the hypothesized region in which peak motor activity would be expected (Vanderah & Gould, 2016) – the proportion of voxels that are now active vs. no longer active with the SpinalCompCor model is larger, suggesting our method improved sensitivity to true neural activation in the BOLD timeseries. However, in the corresponding ROI for left hand-grasping, the proportion of voxels that are now active with the SpinalCompCor model is smaller compared to those that are no longer active. In fact, in most ventral gray matter ROIs, a higher proportion of voxels are no longer active with the SpinalCompCor model, while the opposite is seen for most dorsal gray matter ROIs. Without having data on hand dominance or a graded force-level motor task, we can speculate that differences in hand dominance and the absolute force being generated by each hand may lead to differences in absolute neural activity as well as correlated noise artifacts in left versus right hand-grasping. These differences may also cause variation in the sensory activity associated with the motor task, impacting our dorsal horn results. Regardless of the reasons, our observations suggest that SpinalCompCor may be impacting task activation mapping in an unpredictable way. Moreover, there may be collinearity between SpinalCompCor regressors and our task regressors that must be addressed to ensure SpinalCompCor is appropriate in task fMRI datasets.

#### 4.3.4 Collinearity

The inconsistent task activation mapping results led us to consider whether there could be any issues with regressor collinearity interfering with our statistical inference (e.g., cross-talk between the SpinalCompCor nuisance regressors and our task regressors of interest). In particular, when including additional regressors, it would be cause for concern if there is a strong correlation between one of these and the task regressor(s). A correlation matrix of model regressors for an example hand-grasp task model (Subject 13, Scan 1, Slice 12) shows the collinearity of each regressor with the others **(Fig. 6A,B)**. In this example, high correlations are observed between nuisance regressors with other nuisance regressors, which indicates some redundancy within our model. In fact, this provides additional insight into what motion or physiological confounds SpinalCompCor is capturing. For example, PC1 and PC2 correlate strongly with CSF and cardiac RETROICOR regressors, respectively. Thus, in the absence of external physiological recordings, using data-driven PCA regressors like these may be particularly beneficial.

Interestingly, there are some non-zero correlations indicated between principal component regressors when they should be orthogonal to each other (lower right quadrant of **Fig. 6B**). This is due to FSL's modeling (e.g., FEAT was set up to apply a high-pass filter, which is applied to both the fMRI timeseries and the regressors). In other words, the principal components initially had zero correlation, but the regressors that were actually used in the model, and evaluated here, regained some small degree of correlation following temporal filtering.

Because our chief interest is whether the principal component regressors would share variance with the task regressors, and thus negatively impact the parameter estimates/inferences for task activation, linear models of the task vs. principal component regressors (Task ~ PC1 + PC2 + ... + PCn) were computed and the associated

adjusted $R^2$ was plotted for every slice of each fMRI run **(Fig. 6C)**. Most adjusted $R^2$ values are low, so we are not particularly concerned with introducing model collinearity when adding principal component regressors. However, there some outliers with higher correlations which could contribute to the ambiguous group-level activation results.

