

8-15-2021

## nnResting state fMRI scanner instabilities revealed by longitudinal phantom scans in a multi-center study

Aras Kayvanrad  
*University of Toronto*

Stephen R. Arnott  
*University of Toronto*

Nathan Churchill  
*Saint Michael's Hospital University of Toronto*

Stefanie Hassel  
*University of Calgary*

Aditi Chemparathy  
*University of Toronto*

*See next page for additional authors*

Follow this and additional works at: <https://ir.lib.uwo.ca/biophysicspub>



Part of the [Medical Biophysics Commons](#)

---

### Citation of this paper:

Aras Kayvanrad, Stephen R. Arnott, Nathan Churchill, Stefanie Hassel, Aditi Chemparathy, Fan Dong, Mojdeh Zamyadi, Tom Gee, Robert Bartha, Sandra E. Black, Jane M. Lawrence-Dewar, Christopher J.M. Scott, Sean Symons, Andrew D. Davis, Geoffrey B. Hall, Jacqueline Harris, Nancy J. Lobaugh, Glenda MacQueen, Cindy Woo, Stephen Strother, Resting state fMRI scanner instabilities revealed by longitudinal phantom scans in a multi-center study, *NeuroImage*, Volume 237, 2021, 118197, ISSN 1053-8119, <https://doi.org/10.1016/j.neuroimage.2021.118197>. (<https://www.sciencedirect.com/science/article/pii/S1053811921004742>)

---

## Authors

Aras Kayvanrad, Stephen R. Arnott, Nathan Churchill, Stefanie Hassel, Aditi Chemparathy, Fan Dong, Mojdeh Zamyadi, Tom Gee, Robert Bartha, Sandra E. Black, Jane M. Lawrence-Dewar, Christopher J.M. Scott, Sean Symons, Andrew D. Davis, Geoffrey B. Hall, Jacqueline Harris, Nancy J. Lobaugh, Glenda MacQueen, Cindy Woo, and Stephen Strother



## Resting state fMRI scanner instabilities revealed by longitudinal phantom scans in a multi-center study



Aras Kayvanrad<sup>a,\*</sup>, Stephen R. Arnott<sup>a</sup>, Nathan Churchill<sup>b</sup>, Stefanie Hassel<sup>c,d</sup>, Aditi Chemparathy<sup>a</sup>, Fan Dong<sup>e</sup>, Mojdeh Zamyadi<sup>a</sup>, Tom Gee<sup>e</sup>, Robert Bartha<sup>f,g</sup>, Sandra E. Black<sup>h</sup>, Jane M. Lawrence-Dewar<sup>i</sup>, Christopher J.M. Scott<sup>h</sup>, Sean Symons<sup>h</sup>, Andrew D. Davis<sup>a,j</sup>, Geoffrey B. Hall<sup>j</sup>, Jacqueline Harris<sup>k</sup>, Nancy J. Lobaugh<sup>l</sup>, Glenda MacQueen<sup>c,1</sup>, Cindy Woo<sup>m,1</sup>, Stephen Strother<sup>a,n</sup>, the ONDRI Founding Investigators<sup>2</sup>, the CAN-BIND Investigators<sup>3</sup>

<sup>a</sup> Rotman Research Institute, University of Toronto, Toronto, ON, Canada

<sup>b</sup> Neuroscience Research Program, St. Michael's Hospital, Toronto, ON, Canada

<sup>c</sup> The Mathison Centre for Mental Health Research and Education, University of Calgary, Calgary, AB, Canada

<sup>d</sup> Department of Psychiatry, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

<sup>e</sup> Indoc Research, Toronto, ON, Canada

<sup>f</sup> Robarts Research Institute, Western University, London, ON, Canada

<sup>g</sup> Medical Biophysics, Western University, London, ON, Canada

<sup>h</sup> Sunnybrook Research Institute, Toronto, ON, Canada

<sup>i</sup> Thunder Bay Regional Health Research Institute, Thunder Bay, ON, Canada

<sup>j</sup> Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada

<sup>k</sup> Department of Computing Science, University of Alberta, Edmonton, AB, Canada

<sup>l</sup> Center for Addiction and Mental Health, Toronto, ON, Canada

<sup>m</sup> Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada

<sup>n</sup> Medical Biophysics, University of Toronto, Toronto, ON, Canada

### ARTICLE INFO

#### Keywords:

Resting state fMRI  
fMRI quality assurance  
MRI scanner instabilities  
Multi-center/Longitudinal fMRI studies

### ABSTRACT

Quality assurance (QA) is crucial in longitudinal and/or multi-site studies, which involve the collection of data from a group of subjects over time and/or at different locations. It is important to regularly monitor the performance of the scanners over time and at different locations to detect and control for intrinsic differences (e.g., due to manufacturers) and changes in scanner performance (e.g., due to gradual component aging, software and/or hardware upgrades, etc.). As part of the Ontario Neurodegenerative Disease Research Initiative (ONDRI) and the Canadian Biomarker Integration Network in Depression (CAN-BIND), QA phantom scans were conducted approximately monthly for three to four years at 13 sites across Canada with 3T research MRI scanners. QA parameters were calculated for each scan using the functional Biomarker Imaging Research Network's (fBIRN) QA phantom and pipeline to capture between- and within-scanner variability. We also describe a QA protocol to measure the full-width-at-half-maximum (FWHM) of slice-wise point spread functions (PSF), used in conjunction with the fBIRN QA parameters. Variations in image resolution measured by the FWHM are a primary source of variance over time for many sites, as well as between sites and between manufacturers. We also identify an unexpected range of instabilities affecting individual slices in a number of scanners, which may amount to a substantial contribution of unexplained signal variance to their data. Finally, we identify a preliminary preprocessing approach to reduce this variance and/or alleviate the slice anomalies, and in a small human data set show that this change in preprocessing can have a significant impact on seed-based connectivity measurements for some individual subjects. We expect that other fMRI centres will find this approach to identifying and controlling scanner instabilities useful in similar studies.

\* Corresponding author.

*E-mail addresses:* [mkayvanr@alumni.uwo.ca](mailto:mkayvanr@alumni.uwo.ca) (A. Kayvanrad), [sarnott@research.baycrest.org](mailto:sarnott@research.baycrest.org) (S.R. Arnott), [nchurchill.research@gmail.com](mailto:nchurchill.research@gmail.com) (N. Churchill), [shassel@ucalgary.ca](mailto:shassel@ucalgary.ca) (S. Hassel), [aditi.chemparathy@gmail.com](mailto:aditi.chemparathy@gmail.com) (A. Chemparathy), [fdong@indocresearch.org](mailto:fdong@indocresearch.org) (F. Dong), [mzamyadi@research.baycrest.org](mailto:mzamyadi@research.baycrest.org) (M. Zamyadi), [tgee@indocresearch.org](mailto:tgee@indocresearch.org) (T. Gee), [rbartha@robarts.ca](mailto:rbartha@robarts.ca) (R. Bartha), [sandra.black@sunnybrook.ca](mailto:sandra.black@sunnybrook.ca) (S.E. Black), [dewarja@tbh.net](mailto:dewarja@tbh.net) (J.M. Lawrence-Dewar), [christopher.scott@sunnybrook.ca](mailto:christopher.scott@sunnybrook.ca) (C.J.M. Scott), [Sean.Symons@sunnybrook.ca](mailto:Sean.Symons@sunnybrook.ca) (S. Symons), [adavis@research.baycrest.org](mailto:adavis@research.baycrest.org) (A.D. Davis), [hallg@mcmaster.ca](mailto:hallg@mcmaster.ca) (G.B. Hall), [jkh@ualberta.ca](mailto:jkh@ualberta.ca) (J. Harris), [nancy.j.lobaugh@gmail.com](mailto:nancy.j.lobaugh@gmail.com) (N.J. Lobaugh), [ssrother@research.baycrest.org](mailto:ssrother@research.baycrest.org) (S. Strother).

<sup>1</sup> Posthumously

<https://doi.org/10.1016/j.neuroimage.2021.118197>.

Received 3 February 2020; Received in revised form 18 May 2021; Accepted 19 May 2021

Available online 21 May 2021.

1053-8119/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

**Table 1**

Description of the thirteen sites with 3T research MRI scanners participating in this study. These sites included 5 GE Discovery, 3 Siemens Trio, 3 Siemens Prisma, 1 Siemens Skyra, and 2 Philips Achieva scanners.

Site	Abbreviation	Manufacturer	Model	Coil	Installation / latest upgrade date
The center for Addiction and Mental Health (CAMH), Toronto	CAM	GE	Discovery	8-Ch Head	2011
McMaster University	MCM	GE	Discovery	32-Ch Head / Head Neck Spine (HNS)	2012
Sunnybrook Hospital, Toronto	SBH	GE	Discovery	8-Ch Head	2009
Toronto Western Hospital	TWH	GE	Signa HDxt	8-Ch Head	2004
University of Calgary	UCA	GE	Discovery	Head Neck Spine (HNS)	2011
Baycrest center, Toronto	BYC	Siemens	Trio	12-Ch Head	2006
Western University	WEU	Siemens	Prisma	20-Ch Head	2013
Queen's University	QNS	Siemens	Trio	12-Ch Head	2005
The Ottawa Hospital	TOH	Siemens	Trio	12-Ch Head	2010
University of Toronto	UTO	Siemens	Prisma	32-Ch Head	2016
St. Mike's Hospital, Toronto	SMH	Siemens	Skyra	20-Ch Head	2013
Thunder Bay Regional Health Sciences center	TBR	Philips	Achieva	8-Ch Head	2009
University of British Columbia	UBC	Philips	Achieva	8-Ch Head	2003

## 1. Introduction

Changes in the blood oxygenation-level dependent (BOLD) signal due to neuronal activity constitute only a fraction of the changes in the raw fMRI signal intensity. Other major sources of variance in the fMRI signal are thermal noise, head motion, physiological noise, and temporal instabilities in the MRI scanner hardware (also known as system noise). Consequently, to accurately measure such small changes due to neuronal activity, it is crucial to measure, understand, and where possible, control other sources of variance. Thermal noise is dominated by the SNR-resolution-acquisition-time trade off, and several studies have attempted to improve this trade-off since the early days of MRI (Griswold et al., 2002; Haase et al., 1986; Larkman and Nunes, 2007; Lustig et al., 2007; Mansfield, 1977; Margosian et al., 1986; Pruessmann et al., 1999; Samsonov et al., 2004; Sodickson and Manning, 1997). There exists a rich literature on controlling head motion and physiological noise (Caballero-Gaudes and Reynolds, 2017; Power et al., 2015; Strother, 2006). Equally important is monitoring the temporal stability of MRI scanner hardware and, where possible, making corrections to the detected instabilities. A number of quality assurance (QA) protocols have been developed to monitor the performance of the scanner (Friedman and Glover, 2006; Glover et al., 2012; Jovicich et al., 2016; Price et al., 1990; Yan et al., 2013).

Quality assurance becomes particularly important in longitudinal and/or multi-site studies, which involve collecting data from a group of subjects over time and/or at different locations. It is crucial to regularly monitor the performance of the scanners over time and at different locations to detect and, where possible, control for intrinsic differences (e.g., due to manufacturers) and changes in the scanner performance (e.g., due to gradual component aging, software and/or hardware upgrades, etc.). If such differences and changes are not accounted for, they can add unexplained variance to the data (Friedman and Glover, 2006).

As part of the Ontario Neurodegenerative Disease Research Initiative (ONDRI) (Farhan et al., 2017) and the Canadian Biomarker Integration Network in Depression (CAN-BIND) (Lam et al., 2016) QA phantom scans were conducted approximately monthly for three to four years at 13 sites across Canada with 3T research MRI scanners (See Table 1). We found considerable variance in the QA parameters over time for many sites as well as substantial variance across sites. We also identified an unexpected range of instabilities affecting individual slices in a number of scanners. These slice anomalies may amount to a substantial contribution to the signal variance and to the best of our knowledge have

not been reported before. The main objectives of this paper are as follows: (1) to assess the range of within- and between-scanner variability revealed by the fBIRN QA pipeline and parameters, (2) to identify primary factors contributing to the scanner-dependent signal variance in fMRI studies, and (3) to identify preprocessing approaches to reduce this variance.

## 2. Material and methods

### 2.1. Sites

Thirteen sites participated in this study across Canada, including MR scanners from different vendors: 5 GE Discovery, 3 Siemens Trio, 2 Siemens Prisma, 1 Siemens Skyra, and 2 Philips Achieva (Table 1).

### 2.2. Phantom scan protocol

The present study uses QA data acquired as part of the ONDRI and CAN-BIND research initiatives, where phantom QA scans were performed contemporaneously with human scan protocols of ONDRI and CAN-BIND for quality control. The functional Biomarker Imaging Research Network (fBIRN) phantom was scanned approximately monthly at each site, amounting to a total of 629 scans across all sites. The fBIRN phantom is a spherical plastic vessel 17 cm in diameter filled with a doped agar gel chosen to reflect T1, T2, magnetization transfer, and RF conductivity characteristics of human brain tissue (Friedman and Glover, 2006). Each site had their own copy of the fBIRN phantom to enable regular QA scans. While there may be minor manufacturing differences between different copies, the effect of these differences are assumed to be negligible for the present study.

The data from these 629 scans have been made open source and are available online at <https://www.braincode.ca/content/open-data-releases>.