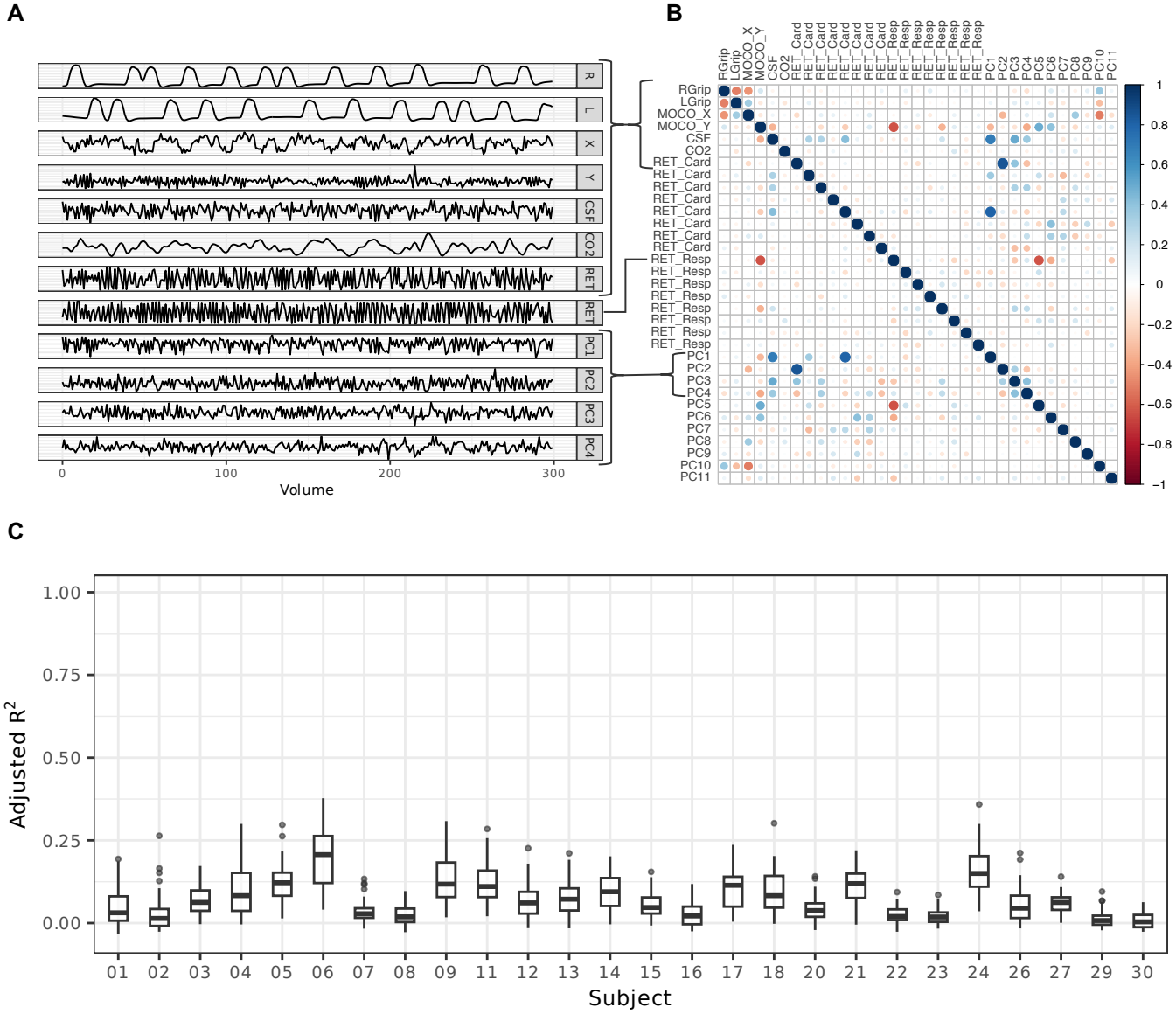


***Fig. 6. Collinearity of fMRI models for an example run and across all hand-grasp data. (A)*** *Select regressors for example model (Subject 13, Scan 1, Slice 12).* ***(B)*** *Example correlation matrix for the same model.* ***(C)*** *Adjusted $R^2$ for the model of Task ~ PC1 + PC2 + … + PCn across all slices, subjects, and scans.*

### 4.4 Limitations and future directions

Spinal cord fMRI data analysis is actively being improved upon by researchers in the community. As such, there were some small changes to data processing and analysis that we have made since the original work using the hand-grasp task dataset (Hemmerling et al., 2023a). While the original work did not observe many significant

596 voxels at the C8 spinal cord segment, we have slightly improved the co-registration and are using the updated
597 Spinal Cord Toolbox version (6.1.0) of the spinal cord segment template. Thus, increased activation in the C8
598 segment is observed here.

599 A main limitation of this work is that we selected and tested a specific analysis pipeline and specific
600 collection of nuisance regressors in our Base model. While the basic fMRI processing steps used are standard,
601 they may be implemented differently by different research groups, depending both on software choice and
602 individuals' use of the tools. The selection of regressors to include in the GLM for the Base model is based on
603 published literature **(Table S1)**, the availability of physiological data recordings, and what has been successful in
604 our own past work. For example, a $P_{ET}CO_2$ regressor is included in the GLM Base model. Recordings of expired
605 gases (i.e., $CO_2$ and $O_2$) are typical for our lab to collect during fMRI experiments because we are particularly
606 interested in studying vascular physiology, but these recordings are not standard for all groups.

607 This work incorporated three 3 T fMRI task/resting state datasets collected on two scanners from different
608 vendors. This variation should increase the applicability of the results of this work. However, the full gamut of
609 possible vendors, scan parameters, tasks, et cetera, cannot be represented by these datasets alone. There is a
610 growing initiative for sharing spinal cord fMRI data across the international research community (Banerjee et al.,
611 2025) and such resources will be invaluable for further standardizing methods in the spinal cord fMRI field, just as
612 big data initiatives in brain fMRI like the Human Connectome Project have done in the last two decades (Elam et
613 al., 2021).

614 Each of these choices affects the results and interpretation of the efficacy of using PCA-derived
615 SpinalCompCor regressors in spinal cord fMRI modeling. Ultimately, we observe that SpinalCompCor can be
616 applied in a systematic, principled manner, and explains significant variance within the spinal cord parenchyma
617 in a variety of datasets. However, the impact of such denoising on task activation was ambiguous and potentially
618 related to task correlated variance in the SpinalCompCor regressors. The nuisance regression aims to remove
619 motion and physiological effects, however, regressors exhibiting task-correlated fluctuations could
620 unintentionally reduce the share of variance that task regressors receive, hindering the detection of activation.
621 Further development of this method could take inspiration from literature in the following ways. In the CompCor
622 method, Behzadi et al. (2007) described a technique to remove stimulus-related components: voxels are
623 excluded from the noise ROI that are too correlated with the stimulus-related reference function. A similar
624 technique could be tested for spinal cord noise ROIs. Alternatively, PCA-derived SpinalCompCor regressors
625 could be orthogonalized with respect to other regressors in the model. However, this should be implemented
626 with caution, and it is advisable to fully understand warnings related to orthogonalization in fMRI modeling (see
627 Mumford et al. (2015)).

628 A different technique that might help to address task-correlated noise is the use of multi-echo independent
629 component analysis (Kundu et al., 2012). Previous work has shown this method to be beneficial in the presence
630 of high task-correlated head motion at the subject-level (Reddy et al., 2024). However, group-level results were
631 similar across levels of head motion (Reddy et al., 2024), indicating that group-level activation mapping may
632 already be fairly robust to variation in noise across subjects. This is similar to what was observed here in the
633 hand-grasp task fMRI data – the SpinalCompCor model explained noise at the subject-level, but a corresponding
634 change in the direction of hypothesized motor activation for group-level maps was not seen. Therefore, advanced
635 denoising methods like SpinalCompCor may be particularly beneficial for precision mapping of the individual
636 spinal cord.

637

## 5 CONCLUSIONS

In this work, we presented a method to derive PCA-based nuisance regressors from a noise ROI for spinal cord fMRI modeling. Components were designated as regressors for denoising using a parallel analysis comparing real and simulated noise ROI timeseries. This automated and principled approach suggested a reasonable number of components to retain as regressors across 3 diverse datasets. Across datasets, SpinalCompCor regressors reflected variance in areas of the fMRI volume *outside* the spinal cord that correspond to anatomical features expected to exhibit physiological fluctuations and systemically confound fMRI timeseries (e.g., blood vessels, CSF). Accordingly, these regressors explained significant signal variance *inside* the spinal cord in subject-level modeling. Resting state fMRI runs showed greater benefit from SpinalCompCor modeling, with a higher proportion of voxels having significant variance explained by the added nuisance regressors than in the two task fMRI datasets. Mixed results were observed, however, for group-level task activation mapping, and this may suggest challenges with task-correlated noise and model collinearity. Overall, we demonstrate benefits of using PCA-derived SpinalCompCor regressors in modeling spinal cord fMRI timeseries. However, we show that these benefits may not be equal across fMRI datasets, and suggest this denoising approach may be most useful in resting state fMRI applications.