The aim of the QA protocol is to measure scanner stability under conditions that match those of human resting state experiments (Friedman and Glover, 2006; Glover et al., 2012). Thus we employed the Canadian Dementia Imaging rs-fMRI protocol (CDIP) (<https://www.cdip-pcid.ca>) for the phantom QA scans (Duchesne et al., 2019) (Appendix Table A1). In a few study sites, the initial scan parameters did not match those of CDIP and were adjusted, during the study, to more closely match CDIP parameters (Table A2).

### 2.3. fBIRN QA parameters

The fBIRN QA pipeline ([https://www.nitrc.org/projects/bxh\\_xcedu\\_tools/](https://www.nitrc.org/projects/bxh_xcedu_tools/)) was employed to calculate QA parameters for each imaging ses-

<sup>2</sup> <https://ondri.ca/publications/>

<sup>3</sup> <https://www.canbind.ca/about-can-bind/our-team/executive-committee/>

**Table 2**  
Summary of fBIRN QA parameters.

• <b>Region of interest (ROI)</b> is by default a 15×15 square centered on the middle slice through the phantom.	
• <b>Signal image</b> is the mean intensity across time by voxel.	
• <b>Static spatial noise image</b> is the sum of all even-numbered images subtracted from the sum of all the odd images.	
• <b>Temporal fluctuation noise image:</b> first, voxel time-series is detrended with a 2nd order polynomial. The fluctuation noise image is the standard deviation (SD) of the residual by voxel.	
• <b>CV (coefficient of variation)</b> is the SD of a time-series divided by the mean of the time-series.	
<b>mean</b>	Average value in the ROI of <b>signal image</b>
<b>SNR</b>	signal-to-noise ratio Noise: variance of the <b>static spatial noise image</b> over the ROI Signal: average of the <b>signal image</b> over the ROI SNR = (signal)/√((Noise)/(length of time series))
<b>SFNR</b>	signal-fluctuation-to-noise ratio <b>Signal image</b> divided by <b>temporal fluctuation noise image</b> by voxel. Summary SFNR value is the average of this within ROI.
For the next 5 variables: a time-series composed of the mean intensity of each volume within the ROI (i.e., 15 × 15 square centered on the middle slice from each volume) is calculated (“raw signal”), and a 2nd order polynomial trend is fit to this data (“fit”).	
<b>msi</b> is mean signal intensity of the raw signal.	
<b>std</b>	SD of residuals after detrending
<b>percentFluc</b>	100*(std)/(msi)
<b>drift</b>	100*(max raw signal - min raw signal)/msi
<b>driftfit</b>	100*(max fit - min fit)/msi
<b>rdc</b>	radius of decorrelation. CV(N) = CV of the raw signal for a square NxN voxel ROI. rdc = CV(1)/CV(Nmax), where Nmax = 15.
<b>FWHM- Full width at half maximum</b>	
The fBIRN QA pipeline uses an old version of AFNI 3dFWHMx (AFNI_2011_12_21_1014) for FWHM calculation, where FWHM is calculated using the “classic” estimate as a function of the ratio of variance of first differences to data variance (Forman et al., 1995).	
<b>minFWHM</b>	min value of all volumes in each direction
<b>maxFWHM</b>	max value of all volumes in each direction
<b>MeanFWHM</b>	mean value of all volumes in each direction
AFNI_2011_12_21_1014 was used for the calculation of the next two variables:	
3dvolreg: motion correction	
3dDetrend: voxel-wise detrending using Legendre polynomials of order up to and including 2	
3dTstat: calculate mean of voxels	
3dAutomask: create mask of brain-only (high signal) voxels, dilated 4 times	
The mask is shifted by N/2 voxels in the phase-encode direction (Nyquist ghosting) to create “ghost mask”.	
<b>MeanGhost</b>	Percentage of the mean intensity of the ghost voxels to non-ghost voxels
<b>MeanBrightGhost</b>	Mean intensity of the top 10 percent of ghost-only voxels

sion (Friedman and Glover, 2006). The QA parameters used in the subsequent analyses are described in Table 2.

#### 2.4. Autocorrelation function (ACF) analysis

Subsequent analyses of the fBIRN QA parameters revealed that FWHM is a prominent factor driving the variance between imaging sessions and between different sites (as outlined in the Results section and previously shown by (Friedman et al., 2006)). Consequently, we developed an independent approach to measuring FWHM to verify and investigate the previous findings as follows:

The (spatial) autocorrelation function (ACF) of a deconstructed image can be used to estimate the FWHM measure of resolution. We calculate the ACF for each slice by first detrending the time-series with a 2nd order Legendre polynomial to subtract out the spatial structure. Detrending is performed using AFNI’s 3dDetrend (AFNI\_17.3.05). The ACF is then calculated for each slice using the Wiener-Khinchin theorem:

$$ACF(x, y) = F_{u,v}^{-1} \{ PSD(u, v) \}$$

where  $F_{u,v}^{-1}$  denotes the inverse Fourier transformation and PSD(u,v) is the power spectral density of the image, which is estimated as the squared magnitude of the image FFT.

The FWHM in each direction (i.e., x and y) is subsequently estimated by fitting a Gaussian plus exponential mixture model to the ACF profile in that direction:

$$ACF(r) = a \exp(-r^2/(2b^2)) + (1-a) \exp(-r/c)$$

The fit is identical to that used by the current version of AFNI 3dFWHM (AFNI\_17.3.05) and is used for consistency with AFNI’s current approach. Unlike the reported fBIRN FWHM measures, the slice ACF approach calculates FWHM values for each slice, which allows the identification of possible slice effects. FWHM values calculated using this approach will be referred to as “ACF FWHM” in the current paper,

to distinguish from fBIRN QA FWHM values. Any slice with  $FWHM_x > \theta \times \min FWHM_x$  or  $FWHM_y > \theta \times \min FWHM_y$  is marked as an anomaly, where  $\min FWHM_x$  and  $\min FWHM_y$  are the minimum FWHM values in the x and y directions in that scan session and  $\theta$  is an anomaly threshold.

#### 2.5. fMRI Preprocessing

##### 2.5.1. Spatial smoothing

Following the approach of (Friedman et al., 2006) to reduce the FWHM variance due to imaging sessions using different scanners, all imaging data were spatially smoothed to the largest average FWHM of all the sites set at 7 mm using AFNI’s 3dBlurToFWHM (Version AFNI\_17.3.05).

##### 2.5.2. Correction of slice anomalies - Censoring

Subsequent investigation revealed a number of anomalous slices in some imaging sessions despite additional spatial smoothing. Where anomalies occur at discontinuous time frames in the time series, the anomalies might be mitigatable by censoring the anomalous slices, i.e., by removing the anomalous slices and replacing them with temporally interpolated values. Moreover, since the anomalies occur at individual slices, a slice-wise correction approach may best suit this problem. SpikeCor (Campbell et al., 2013) is a PCA-based spike correction technique that can identify volume and / or slice anomalies. While the original version of SpikeCor detected volume anomalies only (Campbell et al., 2013), a slice version has been made available by the author where in addition to volume defects, individual slice defects can also be detected ([https://www.nitrc.org/projects/spikecor\\_fmri/](https://www.nitrc.org/projects/spikecor_fmri/)). We employed this version of SpikeCor to investigate the possibility of mitigating these slice anomalies by preprocessing. SpikeCor uses a statistical p-value threshold for outlier detection, which we set at the default value of 0.05.

## 2.6. Principal component analysis

Principal component analysis (PCA) was used to gain an overview of the within- and between-scanner normalised variance and correlation structure of the 14 fBIRN QA variables for the 13 MRI scanners over time. PCA was performed by singular value decomposition of the data matrix, where each row corresponds to one scan session and the columns correspond to the QA variables. The variables were shifted to be zero centered and scaled to have unit variance prior to the analysis. Nonetheless, no statistical inference was drawn based on PCA.

## 2.7. Statistical analysis

One-way multivariate analysis of variance (MANOVA) was used to test for significant differences in QA parameters between sites and, separately, between scanner manufacturers and the effect size was quantified using partial associations ( $\eta^2$ ). Linear mixed effects models were fit to QA parameters, allowing the intercept of the model to vary randomly between scanners (i.e., sites) or scanner manufacturers. Significance of the estimates and the random effects was quantified using likelihood ratio tests. The effect size was quantified based on the correlation between the fitted and the observed values. Intra-class correlation (ICC) was calculated by dividing the random effect variance by the total variance, i.e., the sum of the random effect variance and the residual variance.

## 2.8. Preliminary human experiments

We used a small human data set to demonstrate the impact of slice vs volume censoring on human data as follows: resting state fMRI (rs-fMRI) and MRI data were acquired from 64 baseline participants, 34 normal controls and 30 participants with major depressive disorder from UCA. The participant and acquisition details are described in (Lam et al., 2016; Wijk et al., 2021).

The resting state fMRI data were preprocessed with the OPPNI pipeline ((Churchill et al., 2017, 2015); software available at (<https://github.com/strotherlab/oppni>) using the following steps: (1) the volume with the least amount of head displacement was determined using a principal component analysis (PCA) and all volumes were registered to this volume with rigid-body motion correction via AFNI's 3dvolreg; (2) significant outlier volumes or slices were identified, removed and replaced by interpolated values using neighbouring volumes or slices, respectively, through censoring as implemented in ((Campbell et al., 2013); software available at: [nitrc.org/projects/spikecor\\_fmri](http://nitrc.org/projects/spikecor_fmri)); (3) slice-timing correction was performed with Fourier interpolation via AFNI's 3dTshift; (4) spatial smoothing across MRI scanners at different sites was matched using the 3dBlurToFWHM module in AFNI to smooth the fMRI images to the smoothness level of FWHM=6 mm in three directions (x,y,z); (5) AFNI's 3dAutomask algorithm was used to obtain a binary mask excluding non-brain voxels using default parameter settings, the resultant mask was applied to all EPI volumes prior to subsequent pipeline steps; (6) neuronal tissue masking was performed by estimating a probabilistic mask to reduce the variance contribution of non-neuronal tissues in the brain (macro-vasculature, ventricles) using the first part of the PHYCAA+ algorithm to estimate task-run and subject-specific neural tissue masks ((Churchill and Strother, 2013); software available at [nitrc.org/projects/phycaa\\_plus](http://nitrc.org/projects/phycaa_plus)); (7) Global signal regressors were estimated using PCA performed on each session's fMRI data as the PC1 time-series tends to be highly correlated with global signal effects but is orthogonal to the PC2+ signal subspace (Carbonell et al., 2011); (8) several nuisance regressors (low frequency temporal trends, head motion effects and global PC1) were calculated and then regressed-out from the data concurrently via multiple linear regression (Churchill et al., 2012a, 2012b); (9) physiological noise components were estimated and removed through data-driven physiological correction using the second part of the data-driven PHYCAA+ algorithm (Churchill and

Strother, 2013); software at [nitrc.org/projects/phycaa\\_plus](http://nitrc.org/projects/phycaa_plus)); (10) low-pass filtering was carried out using a linear filter to remove BOLD frequencies above 0.10 Hz; (11) spatial normalization to a structural template (sNORM) was carried out with all scans aligned to the MNI152 template (MNI-normalized, 4 mm resolution) using two transformations (fMRISubj -> MRISubj, and MRISubj->MNITemp, combined into one aggregated transform) via FSL's FMRIB's Linear Image Registration Tool (FLIRT) module.

For each rs-fMRI scanning session we generated two preprocessed data sets: one with outlier volumes censored (VC) and one with outlier slices censored (VS). As a preliminary test of the impact of these different censoring choices we examined changes between the two preprocessed data sets per session for the Posterior Cingulate Cortex connectivity seed PCC, MNI:[0, -60, 28] used by (Wijk et al., 2021). The mean time series within each ROI was extracted using AFNI's 3dROIstats for all of the 64 MNI-normalized preprocessed scans. AFNI's 3dTcorr1D was used to generate voxel-wise Pearson correlation between the mean time series within each ROI and every other voxel of each scan. Finally, correlation values were converted to z-scores using the variance stabilizing Fisher's transformation for correlations using AFNI's 3dCalc.

### 2.8.1. Split-half reproducible PCA

For a set of brain maps, PCA produces eigenimages that correspond to multivariate patterns explaining the greatest variance within the dataset. The reproducibility of these eigenimages can be quantified using a split-half resampling procedure, previously published in (Churchill et al., 2015). In this approach, a matrix of brain images  $X$ , of dimensions ( $V$  voxels  $\times$   $S$  samples) is randomly split in half, producing matrices  $X_1$  and  $X_2$ , of dimensions ( $V \times S/2$ ). Applying PCA obtains from  $X_1$  a set of eigenimages  $u_{1(k)}$ , and from  $X_2$  a set of eigenimages  $u_{2(k)}$ , where  $k = 1 \dots S/2$  indexes over eigenimages, ordered by decreasing amount of variance explained. We then obtain a z-scored map of voxel-wise reproducibility for each component pair  $k$ , using the procedure described in (Strother et al., 2002). This involves z-transforming  $u_{1(k)}$  and  $u_{2(k)}$ , before calculating signal-axis projection  $x_{sig(k)} = (u_{1(k)} + u_{2(k)})/\sqrt{2}$  and noise-axis projection  $x_{noi(k)} = (u_{1(k)} - u_{2(k)})/\sqrt{2}$ , then obtaining  $rSPM Z_{(k)} = x_{sig(k)}/SD(x_{sig(k)})$ . This process is repeated for 100 random split-half partitions, and the mean  $rSPM Z_{(k)}$  maps are retained.