**Table S1. Inclusion of regressors for noise modeling in recent human spinal cord fMRI studies.** Motion includes motion parameters from motion correction and/or motion outliers. CSF regressors were typically calculated as the top percent variance in a CSF mask, as the mean signal, or using PCA methods. Published studies in the past two years (from November 2, 2024) which detailed their noise and fMRI modeling methods were included. (RETROICOR=retrospective image correction (Glover et al., 2000), PNM=physiological noise modelling (Brooks et al., 2008), CSF=cerebrospinal fluid, RVT=respiration volume per time, HRV=heart rate variability).

| | | | Included Regressors | | | |
|---|---|---|---|---|---|---|
| **Authors (Year), Journal** | **Title** | **Task/Rest** | **RETROICOR / PNM** | **Motion** | **CSF** | **Other** |
| Dabbagh et al. (2024), *Imaging Neuroscience* | Reliability of task-based fMRI in the dorsal horn of the human spinal cord | Thermal Stimulus | X | X | X | |
| Seifert et al. (2024), *Hum Brain Mapp* | Thermal stimulus task fMRI in the cervical spinal cord at 7 Tesla | Thermal Stimulus | X | X | X | |
| Kowalczyk et al. (2024), *Hum Brain Mapp* | Spinal fMRI demonstrates segmental organisation of functionally connected networks in the cervical spinal cord: A test-retest reliability study | Resting state | X | X | X | |
| Braaß et al. (2023), *Hum Brain Mapp* | Association between activity in the ventral premotor cortex and spinal cord activation during force generation-A combined cortico-spinal fMRI study | Motor Task | X | X | X | RVT<br>HRV |
| Combes et al. (2023), *Sci Rep* | Detection of resting-state functional connectivity in the lumbar spinal cord with 3T MRI | Resting State | X | X | X (PCA) | Not spine ROI (PCA) |
| Hemmerling et al. (2023), *Hum Brain Mapp* | Spatial distribution of hand-grasp motor task activity in spinal cord functional magnetic resonance imaging | Motor Task | X | X | X | $P_{ET}CO_2$<br>SpinalCompCor |
| Kinany et al. (2023), *NeuroImage* | Decoding cerebro-spinal signatures of human behavior: Application to motor sequence learning | Motor Task | X | X | X | Mean global signal outside the spinal cord |
| Landelle et al. (2023), *Mov Disord* | Altered Spinal Cord Functional Connectivity Associated with Parkinson's Disease Progression | Resting State | X | X | X (PCA) | RVT<br>HRV |

**Data and Code Availability:** Code to create regressors and perform the parallel analysis is available at https://github.com/BrightLab-ANVIL/spinalcompcor. Datasets from three sources were used and are available according to their original publication data availability. The hand-grasping task fMRI data are publicly available on OpenNeuro at doi:10.18112/openneuro.ds004616.v1.1.1 (Hemmerling et al., 2023b). The breath-holding task fMRI data are publicly available on OpenNeuro at doi:10.18112/openneuro.ds004974.v1.0.0 (Hemmerling et al., 2024b). (*Note: The second dataset has not yet been published but will be by the time of publication of this article.*) The resting state fMRI data are available upon reasonable request (see Barry et al. (2018)).

**Author Contributions:** K.J.H: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration. A.D.V.: Methodology, Writing - Review & Editing. C.G.: Software, Formal analysis, Writing - Review & Editing. R.L.B.: Conceptualization, Methodology, Investigation, Data curation, Writing - Review & Editing, Funding acquisition. M.G.B.: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Project administration, Funding acquisition

**Funding:** Research supported by the Craig H. Neilsen Foundation (595499). KJH was supported by a NIH NIBIB-funded training program (T32EB025766) and an NINDS-funded predoctoral fellowship (F31NS134222). ADV was supported by an NIH NINDS-funded predoctoral fellowship (F31NS126012). Resting state spinal cord fMRI data were acquired with the support of K99EB016689 and R21NS081437, and RLB was supported by NIBIB through awards R00EB016689, R01EB027779 and R21EB031211. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

**Declaration of Competing Interests:** Since January 2024, Dr. Barry has been employed by the National Institute of Biomedical Imaging and Bioengineering at the NIH. This article was co-authored by Robert Barry in his personal capacity. The opinions expressed in the article are his own and do not necessarily reflect the views of the NIH, the Department of Health and Human Services, or the United States government. The other authors declare no competing interests.

**Acknowledgments:** This work was supported by the Center for Translational Imaging at Northwestern University. This research was also supported in part through the computational resources and staff contributions provided for the Quest high performance computing facility at Northwestern University which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology.