## 3. Results

### 3.1. fBIRN QA measurements

To give a sense of the range of the values produced by the fBIRN QA pipeline for each site, the fourteen QA parameters are plotted for all sites in Fig. 1. The sites are labeled using a three-letter code defined in Table 1. Multiple potential site differences are seen in Fig. 1 and manufacturer differences in mean, SNR, and max and minFWHM values are clearly visible between the first five GE sites (CAM, MCM, SBH, TWH & UCA), the next six Siemens sites (BYC, WEU, QNS, TOH, UTO, SMH) and the remaining two Phillips sites (TBR, UBC). The differences are significant between sites (One-way MANOVA,  $p < 0.001$ ,  $\eta^2 = 0.43$ ) and between manufacturers (One-way MANOVA,  $p < 0.001$ ,  $\eta^2 = 0.90$ ).

Principal component analysis (PCA) provides a visualization of the normalised variance and correlation of the QA variables within imaging sites as well as between different sites and manufacturers (Fig. 2). 66% of the total variance is captured by the first two components and a substantial part of that is between-manufacturer variance in agreement with the relatively large partial association effect size ( $\eta^2 = 0.90$ ) for the between manufacturer MANOVA. For instance, in Fig. 2, Siemens scanners are mostly in the first quadrant (top right quadrant), while GE scanners are mostly located in the third (bottom left) and Phillips scanners in the second (top left) quadrants on the PCA plane. Furthermore, the within-manufacturer variance varies substantially between different manufacturers and sites. For example, for PC1 Siemens has

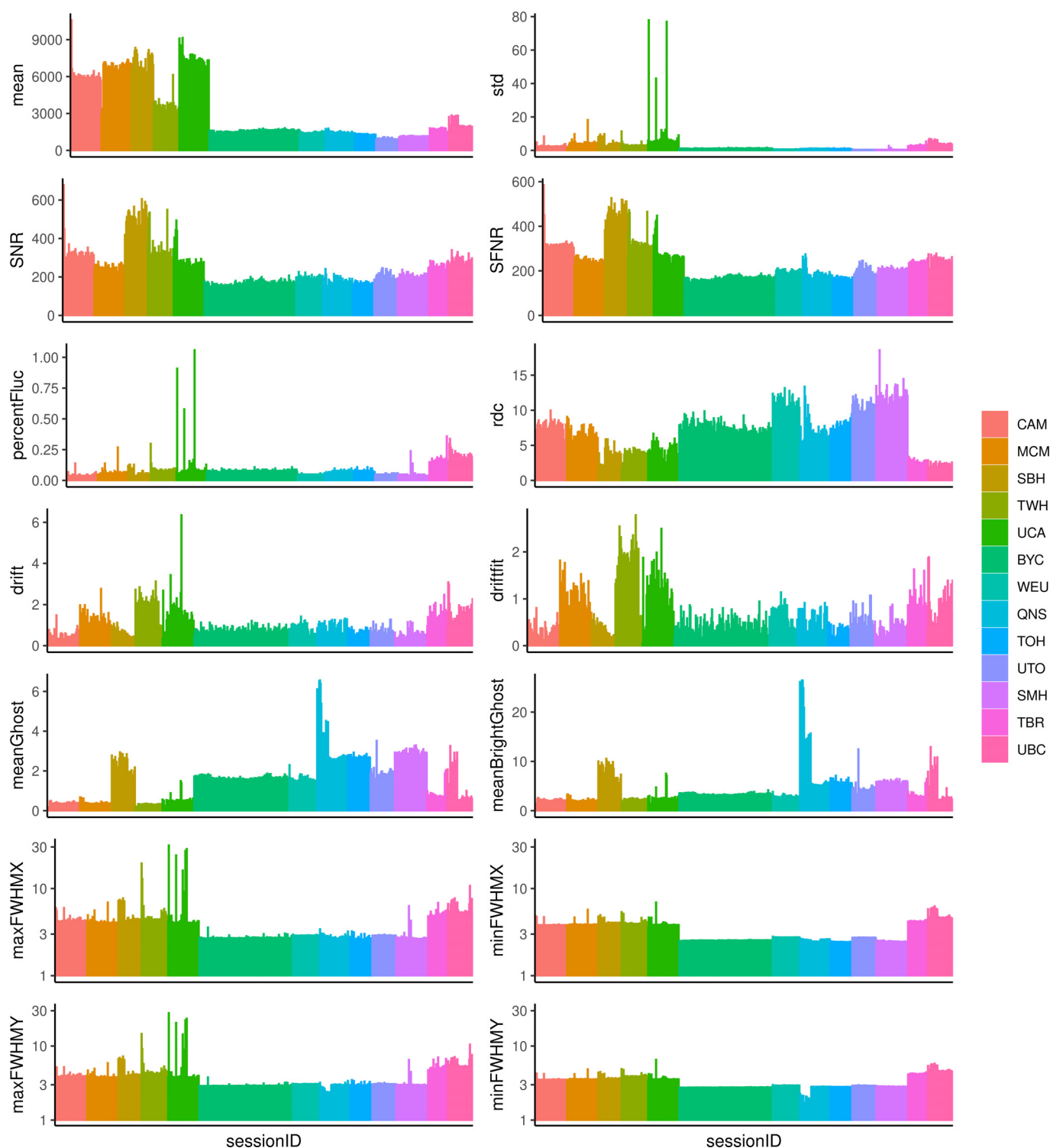
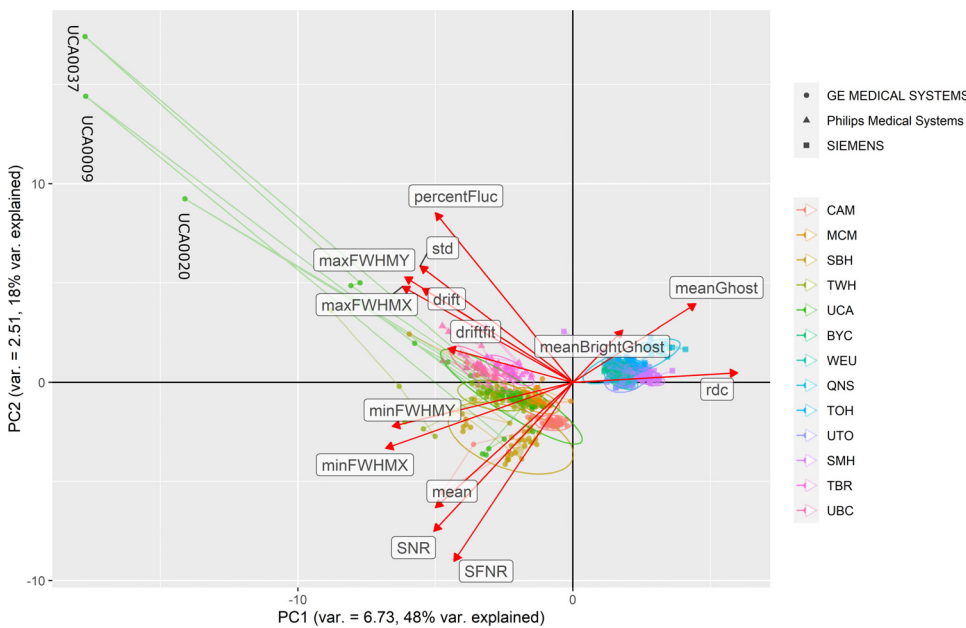


Fig. 1. fBIRN QA values for the 13 participating sites. Each vertical bar corresponds to an imaging session. Sessions are ordered chronologically for each site. The sites are labeled with a three-letter code as defined in Table 1. Note that FWHM values are plotted in mm on log scales and are based on the fBIRN use of AFNI’s “classic” FWHM estimate as a function of the ratio of variance of first differences to data variance.

the smallest within-manufacturer variance and GE has the largest (Appendix Table A4). Moreover, we observe prominent anomalous scans in a number of sites (most strikingly in UCA, TWH, SMH, UBC, and TBR), as shown in Fig. 2, where the imaging sessions within each site are connected in temporal order of acquisition.

Finally, the 14 QA variables shown as red arrows with lengths proportional to their variable loadings in Fig. 2 suggest two primary clusters

of variable effects over time, and site and manufacturer. The first cluster of variables is roughly arrayed around the 45° diagonal (i.e., correlated; see Fig. A1 in Appendix) within the bottom left quadrant and includes minFWHM<sub>y</sub>, minFWHM<sub>x</sub>, mean, SNR and SFNR with negative variable loadings on PC1 and PC2. Therefore the individual scan sessions from GE and Phillips scanners with their negative factor/observation loadings x the negative variable loadings for these five variables all have incre-



**Fig. 2.** Principal component analysis (PCA) of the fBIRN QA variables. Shown is the scatter plot of the first two principal components (PC1 and PC2) for all the imaging sessions (including outliers). Each imaging session is color-coded by scan site and shape-coded by scanner manufacturer. The imaging sessions within each site are connected in temporal order of acquisition. The ellipses show 95% confidence bounds assuming a multivariate t distribution. A few examples of prominent anomalous sessions are highlighted with labels for UCA. The fBIRN QA variables are also plotted using red arrows on the PCA plane with lengths proportional to the variable loadings. An interactive version of this plot can be viewed at [https://kayvanrad.github.io/phantomQA/#fbirn\\_qa\\_pca](https://kayvanrad.github.io/phantomQA/#fbirn_qa_pca).

**Table 3**  
Linear mixed effects fit of the QA parameters to the scanners (i.e., sites) as a random effect.

QA parameter	Intercept estimate (95%-CI)	Scanner (Site) SD (ICC)	Residual SD	Effect size (R <sup>2</sup> )
mean	3171* (-330.8, 6659)	2546* (0.97)	443	0.97
std	2.96* (1.42, 4.50)	2.65* (0.27)	4.34	0.27
SNR	259* (207, 311)	92.7* (0.88)	34.0	0.86
SFNR	250* (204, 296)	81.21* (0.89)	28.6	0.88
percentFluc	0.0896* (0.0622, 0.117)	0.0479* (0.40)	0.0584	0.34
rdc	6.55* (4.80, 8.30)	3.10* (0.87)	1.22	0.83
drift	1.03* (0.752, 1.31)	0.490* (0.63)	0.379	0.58
driftfit	0.671* (0.435, 0.906)	0.415* (0.66)	0.300	0.63
meanGhost	1.57* (0.973, 2.16)	1.05* (0.85)	0.440	0.83
meanBrightGhost	4.44* (3.08, 5.79)	2.38* (0.54)	2.19	0.51
minFWHMx	3.46* (2.96, 3.96)	0.883* (0.92)	0.261	0.91
maxFWHMx	4.19* (3.35, 5.02)	1.45* (0.31)	2.16	0.31
minFWHMy	3.44* (3.04, 3.83)	0.700* (0.90)	0.243	0.88
maxFWHMy	4.10* (3.41, 4.79)	1.19* (0.30)	1.80	0.30

\*  $p < 0.05$ .

mentally positively larger values than for the Siemens’s scanner sessions in the upper right quadrant (i.e., positive factor/observation loadings x negative variable loadings). This shows that the manufacturer dependent blocks seen in Fig. 1, with GE and Philips FWHM greater than Siemens, are quantitatively dominant effects in the overall data set. We also note that in terms of these main effects, multiple QA variables appear to be somewhat redundant with meanGhost, rdc and the weaker meanBrightGhost (shorter red vector) inversely correlated with the positively correlated minFWHMy, minFWHMx, mean, SNR and SFNR. The second cluster of variables occurs around the 45° diagonal (i.e., correlated) within the upper left quadrant and includes percentFluc, std, drift, maxFWHMx, maxFWHMy and the weaker driftfit, and all record temporal variation within- and between- scan sessions within a site, likely driven by the multiple outlying scan sessions noted in GE and Philips scanners above.

### 3.1.1. Main drivers of variance and anomalous scan sessions

To further quantitatively understand the structure of these effects we used linear mixed effects models fitting the QA parameters to scanners (i.e., sites) and, separately, to scanner manufacturers as random effects. These results are summarized in Tables 3 and 4, respectively. Substantial variance in mean, SNR, SFNR, meanGhost, minFWHMx, and minFWHMy is explained by scanner (ICC  $\geq 0.85$ ,  $p < 0.05$  for SFNR and  $p < 0.001$  for others) and substantial variance in mean, minFWHMx and

minFWHMy is explained by scanner manufacturer random effects (ICC  $\geq 0.85$ ,  $p < 0.001$ ). This is consistent with the PCA factor loadings in Fig. 2, where minFWHMx, minFWHMy, and mean drive the between-manufacturer difference between the GE/Philips versus Siemens clusters. Notably, however, limited variance is explained in measures of max FWHM, i.e., maxFWHMx and maxFWHMy, by scanner and manufacturer (ICC  $< 0.40$ ,  $p < 0.001$ ). In consistency with these low ICC values, measures of maxFWHM, which show weaker correlations with minFWHM (Fig. A1) and are almost orthogonal to minFWHM on the PCA plane (Fig. 2), do not appear to drive substantial between-manufacturer variance. Nevertheless, a large number of the anomalous scan sessions appear to be associated with maxFWHM (Fig. 2). This furthermore indicates that min and max FWHM are potentially driven by different factors.

### 3.2. Spatial smoothing reduces variance in minFWHM

We tested the extent to which scanner differences in imaging resolution can be alleviated by smoothing the images to the greatest mean FWHM of all sites, as advocated by (Friedman et al., 2006). Our data show that after spatial smoothing to 7 mm by AFNI’s 3dBlurToFWHM, scanner effects contribute to a lesser extent to the variance in minFWHM in the x (LMM, ICC = 0.81,  $p < 0.001$ ) and y (LMM, ICC = 0.72,  $p < 0.001$ ) directions, amounting to 12% and 20% reduction in ICC for



**Table 4**  
Linear mixed effects fit of the QA parameters to the scanner manufacturer as a random effect.

QA parameter	Intercept estimate (95%-CI)	Manufacturer SD (ICC)	Residual SD	Effect size (R <sup>2</sup> )
mean	3206 ( $p = 0.17$ ) (-235, 6648)	2614* (0.89)	910	0.86
std	3.32 ( $p = 0.12$ ) (0.437, 6.22)	2.16* (0.18)	4.61	0.16
SNR	261* (162, 361)	75.3* (0.61)	59.9	0.58
SFNR	248* (160, 337)	67.3* (0.63)	51.3	0.59
percentFluc	0.112 ( $p = 0.096$ ) (0.0274, 0.197)	0.0644* (0.52)	0.0619	0.25
rdc	5.45 ( $p = 0.095$ ) (1.32, 9.56)	3.13* (0.71)	1.99	0.55
drift	1.16* (0.617, 1.70)	0.409* (0.40)	0.505	0.25
drifffit	0.712* (0.374, 1.05)	0.255* (0.26)	0.430	0.22
meanGhost	1.34 ( $p = 0.094$ ) (0.324, 2.36)	0.770* (0.49)	0.791	0.44
meanBrightGhost	4.085* (2.852, 5.32)	0.903* (0.080)	3.00	0.072
minFWHMx	3.76* (2.35, 5.16)	1.069* (0.92)	0.308	0.88
maxFWHMx	4.58* (2.52, 6.64)	1.56* (0.33)	2.23	0.25
minFWHMy	3.71* (2.52, 4.90)	0.905* (0.90)	0.295	0.82
maxFWHMy	4.47* (2.69, 6.27)	1.35* (0.34)	1.87	0.23

\*  $p < 0.05$ .

**Table 5**

Percentage of scan sessions with anomalies for each site for different anomaly thresholds  $\theta$ . A slice is identified as anomalous if  $\text{FWHM}_x > \theta \times \text{minFWHM}_x$  or  $\text{FWHM}_y > \theta \times \text{minFWHM}_y$ ,  $\text{minFWHM}_x$  and  $\text{minFWHM}_y$  are the minimum FWHM values in the x and y directions in that scan session and  $\theta$  is an anomaly threshold. U: percentage of anomalous scan sessions for the unprocessed raw data, S: percentage of anomalous scan sessions for data processed using slice SpikeCor, V: percentage of anomalous scan session for volume SpikeCor processed data.

Site	# sessions	$\theta = 2$			$\theta = 3$			$\theta = 5$			$\theta = 7$			$\theta = 10$		
		U(%)	S(%)	V(%)	U(%)	S(%)	V(%)	U(%)	S(%)	V(%)	U(%)	S(%)	V(%)	U(%)	S(%)	V(%)
BYC	141	99	76	99	41	4	40	6	0	4	3	0	1	2	0	0.7
CAM	48	100	100	100	46	38	40	8	8	8	8	8	8	8	4	6.2
MCM	47	100	96	96	55	19	23	4	4	4	4	2	4	2	2	2.1
QNS	46	87	57	87	61	2	61	48	2	48	15	2	11	2	0	2.2
SBH	34	103	103	103	103	88	88	74	44	44	35	32	32	29	29	29.4
TBR	30	100	100	100	100	100	100	100	97	100	90	50	90	80	10	73.3
SMH	48	10	6	10	8	6	8	6	4	6	4	4	4	4	2	2.1
TOH	33	97	45	97	61	9	48	9	0	6	6	0	3	3	0	0
TWH	40	100	100	100	100	100	100	92	88	92	50	32	48	18	12	17.5
UBC	38	100	100	100	100	100	100	100	100	100	100	95	100	89	74	92.1
UCA	47	100	100	100	98	68	77	36	26	30	17	15	15	15	15	14.9
UTO	36	8	11	8	0	0	0	0	0	0	0	0	0	0	0	0
WEU	41	10	10	12	0	0	0	0	0	0	0	0	0	0	0	0

$\text{minFWHM}_x$  and  $\text{minFWHM}_y$ , respectively. Smoothing brings the minimum FWHM up to an approximately common value for all the scanners (Appendix Table A6) and reduces the variance of  $\text{minFWHM}_x$  and  $\text{minFWHM}_y$  by 81% and 71%, respectively. Nevertheless, the scanner effect is far from eliminated. Moreover, smoothing does not appear to remove and/or alleviate the anomalous sessions, indicating that these anomalies are potentially driven by factors other than differences in reconstruction resolution (Appendix Fig. A4).

### 3.3. ACF Analysis

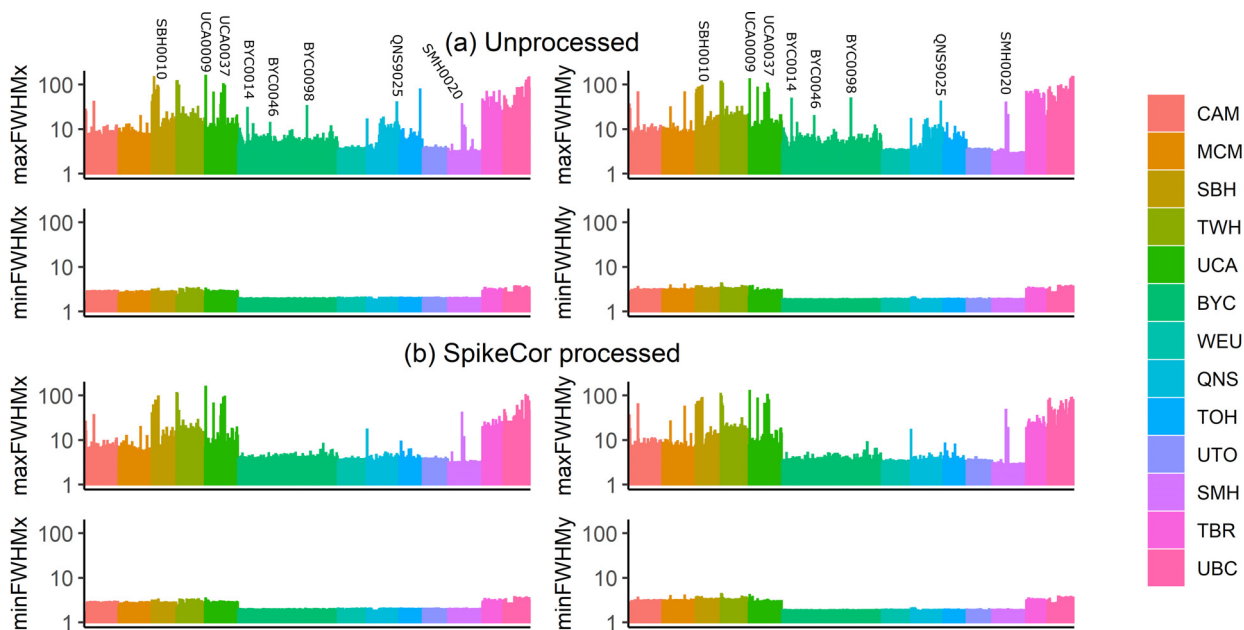
As noted, imaging resolution (measured in terms of FWHM) is one of the main factors driving the within- and between-site variance and temporal anomalies. To further investigate imaging resolution, we looked at an independent slice-wise measure of resolution using ACF. This was motivated by two factors: (1) to provide an independent measure of FWHM compared to the fBIRN QA pipeline FWHM measures, and (2) to calculate slice-wise measures of FWHM, as our preliminary investigation of the anomalous images showed possible slice effects (See, for example, Fig. 6) and the fBIRN QA pipeline does not provide slice measurements of FWHM - as noted the fBIRN QA pipeline uses the “classic” estimate of FWHM, defined as a function of the ratio of variance of first differences to data variance (Forman et al., 1995).

Fig. 3(a) shows the min and max slice ACF FWHM values for the 13 participating sites. ACF FWHM values reveal relatively stable but distinct manufacturer differences in minimum imaging resolution in x (LMM, ICC = 0.90,  $p < 0.001$ ) and y (LMM, ICC = 0.92,  $p < 0.001$ ) directions, where  $\text{minFWHM}$  ranges around 2 mm for Siemens and around

3 mm for GE and Philips scanners (Appendix Table A6). Scanner manufacturer to a lesser extent explains the variance in maximum slice ACF FWHM in x (LMM, ICC = 0.60,  $p < 0.001$ ) and y (LMM, ICC = 0.72,  $p < 0.001$ ) directions. Nonetheless, these measures reveal multiple anomalous scan sessions with suspiciously high  $\text{maxFWHM}$  values. While some of these anomalies are observed in the fBIRN QA measures (e.g., those of UCA), these anomalies are more prominently revealed by the slice-based ACF measures - for example, for BYC0098, compare the max ACF FWHM values in Fig. 3(a), with max fBIRN QA (volume-based) FWHM values in Fig. 1.

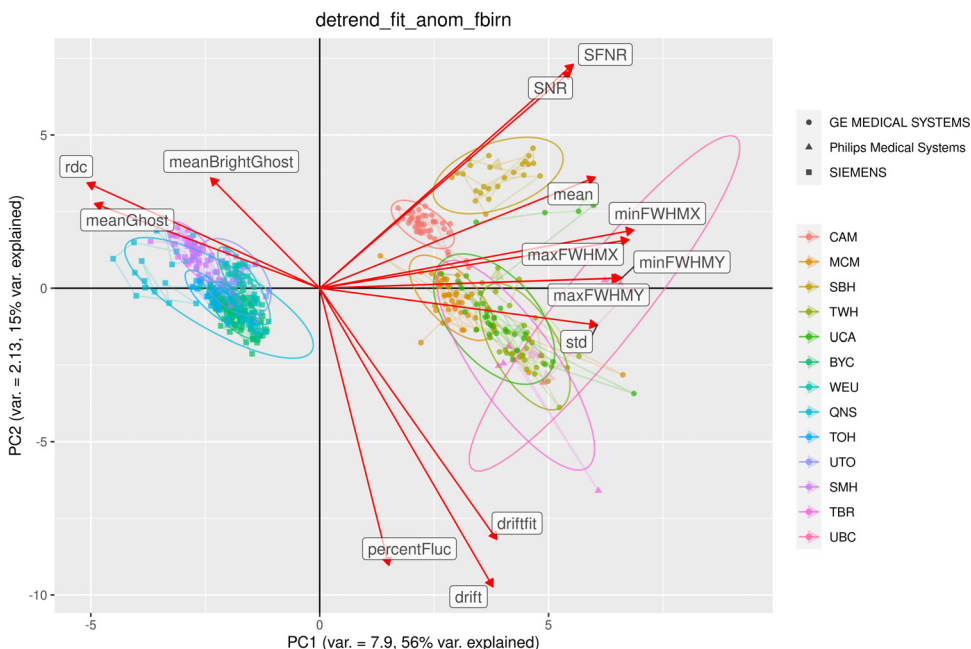
### 3.4. Scan sessions with anomalies

Table 5 summarizes the number of anomalous scan sessions detected using the slice ACF FWHM values in each site for different thresholds  $\theta$ . The number of anomalous sessions incrementally decreases with increasing the threshold. While the number does not decrease at a uniform rate, with sharper decrease between some threshold values compared to the others, our data do not show any consistent pattern for different sites. For example, for WEU the number of anomalies drops from 4 to 0 between  $\theta = 2$  and  $\theta = 3$ , but for BYC, the number drops from 58 to 8 between  $\theta = 3$  and  $\theta = 5$ . Appendix Table A3(a) summarizes the anomalous scan sessions detected using a conservative threshold of  $\theta = 10$ . Presuming the min FWHM value reflects the manufacturer’s intended resolution, any slice with a FWHM value one order of magnitude larger than the intended resolution is considered anomalous as such a huge discrepancy is undesired and can have detrimental effects on the data.



**Fig. 3.** Slice ACF FWHM values for the 13 participating sites for (a) Unprocessed data, and (b) data processed with SpikeCor. Each bar corresponds to an imaging session. Sessions are ordered chronologically for each site. The sites are labeled with a three-letter code as defined in Table 1. A few anomalous sessions are labeled on the plot. Note that FWHM values are plotted in mm on log scales. While SpikeCor effectively removed anomalies in some sessions (e.g., BYC0014, BYC0046, BYC0098, and QNS9025), it was not able to ameliorate the anomalies in some other sessions (e.g., SBH0010, SMH0020, UCA009, and UCA0037).

Principal component analysis of the fBIRN QA parameters of the data excluding the  $\theta = 10$  anomalous sessions verify the impact of anomaly detection based on slice ACF FWHM on the fBIRN QA variance (Fig. 4). With the anomalies removed, maxFWHM and minFWHM, which were previously almost orthogonal to each other (Figs. 2 and A1(a)), become strongly correlated (Figs. 4, and A1(b)), together driving the between-manufacturer variance, which is now more clearly reflected by PC1 than in Fig. 2. Consistently, with anomalies excluded, substantial variance is now explained by the scanner manufacturer effect in maxFWHMx (LMM, ICC = 0.98,  $p < 0.001$ ) and maxFWHMy (LMM, ICC = 0.98,  $p < 0.001$ ), in addition to minFWHMx (LMM, ICC = 0.98,  $p < 0.001$ ) and minFWHMy (LMM, ICC = 0.96,  $p < 0.001$ ) (Compare with Table 4).