## References


Banerjee, R., Kaptan, M., Tinnermann, A., Khatibi, A., Dabbagh, A., Kündig, C. W., Law, C. S. W., Pfyffer, D., Lythgoe, D. J., Tsivaka, D., Van De Ville, D., Eippert, F., Muhammad, F., Glover, G. H., David, G., Haynes, G., Haaker, J., Brooks, J. C. W., Finsterbusch, J., … Cohen-Adad, J. (2025). EPISeg: Automated segmentation of the spinal cord on echo planar images using open-access multi-center data. *BioRxiv*. https://doi.org/10.1101/2025.01.07.631402

Barry, R. L., Conrad, B. N., Smith, S. A., & Gore, J. C. (2018). A practical protocol for measurements of spinal cord functional connectivity. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-34841-6

Barry, R. L., Rogers, B. P., Conrad, B. N., Smith, S. A., & Gore, J. C. (2016). Reproducibility of resting state spinal cord networks in healthy volunteers at 7 Tesla. *NeuroImage*, *133*, 31–40. https://doi.org/10.1016/j.neuroimage.2016.02.058

Barry, R. L., Smith, S. A., Dula, A. N., & Gore, J. C. (2014). Resting state functional connectivity in the human spinal cord. *ELife*, *2014*(3). https://doi.org/10.7554/eLife.02812

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method ( CompCor ) for BOLD and perfusion based fMRI. *Human Brain Mapping Journal*, *37*(1), 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

Braaß, H., Feldheim, J., Chu, Y., Tinnermann, A., Finsterbusch, J., Büchel, C., Schulz, R., & Gerloff, C. (2023). Association between activity in the ventral premotor cortex and spinal cord activation during force generation—A combined cortico-spinal fMRI study. *Human Brain Mapping*, *44*(18). https://doi.org/10.1002/hbm.26523

Brooks, J. C. W., 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*(2), 680–692. https://doi.org/10.1016/j.neuroimage.2007.09.018

Combes, A., Narisetti, L., Sengupta, A., Rogers, B. P., Sweeney, G., Prock, L., Houston, D., McKnight, C. D., Gore, J. C., Smith, S. A., & O'Grady, K. P. (2023). Detection of resting-state functional connectivity in the lumbar spinal cord with 3T MRI. *Scientific Reports*, *13*(1), 18189. https://doi.org/10.1038/s41598-023-45302-0

Dabbagh, A., Horn, U., Kaptan, M., Mildner, T., Müller, R., Lepsien, J., Weiskopf, N., Brooks, J. C. W., Finsterbusch, J., & Eippert, F. (2024). Reliability of task-based fMRI in the dorsal horn of the human spinal cord. *Imaging Neuroscience*, *2*, 1–27. https://doi.org/10.1162/IMAG_A_00273

De Leener, B., Fonov, V. S., Collins, D. L., Callot, V., Stikov, N., & Cohen-Adad, J. (2018). PAM50: Unbiased multimodal template of the brainstem and spinal cord aligned with the ICBM152 space. *NeuroImage*, *165*, 170–179. https://doi.org/10.1016/j.neuroimage.2017.10.041

Deshpande, R., & Barry, R. (2022). "Neptune: a toolbox for spinal cord functional MRI data processing and quality assurance." *Proceedings 30th Scientific Meeting, International Society for Magnetic Resonance in Medicine*.

Eippert, F., Kong, Y., Jenkinson, M., Tracey, I., & Brooks, J. C. W. (2017). Denoising spinal cord fMRI data: Approaches to acquisition and analysis. *NeuroImage*, *154*, 255–266. https://doi.org/10.1016/j.neuroimage.2016.09.065

Elam, J. S., Glasser, M. F., Harms, M. P., Sotiropoulos, S. N., Andersson, J. L. R., Burgess, G. C., Curtiss, S. W., Oostenveld, R., Larson-Prior, L. J., Schoffelen, J. M., Hodge, M. R., Cler, E. A., Marcus, D. M., Barch, D. M., Yacoub, E., Smith, S. M., Ugurbil, K., & Van Essen, D. C. (2021). The Human Connectome Project: A retrospective. *NeuroImage*, *244*. https://doi.org/10.1016/j.neuroimage.2021.118543

Glover, G. H., Li, T. Q., & Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magnetic Resonance in Medicine*, *44*(1), 162–167. https://doi.org/10.1002/1522-2594(200007)44:1<162::AID-MRM23>3.0.CO;2-E