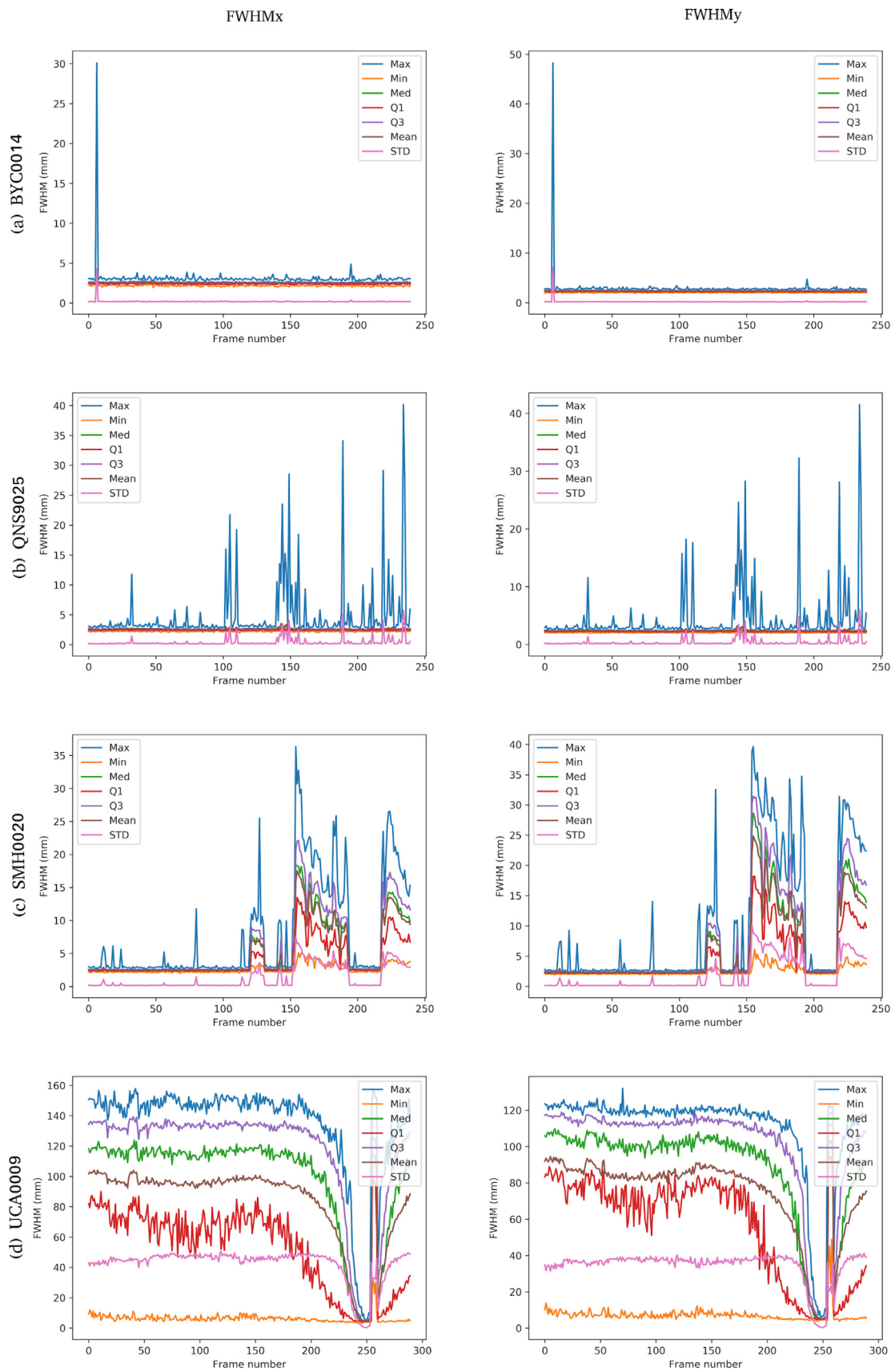


**Fig. 4.** Principal component analysis (PCA) of the fBIRN QA variables after the removal of the anomalous scan sessions within each site using slice ACF FWHM measurements with  $\theta = 10$ . Shown is the scatter plot of the first two principal components (PC1 and PC2) of the fBIRN QA variables for all the imaging sessions. Each imaging session is color-coded by scan site and shape-coded by scanner manufacturer. The ellipses show 95% normal probabilities. The original fBIRN QA variables are also plotted using red arrows on the PCA plane with lengths proportional to the variable loadings. An interactive version of this plot can be viewed at [https://kayvanrad.github.io/phantomQA/#anomalies\\_removed\\_pca](https://kayvanrad.github.io/phantomQA/#anomalies_removed_pca).

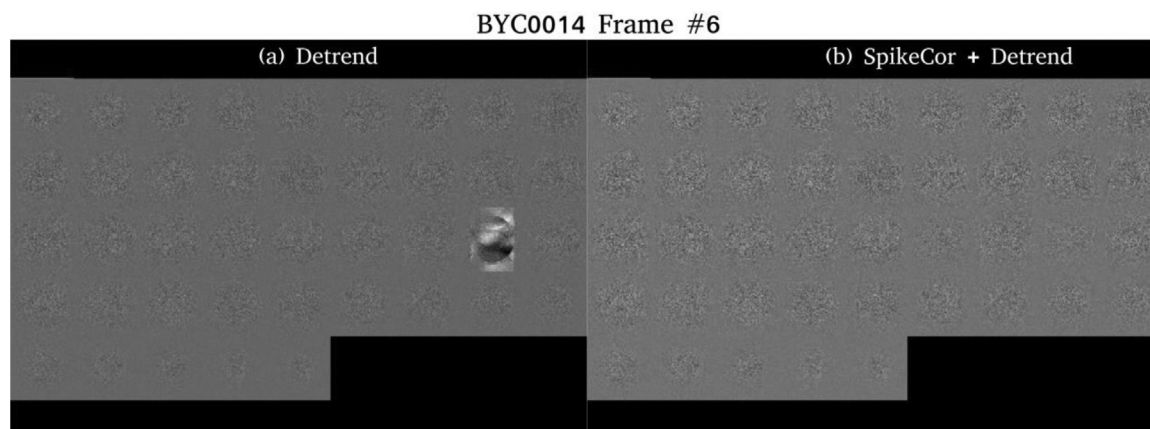
Fig. 5 summarizes slice ACF FWHM values for four representative anomalous sessions, namely, BYC0014, QNS9025, SMH0020, and UCA0009, identified in Fig. 3. Slice ACF FWHM values are shown for a normal session (SMH0019) in Fig. A2. The following examples demonstrate various cases that we observed in our data:

BYC0014 shows one prominent spike in maxFWHM in only one time frame (frame #6) (Fig. 5(a)). A visual investigation of this acquisition reveals only one anomalous slice in that volume (Fig. 6(a)), which is manifest in the maxFWHM values.

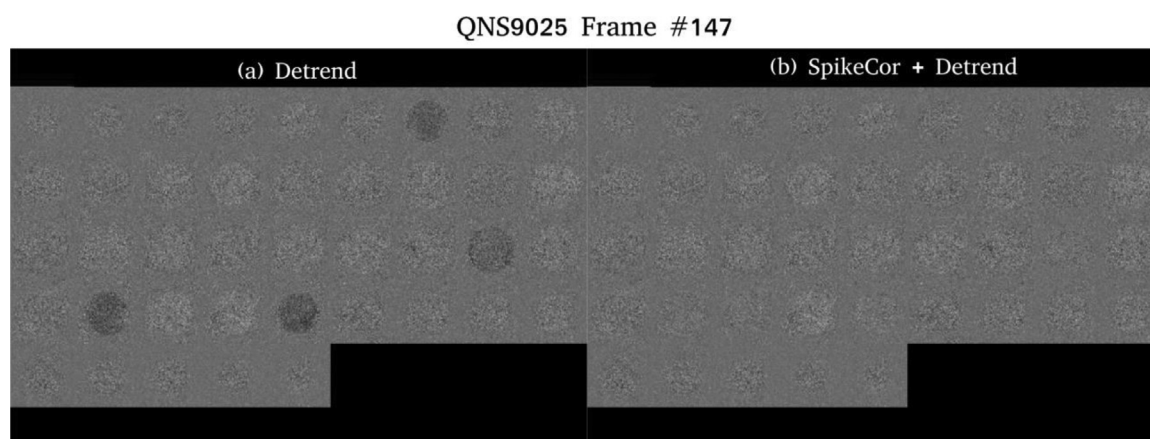
QNS9025 shows multiple spikes in maxFWHM over several time-frames (Fig. 5(b)). Nevertheless, these spikes do not manifest in the mean, median, or the quartiles, indicating that these effects are due to



**Fig. 5.** Summary of slice ACF measures of FWHM for four anomalous sessions: (a) BYC0014, (b) QNS9015, (c) SMH0020, and (d) UCA0009. Plotted are max, min, median, first quartile (Q1), third quartile (Q3), mean, and standard deviation (STD) values of slice ACF FWHM for each volume in the time series in the x and y directions from the 41 slices in each 2400 ms scan frame.



**Fig. 6.** BYC0014 Frame #6- Volume with a single anomalous slices. (a) Detrended data. A single anomalous slice stands out. (b) Data preprocessed by SpikeCor censoring followed by detrending. The anomalous slice is censored out (i.e., replaced with interpolated data) by SpikeCor.



**Fig. 7.** QNS9025 Frame #147- Volume with a few anomalous slices. (a) Detrended data. A few anomalous slices are apparent. (b) Data preprocessed by SpikeCor censoring followed by detrending. The anomalous slices are censored out by SpikeCor.

one or a few individual slices but not the majority of slices in the corresponding volumes. Fig. 7(a) shows an example of one such volume with a few anomalous slices.

SMH0020 shows blocks of defective volumes with unusually high measures of FWHM (Fig. 5(c)). The anomalies are manifest not only in maxFWHM values, but also in other measures of center and spread (mean, median, quartiles). These anomalies are due to multiple and sometimes many anomalous slices in the corresponding volumes. Fig. 8(a) displays these anomalies in one such volume. For comparison, a normal volume in the same scan session is shown in Fig. 8(c). Fig. A2 displays slice FWHM summary values for SMH0019 (the monthly scan session immediately prior to SMH00200), which has normal FWHM values.

Most of the frames (i.e., volumes) in UCA0009 are defective with slice instabilities (Fig. 5(d)). Moreover, these anomalies are manifest in the median and the quartiles in addition to the max FWHM values, indicating that a large portion of the slices are defective in each volume.

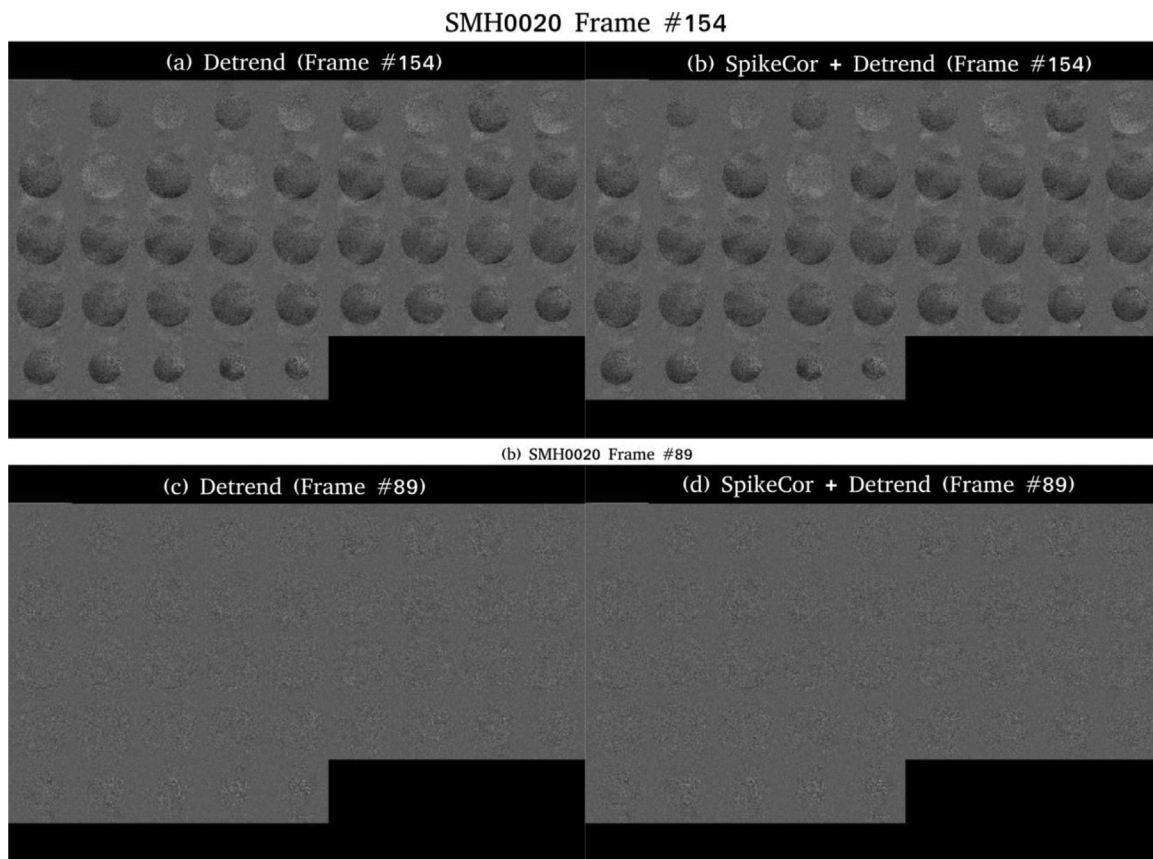
#### 3.4.1. Prevalence of slice anomalies per scan session

Fig. A3(a) summarizes the total number of slice anomalies and the median of the number of slice anomalies per volume for all the scan sessions. As noted, the distribution of slice anomalies varies greatly between different scan sessions. For example, while there are a number of anomalous slices in QNS9025 and SMH0020, the median of the number of anomalies per volume is zero for those scan session, indicating that less than half of the frames (i.e., volumes) in the acquisition were affected by these instabilities. On the other hand, UCA0009 has a large to-

tal number of anomalies, and a large median of the number of anomalies per volume, indicating that a substantial number of slices in a substantial number of volumes (more than half) were affected by these anomalies. These observations are consistent with those of Fig. 5.

#### 3.5. Correction of anomalous slices

Our data show that slice-based SpikeCor can effectively identify anomalous slices under certain circumstances, namely, when only a fraction of the slices are defective, but fails when they constitute a large fraction of the slices. For example, slice-based SpikeCor succeeds in removing or alleviating the anomalies in BYC0014 (Figs. 9(a) and 6(b)) and QNS9025 (Figs. 9(b) and 7(b)). On the other hand, SpikeCor does not appear to be effective in removing the anomalies for SMH0020 (Figs. 9(c) and 8(b)) and for UCA0009 (Fig. 9(d)). Furthermore, SpikeCor does not affect normal acquisition, e.g. SMH 0019 (Fig. A2). Our data also indicate that slice-based SpikeCor is more effective than the volume-based SpikeCor in removing the anomalies, regardless of the threshold used to detect anomalies (Table 5). Appendix Table A3(b) lists the anomalous scan sessions, identified using a conservative threshold of  $\theta = 10$ , for the unprocessed data as well as for volume-based and slice-based SpikeCor preprocessing for each center. With the exception of volume-based SpikeCor for UBC, SpikeCor (both slice- and volume-based) processed anomalies are a subset of those of the raw data. That is, SpikeCor does not produce additional synthetic anomalies.



**Fig. 8.** SMH0020 Frame #154 (a,b)- Volume with several anomalous slices. (a) Detrended data. Several anomalous slices strike out. (b) Data preprocessed by SpikeCor censoring followed by detrending. SpikeCor fails to detect and remove anomalous slices. Frame #89 (c,d)- Normal volume. (c) Detrended data. No anomalous slices are observed. (d) Data preprocessed by SpikeCor censoring followed by detrending. SpikeCor does not appear to affect the volume.

### 3.6. Preliminary human experiments

For the PCC seed and the preprocessing pipelines using volume- (VC) or slice-based (SC) censoring, z-mapped session connectivity volumes (CONNz) were analysed using split-half PCA to ensure a significant reinforcing connectivity map that is stable and reproducible exists across the 64 CONNz volumes. Individual session variance was calculated for the VC CONNz (VCz) and SC CONNz (SCz) volumes using all PC components from the split-half PCA.

Split-half PCA shows one highly-significant (VC: $p < 0.01$ ; SC: $p < 0.01$ ) highly reproducible (VC: $R = 0.95$ ; SC: $R = 0.95$ ) connectivity pattern accounting for (VC:44%; SC:45%) of total variance (Fig. 10). This reflects some of the classical default mode regions seen when using a PCC seed in resting state data with the large Z-values and associated high pattern reproducibilities indicating a strong stable connectivity pattern being reinforced across most sessions.

Using either a paired  $t$ -test ( $p = 0.56$ ) or Wilcoxon non-parametric run test ( $p = 0.64$ ) there is no difference between VCz and SCz session variance distributions.

To explore the possible changes in seed connectivity values per voxel we regressed the brain-masked SCz voxel values against the VCz voxel values separately for each session. Fig. 11 demonstrates these regression slope values plotted as a function of the 64 session numbers so that scan date progresses from left (Session 1) to right (Session 77) from October 2013 through August 2016. Note that there are missing sessions with a total of only 64 rs-fMRI data sets because session numbers were taken from CAN-BIND study enrollment numbers, and for a variety of reasons not all participants had a MRI scan. There is no significant trend of slope with session number ( $p = 0.41$ ;  $F = 0.70$ ). However, the 32 sessions up to and including session 42 (large red disk) have a lower slope variance than for sessions after 42. The first 32 slope values have a variance ratio

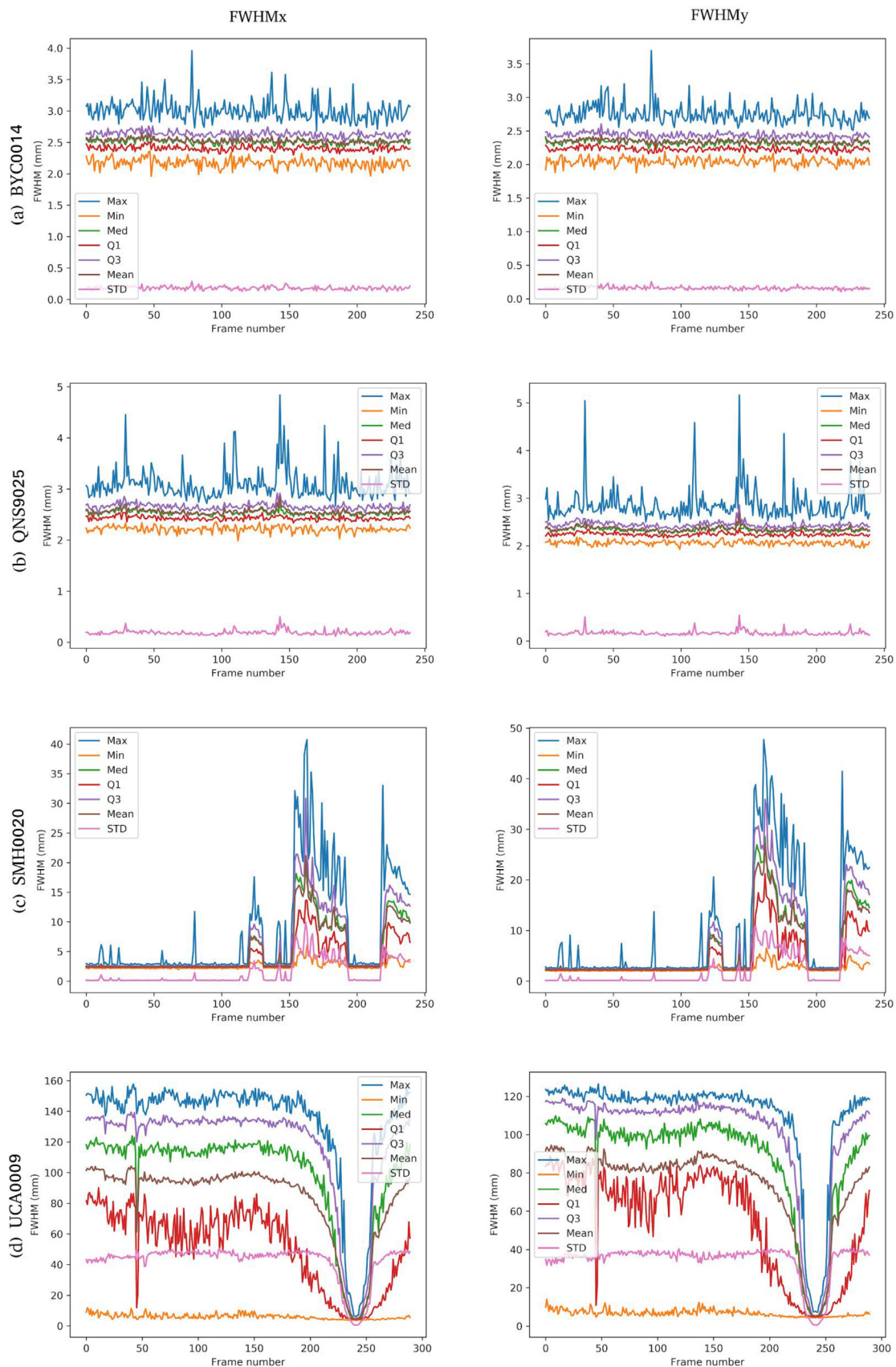
of 0.425 (two-tailed F-test:  $F = 0.425$ ,  $p < 0.05$ , 95%-CI = [0.21, 0.87]) of the slope value variance for the 32 session  $\#s > 42$ . This fits with our monthly fBIRN phantom measurements from Calgary showing that the scanner became more unstable in the later period of human data collection for CAN-BIND.

In Fig. 11 we see both large increases and decreases in overall z-scores above 20%. Fig. 12(a) shows session 64 with a large decrease (red disk in Fig. 11), and Fig. 12(b) shows session 42 with a large increase (purple disk in Fig. 11). Overall 46 of the 64 (72%) sessions have slopes less than or equal to 1.0 showing that the primary impact of slice censoring versus volume censoring is to reduce z-scores.

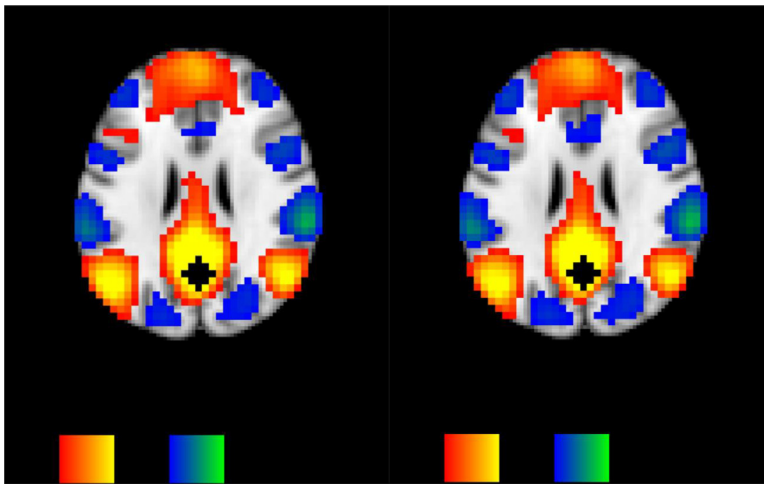
## 4. Discussion

Scanner stability is an important concern for fMRI experiments (Friedman and Glover, 2006). Instabilities can add unexplained variance to the data. This is particularly relevant to multi-site longitudinal studies. Not only is it important to ensure the scanner remains stable throughout a run (set of volumes acquired using a prescribed pulse sequence), it is also very important to ensure stability within each site over time (i.e., between different imaging sessions in one site), and between sites. Our data highlight the importance and efficacy of phantom scans for monitoring the stability of MRI scanners and identify an additional important stability dimension across slices within a volume. Routine QA phantom scans at all sites effectively detected instabilities across all of these possible stability dimensions across multiple scanners.

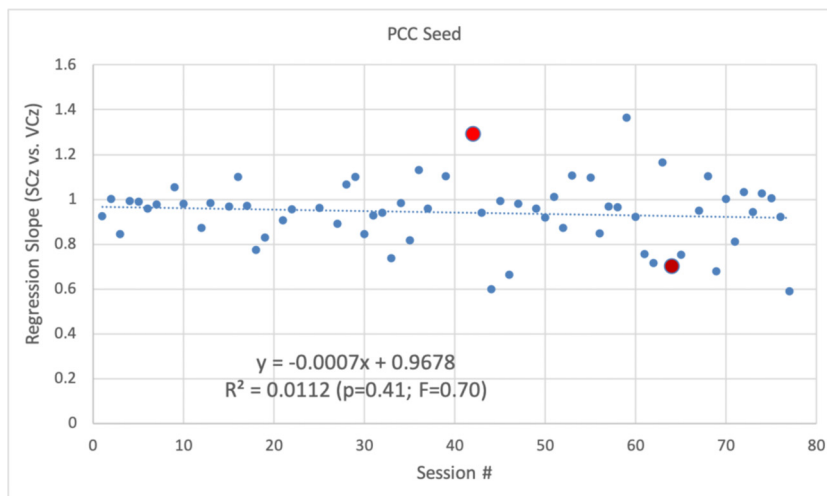
As noted above, we used the fBIRN QA pipeline to calculate 14 QA parameters, the utility of which has been described in prior works (Friedman and Glover, 2006; Glover et al., 2012). It was unknown to what extent these 14 variables capture different variance / noise sources and one of the goals of the current study was to evaluate the utility of



**Fig. 9.** Summary of slice ACF measures of FWHM after SpikeCor censoring for the scan sessions of Fig. 5(a) BYC0014, (b) QNS9015, (c) SMH0020, and (d) SMH0019. Plotted are max, min, median, first quartile (Q1), third quartile (Q3), mean, and standard deviation (STD) values of ACF FWHM for each volume in the time series in the x and y directions.