Gore, J. C., Li, M., Gao, Y., Wu, T. L., Schilling, K. G., Huang, Y., Mishra, A., Newton, A. T., Rogers, B. P., Chen, L. M., Anderson, A. W., & Ding, Z. (2019). Functional MRI and resting state connectivity in white matter - a mini-review. *Magnetic Resonance Imaging*, *63*. https://doi.org/10.1016/j.mri.2019.07.017

Gros, C., De Leener, B., Badji, A., Maranzano, J., Eden, D., Dupont, S. M., Talbott, J., Zhuoquiong, R., Liu, Y., Granberg, T., Ouellette, R., Tachibana, Y., Hori, M., Kamiya, K., Chougar, L., Stawiarz, L., Hillert, J., Bannier, E., Kerbrat, A., … Cohen-Adad, J. (2019). Automatic segmentation of the spinal cord and intramedullary

multiple sclerosis lesions with convolutional neural networks. *NeuroImage*, *184*, 901–915. https://doi.org/10.1016/J.NEUROIMAGE.2018.09.081

Hemmerling, K. J., Hoggarth, M. A., Parrish, T., & Bright, M. G. (2022). Spinal cord fMRI heatmaps reveal a structured cardiac artifact "traveling" along the length of the cord. *Proceedings 30th Scientific Meeting, International Society for Magnetic Resonance in Medicine*.

Hemmerling, K. J., Hoggarth, M. A., Sandhu, M. S., Parrish, T. B., & Bright, M. G. (2023a). Spatial distribution of hand-grasp motor task activity in spinal cord functional magnetic resonance imaging. *Human Brain Mapping*, *44*(17), 5567–5581. https://doi.org/10.1002/hbm.26458

Hemmerling, K. J., Hoggarth, M. A., Sandhu, M. S., Parrish, T. B., & Bright, M. G. (2023b). Spatial distribution of hand-grasp motor task activity in spinal cord functional magnetic resonance imaging. In *OpenNeuro*.

Hemmerling, K. J., Hoggarth, M. A., Sandhu, M. S., Parrish, T. B., & Bright, M. G. (2024a). MRI mapping of hemodynamics in the human spinal cord. *BioRxiv*, 2024.02.22.581606. https://doi.org/10.1101/2024.02.22.581606

Hemmerling, K. J., Hoggarth, M. A., Sandhu, M. S., Parrish, T. B., & Bright, M. G. (2024b). MRI mapping of hemodynamics in the human spinal cord. In *OpenNeuro*.

Hu, Y., Jin, R., Li, G., Luk, K. D. K., & Wu, E. X. (2018). Robust spinal cord resting-state fMRI using independent component analysis-based nuisance regression noise reduction. *Journal of Magnetic Resonance Imaging*, *48*(5), 1421–1431. https://doi.org/10.1002/jmri.26048

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Kinany, N., Khatibi, A., Lungu, O., Finsterbusch, J., Büchel, C., Marchand-Pauvert, V., Van De Ville, D., Vahdat, S., & Doyon, J. (2023). Decoding cerebro-spinal signatures of human behavior: Application to motor sequence learning. *NeuroImage*, *275*. https://doi.org/10.1016/j.neuroimage.2023.120174

Kinany, N., Pirondini, E., Mattera, L., Martuzzi, R., Micera, S., & Van De Ville, D. (2022). Towards reliable spinal cord fMRI: Assessment of common imaging protocols. *NeuroImage*, *250*. https://doi.org/10.1016/j.neuroimage.2022.118964

Kong, Y., Jenkinson, M., Andersson, J., Tracey, I., & Brooks, J. C. W. (2012). Assessment of physiological noise modelling methods for functional imaging of the spinal cord. *NeuroImage*, *60*(2), 1538–1549. https://doi.org/10.1016/j.neuroimage.2011.11.077

Kowalczyk, O. S., Medina, S., Tsivaka, D., McMahon, S. B., Williams, S. C. R., Brooks, J. C. W., Lythgoe, D. J., & Howard, M. A. (2024). Spinal fMRI demonstrates segmental organisation of functionally connected networks in the cervical spinal cord: A test–retest reliability study. *Human Brain Mapping*, *45*(2). https://doi.org/10.1002/hbm.26600

Kundu, P., Inati, S. J., Evans, J. W., Luh, W. M., & Bandettini, P. A. (2012). Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *NeuroImage*, *60*(3). https://doi.org/10.1016/j.neuroimage.2011.12.028