**Fig. 10.** Split-half PCC-seed, group zMAP from VCz (left) and SCz (right) sessions. PCC seed shown as black cross. Negative color scale from dark blue ( $Z=-5$ ) to green-white ( $Z=-25$ ), and positive color scale from red ( $Z=5$ ) to yellow-white ( $Z=25$ ).



**Fig. 11.** Slope of SCz voxel values regressed against the VCz voxel values separately for each session plotted chronologically as a function of session number.

the published fBIRN parameters for monitoring scanner stability and precision, and identify the most important QA variables to monitor.

#### 4.1. Between-manufacturer variance

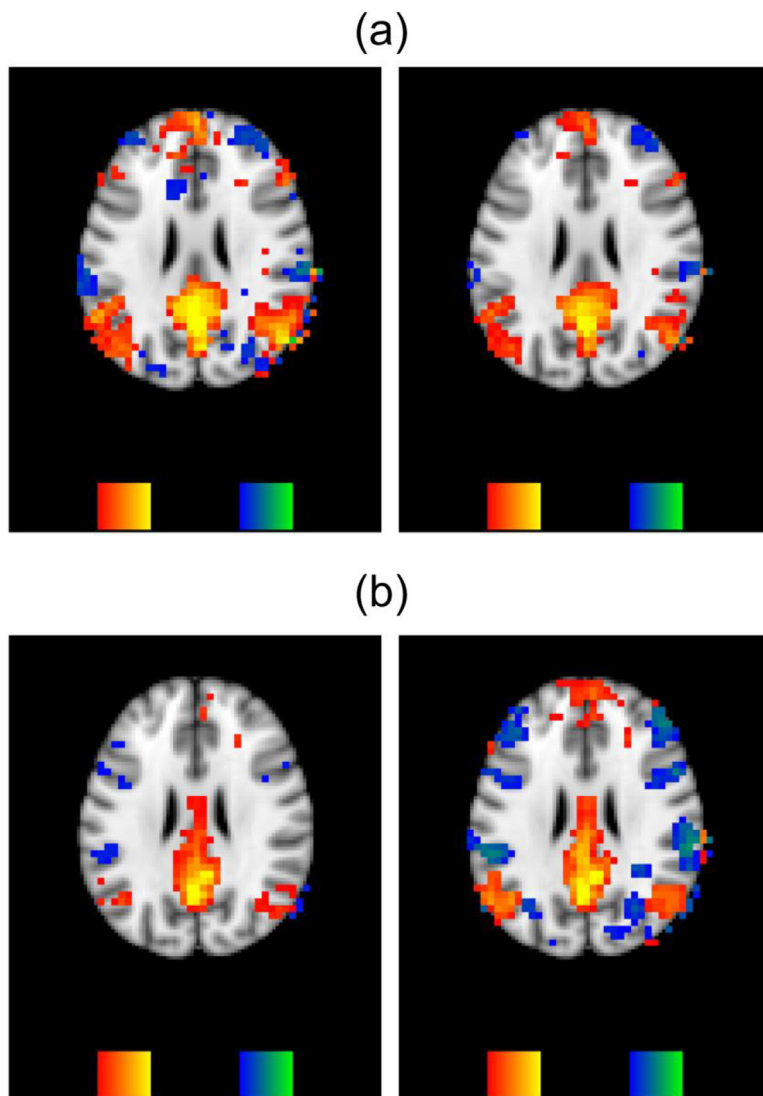
Our data indicate that between manufacturer differences in imaging resolution are a major source of variance in multisite studies. In particular, we identified two distinct mechanisms by which these differences can contribute to the variance: (1) differences in minFWHM values, attributed to the differences intrinsic resolution of the scanners, and (2) differences in maxFWHM values, attributed to slice anomalies.

We identified two primary clusters of variable effects over time, and site and manufacturer (Fig. 2). The first cluster of variables includes minFWHM<sub>y</sub>, minFWHM<sub>x</sub>, mean, SNR, and SFNR. The positive correlation between these variables makes sense since the higher FWHM in GE and Phillips scanners leads to higher SNR/SFNR values. We also noted that meanGhost, meanBrightGhost, and rdc are inversely correlated with the positively correlated minFWHM<sub>y</sub>, minFWHM<sub>x</sub>, mean, SNR, and SFNR. While minFWHM<sub>x</sub>, minFWHM<sub>y</sub>, and mean may be considered manufacturer traits, SNR, SFNR, rdc, and meanGhost are more site dependent and capture significant scanner differences within manufacturers. Furthermore, these variables might be expected to be relatively stable within sessions, whereas the second cluster of variables, including percentFluc, std, drift, maxFWHM<sub>x</sub>, maxFWHM<sub>y</sub>, and riftfit, which are located at approximately 90° to the first cluster on the PCA plane, largely reflect within session temporal fluctuations. The correlation (positive or negative) between the QA variables within each cluster indicate that multiple QA variables appear to be somewhat redundant.

We note that while there exists a nested dependence among the observations due to site effects, the main objective of the PCA illustrations is descriptive rather than inferential, which is not affected by complications due to non-independence of the observations. As noted above, the PCA illustrations provide insight into the structure of the data, including the between- and with-site effects.

We observe significant contributions of scanner manufacturer effects to the variance in minFWHM (LMM, ICC  $\geq 0.90$ ,  $p < 0.001$ ), which presumably reflects the manufacturer's intended imaging resolution. Our data show clear differences in minFWHM between manufacturers (Figs. 1, 3, and Appendix Table A6), which are attributed to the difference in the reconstruction techniques used by different vendors (e.g., differences in apodization, etc.) (Friedman et al., 2006). Moreover, we noted that preprocessing by blurring all images to a common spatial resolution (e.g., the coarsest spatial resolution of all scanners) can reduce scanner-related variance in imaging resolution, reflected in minFWHM values (Table A6). However, mean differences of  $> 1$  mm remain suggesting that a new algorithm other than the current AFNI 3dBlurToFWHM (Version AFNI\_17.3.05) is needed to further reduce this major inter-scanner difference and its impact on SNR.

The anomalous scan sessions, on the other hand, are reflected in maxFWHM, which is almost orthogonal to the minFWHM (Fig. 2). This indicates that maxFWHM, which is more than one order of magnitude larger than the minFWHM for some scan sessions, is driven by factors other than the vendor-specific reconstruction resolution. Nevertheless, min and max FWHM become more strongly correlated if the anomalous sessions are excluded (Fig. 4) and scanner manufacturer becomes a substantial driver of variance in maxFWHM in addition to minFWHM (LMM, ICC  $\geq 0.90$ ,  $p < 0.001$ ). Furthermore, with these sessions ex-



**Fig. 12.** (a) Session 64 CONN zMAP for VCz (left) and SCz (right). Negative color scale from dark blue ( $Z=-0.5$ ) to green-white ( $Z=-1.5$ ), and positive color scale from red ( $Z = 0.5$ ) to yellow-white ( $Z = 1.5$ ). Regression Eqn:  $SCz_{64} = 0.70 \cdot VCz_{64} + 0.032$ ,  $R^2=0.94$ . (b) Session 42 CONN zMAP for VCz (left) and SCz (right). Negative color scale from dark blue ( $Z=-0.5$ ) to green-white ( $Z=-1.5$ ), and positive color scale from red ( $Z = 0.5$ ) to yellow-white ( $Z = 1.5$ ). Regression Eqn:  $SCz_{42} = 1.29 \cdot VCz_{42} - 0.012$ ,  $R^2 = 0.85$ .

cluded, measures of min and max FWHM show stronger correlation with center and spread measures of FWHM, i.e., median and quartiles (Fig. A1). In addition, with the anomalies excluded, slice ACF measures of FWHM also show stronger correlation with fBIRN QA measures of (volume) FWHM (Fig. A1). These observations confirm that these anomalies are caused by instabilities affecting individual slices rather than by global instabilities affecting entire volumes and/or the entire acquisition.

#### 4.2. Slice anomalies

Our data provide evidence of a potentially significant new source of variance due to anomalous individual slices without subject motion or any sources of physiological noise. These anomalies manifest in abnormally high values of slice ACF FWHM. They can go unnoticed with the fBIRN QA measurements but can be more effectively revealed using slice ACF FWHM measures of imaging resolution. Since these anomalies are caused by instabilities affecting individual slices rather than an entire volume, they are best captured by slice measures of FWHM, rather than volume FWHM where a single FWHM value (in each direction) is calculated for each volume using a volume ACF or by averaging over slices.

We presented a simple anomaly detection criterion based on the slice ACF FWHM, where the minFWHM is used as a reference and anomalies are detected using a multiplicative threshold. We also noted that the number of scan sessions with slice anomalies incrementally decreases

with increasing the threshold. The value of the threshold can be chosen empirically depending on how strictly the user wishes to identify anomalies based on the application.

Slice anomalies are observed more frequently in older scanners. We almost never observed slice anomalies in WEU and UTO scans, which have relatively new Prisma scanners. We speculate these anomalies are related to aging hardware, resulting in spontaneous instabilities in the scanners' acquisition performance.

#### 4.3. Correcting anomalous slices

We noted a preprocessing approach by which these slice anomalies can be corrected under certain circumstances. Since these scanner instabilities affect individual slices (rather than entire volumes) a slice-based censoring scheme is more effective in removing these anomalies than a volume-based censoring scheme. Among the available censoring techniques, SpikeCor provides the option for slice-based censoring. Our results confirm that slice-based SpikeCor more effectively removes the anomalies than the volume-based SpikeCor (Table 5).

For correction to be effective in removing slice anomalies (1) the anomalies should occur in discontinuous time frames or discontinuous blocks of at most a few frames length (hence enabling interpolation), and (2) the censoring method should succeed in correctly identifying the anomalous slices. SpikeCor fails to correctly identify the anomalous slices when they constitute a large fraction of the total number of slices in the volume. This is most prominently evident in several scan sessions



at SBH, UBC, and UCA, where SpikeCor has limited efficacy in detecting anomalous slices within the scans.

It should be noted that different independent metrics are involved in SpikeCor and in ACF anomaly detection. The ACF anomaly detection is based on the physical measurements of ACF FWHM, where anomalies are detected using a threshold with respect to the minimum FWHM value. SpikeCor, on the other hand, uses a statistical threshold for anomaly detection. While one could also use the slice ACF threshold for anomaly detection for censoring, the goal of the current study is not to propose yet another censoring technique, nor is it to compare the efficacy of different censoring techniques. The current study demonstrates slice anomalies on phantom data and provides a proof-of-concept approach to removing them under certain circumstances.

For the next step of testing the impact of slice- versus volume-based censoring in resting state human data we note that it may be difficult to separate and detect slice anomalies of the sort demonstrated here from motion and physiological effects. As a result it is unclear exactly what metric should be used to show improved performance with slice-based censoring in resting state human scans. This is particularly the case since temporal variance itself is considered a cognitive signal in some studies (Garrett et al., 2013) and while presumably it should decrease in cases of extreme anomalies it is unclear if it should generally go up or down after slice-based censoring. We have presented some preliminary human data to illustrate these issues and discuss this further below.

#### 4.4. Hardware and/or software changes

Hardware and/or software changes are reflected in and can be identified by the QA parameters. Our data show an abrupt step change in several QA parameters between SBH0010 and SBH0011, including in FWHM (both fBIRN QA measures and slice ACF measures), percent-Fluc, and std (Figs. 1 and 3). A follow up with the imaging site revealed that the RF coil was replaced in between those scans. QNS updated their scan parameters by using a lower pixel bandwidth (Table A2) and switching from MSENSE to GRAPPA to resolve ghosting issues that were occurring in their images. These changes are reflected as prominent step changes in measures of meanGhost and meanBrightGhost in Fig. 1. This change is also reflected in slice ACF measures of maxFWHM, as a step increase after the change in parameters (Fig. 3). We also observed step changes in ghosting and FWHM parameters for UBC (Fig. 1). An investigation of the scan parameters revealed that these parameters were updated to match those of the CDIP protocol (Table A2), which in turn resulted in step changes in QA parameters of ghosting and FWHM (where these measures improved, i.e., their value decreased, once the scan parameters were set to those of the CDIP protocol). The crop of large anomalies in UCA prior to and including UCA0037 was due to a failing head coil that was eventually replaced after the UCA0037 scan in December 2016. As a result, the UCA scans following UCA0037 have normal FWHM values. Note that the last human scan, session 77, was acquired in August 2016, four months before the failing head coil was replaced. These results indicate that the fBIRN QA parameters and the FWHM measures in particular may provide a sensitive means of monitoring MRI scanners for early signs of head coil failure.