Kuypers, H. G. J. M. (1982). A New Look at the Organization of the Motor System. *Progress in Brain Research*, *57*(C). https://doi.org/10.1016/S0079-6123(08)64138-2

Landelle, C., Dahlberg, L. S., Lungu, O., Misic, B., De Leener, B., & Doyon, J. (2023). Altered Spinal Cord Functional Connectivity Associated with Parkinson's Disease Progression. *Movement Disorders*, *38*(4). https://doi.org/10.1002/mds.29354

Leinberry, C. F., & Wehbé, M. A. (2004). Brachial plexus anatomy. *Hand Clinics*, *20*(1). https://doi.org/10.1016/S0749-0712(03)00088-X

Lindquist, M. A., Geuter, S., Wager, T. D., & Caffo, B. S. (2019). Modular preprocessing pipelines can reintroduce artifacts into fMRI data. *Human Brain Mapping*, *40*(8). https://doi.org/10.1002/hbm.24528

Mumford, J. A., Poline, J. B., & Poldrack, R. A. (2015). Orthogonalization of regressors in fMRI models. *PLoS ONE*, *10*(4). https://doi.org/10.1371/journal.pone.0126255

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric Permutation Tests For Functional Neuroimaging: A Primer with Examples. *Human Brain Mapping*, *15*(1), 1–25. https://doi.org/10.1002/hbm.1058

Patriat, R., Molloy, E. K., & Birn, R. M. (2015). Using Edge Voxel Information to Improve Motion Regression for rs-fMRI Connectivity Studies. *Brain Connectivity*, *5*(9), 582. https://doi.org/10.1089/BRAIN.2014.0321

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. *Magnetic Resonance Imaging*, *27*(3), 300–310. https://doi.org/10.1016/j.mri.2008.07.019

Raj, D., Anderson, A. W., & Gore, J. C. (2001). Respiratory effects in human functional magnetic resonance imaging due to bulk susceptibility changes. *Physics in Medicine and Biology*, *46*(12), 3331–3340. https://doi.org/10.1088/0031-9155/46/12/318

Reddy, N. A., Zvolanek, K. M., Moia, S., Caballero-Gaudes, C., & Bright, M. G. (2024). Denoising task-correlated head motion from motor-task fMRI data with multi-echo ICA. *Imaging Neuroscience*, *2*. https://doi.org/10.1162/imag_a_00057

Rieseberg, S., Frahm, J., & Finsterbusch, J. (2002). Two-dimensional spatially-selective RF excitation pulses in echo-planar imaging. *Magnetic Resonance in Medicine*, *47*(6), 1186–1193. https://doi.org/10.1002/MRM.10157

Seifert, A. C., Xu, J., Kong, Y., Eippert, F., Miller, K. L., Tracey, I., & Vannesjo, S. J. (2024). Thermal stimulus task fMRI in the cervical spinal cord at 7 Tesla. *Human Brain Mapping*, *45*(3). https://doi.org/10.1002/hbm.26597

Sengupta, A., Mishra, A., Wang, F., Chen, L., & Gore, J. (2023). Identification of synchronous BOLD signal patterns in white matter of primate spinal cord. *Research Square*. https://doi.org/10.21203/rs.3.rs-2389151/v1

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*(1), 83–98. https://doi.org/10.1016/J.NEUROIMAGE.2008.03.061

Vanderah, T. W., & Gould, D. J. (2016). Spinal Cord. In *Nolte's the Human Brain* (7th ed., pp. 233–271). Elsevier Inc.

Venema, V., Meyer, S., García, S. G., Kniffka, A., Simmer, C., Crewell, S., Löhnert, U., Trautmann, T., & Macke, A. (2006). Surrogate cloud fields generated with the iterative amplitude adapted Fourier transform algorithm. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *58*(1). https://doi.org/10.1111/j.1600-0870.2006.00160.x

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, *92*(100), 381–397. https://doi.org/10.1016/J.NEUROIMAGE.2014.01.060

Xie, G., Piché, M., Khoshnejad, M., Perlbarg, V., Chen, J. I., Hoge, R. D., Benali, H., Rossignol, S., Rainville, P., & Cohen-Adad, J. (2012). Reduction of physiological noise with independent component analysis improves the detection of nociceptive responses with fMRI of the human spinal cord. *NeuroImage*, *63*(1), 245–252. https://doi.org/10.1016/j.neuroimage.2012.06.057