#### 4.5. Translation to human studies

The ability to detect slice anomalies and the efficacy of the preprocessing strategies for controlling them needs to be more comprehensively verified on human data. Our preliminary data show that there is an impact of slice despiking versus volume despiking for PCC seed connectivity. While 46 of the 64 (72%) sessions have slopes less than or equal to 1.0 (Fig. 11), these regression plots of individual session values show both highly reduced and highly increased voxel connectivity values of more than 30%. Therefore, while on average the SC CONNz

map connectivity values are reduced relative to VC CONNz there are individual sessions that have strongly increased SC CONNz map connectivity values. We speculate that the slice censoring may be overcorrecting in the majority of sessions with reduced connectivity values as they may not contain significant slice anomalies, but testing this hypothesis is beyond the scope of this paper. While the group connectivity maps are almost identical (i.e., Fig. 10) our results indicate that slice-versus volume-censoring may significantly change inter-individual differences in connectivity and is therefore an important issue to be further understood as the field focuses more and more on individual effects, and not only group results. The presence of physiological processes and head motion adds substantial variance in human data, which can consequently induce challenges for accurate detection and removal of the anomalous slices and the effects we observe may be the result of interactions between multiple preprocessing steps. Additionally, previous literature has investigated the effect of a number of environmental and physiological variables on within-subject variance in longitudinal studies (Karch et al., 2019). Notably, it has been shown that days since the first scan and time of day are predictors of within-subject variance in structural MRI (Karch et al., 2019; Nakamura et al., 2015). The present human data provide a preliminary investigation of the impact of slice vs volume censoring on human resting state fMRI scans. Our phantom data identify unexpected slice anomalies and provide proof-of-concept preprocessing approaches for controlling these anomalies. We speculate these scanner dependent differences and fluctuations are equally present in human scans. Nevertheless, their relative magnitude and impact compared to other variables affecting in vivo scans remains to be investigated. Subsequent studies need to focus on translating the findings to human scans, taking the effect of the present scanner-related differences and fluctuation into account in relation to subject-specific variables.

#### 4.6. Concluding remarks

In conclusion, the present study provides evidence of slice instabilities in fMRI scans, which can result in unexplained variance in the data. These instabilities can be detected using slice ACF measurements of FWHM and can, under certain circumstances, be controlled by preprocessing. When preprocessing is not effective in controlling these instabilities, the affected scan sessions can be detected using slice ACF FWHM values and excluded in subsequent analyses. Between-manufacturer differences in spatial resolution is another main factor driving variance in multisite studies, which can also be at least partially controlled by preprocessing. The utilization of regular QA scan protocols can reveal problems with the performance of the scanners, some of which may be addressed by hardware and/or software modifications, as was the case for at least two of the participating sites where increasing slice instabilities detected with slice-based FWHM measurements eventually led to head coils being replaced. Furthermore, our phantom results have led to preliminary human studies that have indicated the existence of quite large interindividual connectivity differences as a result of applying the same slice censoring technique in humans that removes some of the slice instabilities detected in phantoms.

#### Authors' note

The opinions, results, and conclusions are those of the authors' and no endorsement by the Ontario Brain Institute is intended or should be inferred.

#### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was conducted with the support of the Ontario Brain Institute, an independent nonprofit corporation, funded partially by the Ontario government. Matching funds were provided by participant hos-

pital and research institute foundations, including the Baycrest Foundation, Bruyère Research Institute, center for Addiction and Mental Health Foundation, London Health Sciences Foundation, McMaster University Faculty of Health Sciences, Ottawa Brain and Mind Research Institute, Queen's University Faculty of Health Sciences, Sunnybrook Health Sciences Foundation, the Thunder Bay Regional Health Sciences center, the University of Ottawa Faculty of Medicine, and the Windsor/Essex County ALS Association. The Temerty Family Foundation provided the major infrastructure matching funds. AK, AC, and MZ were partially supported by a Canadian Institutes of Health Research (CIHR) grant (MOP201403).

### Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: S. Strother is the Chief Scientific Officer of ADMdx, Inc., which receives NIH funding, and he currently has research grants from Brain Canada, Canada Foundation for Innovation (CFI), Canadian Institutes of Health Research (CIHR), and the Ontario Brain Institute in Canada.

**Data and code availability statement:** The data used in the current study are available online at <https://www.braincode.ca/content/open-data-releases>. A dashboard containing interactive figures and links to github repositories containing the code used for the analysis is freely accessible online at <https://kayvanrad.github.io/phantomQA/>.

### Credit authorship contribution statement

**Aras Kayvanrad:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Stephen R. Arnott:** Conceptualization, Validation, Data curtion, Formal analysis, Investigation, Resources, Data curtion, Writing – review & editing, Supervision, Project administration. **Stefanie Hassel:** Investigation, Resources, Data curtion, Writing – review & editing, Supervision. **Aditi Chemparathy:** Methodology, Software, Validation, Formal analysis, Investigation, Resources, Visualization. **Fan Dong:** Software, Resources, Data curtion. **Mojdeh Zamyadi:** Software, Validation. **Tom Gee:** Software, Resources, Data curtion. **Robert Bartha:** Methodology, Resources, Funding acquisition, Investigation. **Sandra E. Black:** Methodology, Resources, Funding acquisition. **Jane M. Lawrence-Dewar:** Investigation, Resources. **Christopher J.M. Scott:** Methodology, Investigation, Resources. **Sean Symons:** Resources, Funding acquisition. **Andrew D. Davis:** Investigation, Writing – review & editing. **Geoffrey B. Hall:** Resources, Funding acquisition. **Jacqueline Harris:** Investigation. **Nancy J. Lobaugh:** Methodology, Investigation. **Glenda MacQueen:** Resources, Supervision, Funding acquisition. **Cindy Woo:** Investigation, Resources. **Stephen Strother:** Conceptualization, Methodology, Validation, Formal analysis, Resources, Data curtion, Writing – review & editing, Supervision, Project administration, Funding acquisition.

### Acknowledgements

We would like to thank the Indoc research team for their data management support. This work was completed on behalf of the Ontario Neurodegenerative Disease Research Initiative (ONDRI) and the Canadian Biomarker Integration Network in Depression (CAN-BIND). The authors would like to acknowledge the ONDRI Founding Authors: Robert Bartha, Sandra E. Black, Michael Borrie, Dale Corbett, Elizabeth Finger, Morris Freedman, Barry Greenberg, David A. Grimes, Robert A. Hegele, Chris Hudson, Anthony E. Lang, Mario Masellis, William E. McIlroy, Paula M. McLaughlin, Manuel Montero-Odasso, David G. Munoz, Douglas P. Munoz, J. B. Orange, Michael J. Strong, Stephen C. Strother, Richard H. Swartz, Sean Symons, Maria Carmela Tartaglia, Angela Troyer, and Lorne Zinman.

**Table A1**

Acquisition parameters for the Canadian Dementia Imaging rs-fMRI protocol (CDIP).

	GE	Philips	Siemens
BO	3T	3T	3T
Model	Discovery	Achieva	Skyra/Trio/Prisma
Sequence name	fMRI EPI	fMRI EPI	fMRI EPI
Imaging options	EDR	GRE EPI, CLEAR SENSE	NA
TE (ms)	30	30	30
TR (ms)	2400	2400	2400
FA (°)	70	70	70
FoV (mm)	224 × 224	224 × 224	224 × 224
Matrix	64 × 64	64 × 64	64 × 64
#slices	41	41	41
Voxel (mm)	3.5 × 3.5 × 3.5	3.5 × 3.5 × 3.5	3.5 × 3.5 × 3.5
Orientation	Axial	Axial	Axial
NEX	1	1	1
Acceleration	2	2	2
BW (Hz/Px)	7812	2441	2440
#acquisitions	250	250	250
Prescan time	00:30	00:30	00:30
Scan time	10:00	10:00	10:00
Total scan time	10:30	10:30	10:30

### Appendix

#### Scan parameters

#### Correlation analysis

The correlation between fBIRN QA and ACF FWHM variables is displayed in Fig. A1(a). fBIRN QA measures of min FWHM are highly correlated with each other (minFWHMX and minFWHMY) and fBIRN measures of max FWHM also highly correlate with each other (maxFWHMX and maxFWHMY). However, fBIRN measures of min FWHM do not show strong correlation with those of max FWHM. Moreover, as expected, SNR and SFNR are highly correlated and so are measures of drift and driftfit. Similar to fBIRN measures, although ACF measures of min FWHM are highly correlated with each other (minFWHMX and minFWHMY) and ACF measures of max FWHM are also highly correlated with each other (maxFWHMX and maxFWHMY), we did not observe strong correlation between ACF measures of min and max FWHM. Moreover, while “center and spread” measures of ACF FWHM (i.e., mean, std, med, Q1, and Q3) are highly correlated, they do not show strong correlation with min and max ACF FWHM measures. This indicates that measures of max and min FWHM are potentially driven by different factors.

The correlation between fBIRN QA measures and slice ACF FWHM measures with the anomalous sessions removed is summarized in Fig. A1(b). After the removal of the anomalous sessions, fBIRN measures of min and max FWHM now show strong correlations with each other. Also, min and max slice ACF FWHM now strongly correlate with each other and with center and spread measures. Furthermore, slice ACF measures of FWHM now show stronger correlation with fBIRN QA measures of FWHM.

#### Scan sessions with slice anomalies

#### Spatial smooting

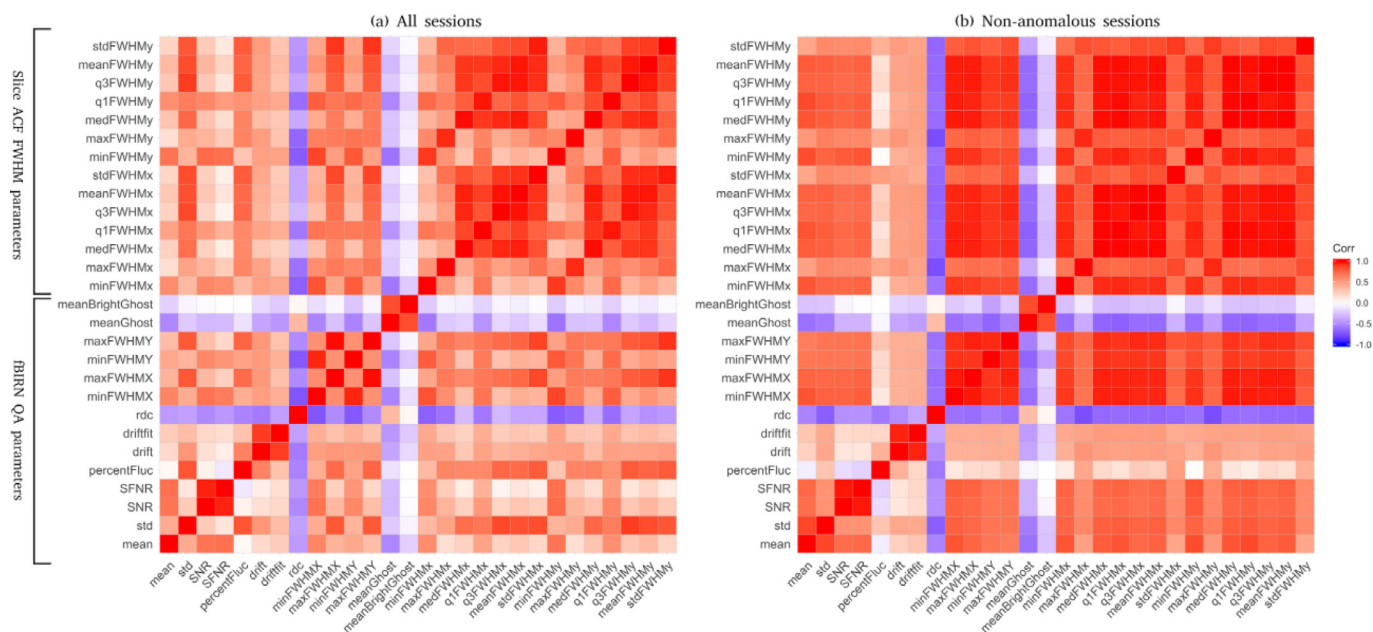
#### Variance within different sites and different manufacturers

#### Table A5.

**Table A2**

Summary of the sites that updated their rs-fMRI scan parameters during the study to match those of the CDIP (Table A1). The echo time (TE), repetition time (TR), flip angle (FA), voxel size, and pixel bandwidth (BW), are shown before (b) and after (a) the update.

	TE(ms)		TR(ms)		FA (°)		Voxel (mm)		BW (Hz/Px)	
	b	a	b	a	b	a	b	a	b	a
<b>CAM</b>	30	30	2000	2400	75	70	4.0 × 4.0	3.5 × 3.5	7812	7812
<b>QNS</b>	25	30	2000	2400	75	70	4.0 × 4.0	3.5 × 3.5	2232	2441
<b>TBR</b>	30	30	2400	2400	70	70	3.5 × 3.5	3.5 × 3.5	3241	3125
<b>TWH</b>	30	30	2000	2400	75	70	4.0 × 4.0	3.5 × 3.5	7812	7812
<b>UBC</b>	30	30	2000	2400	90	70	4.0 × 4.0	3.5 × 3.5	4783	3127
<b>UCA</b>	30	30	2000	2400	75	70	4.0 × 4.0	3.5 × 3.5	7812	7812



**Fig. A1.** Correlation between fBIRN QA and ACF FWHM variables: (a) all sessions (left), (b) anomalous sessions excluded (right).

**Table A3**

Summary of anomalous sessions within each imaging site detected using a conservative anomaly threshold of  $\theta = 10$ : (a) unprocessed data, (b) subsequent to processing with slice-based SpikeCor, and (c) subsequent to processing with volume-based SpikeCor.

Site	Sessions	(a) Unprocessed		(b) Slice SpikeCor		(c) Volume SpikeCor	
		Anomalies	Anomalous sessions	Anomalies	Anomalous sessions	Anomalies	Anomalous sessions
BYC	141	3	BYC01_PHA_FBN1394_0014 BYC01_PHA_FBN1394_0046 BYC01_PHA_FBN1394_0098	0		1	BYC01_PHA_FBN1394_0046
CAM	48	4	CAM01_PHA_FBN1299_0001 CAM01_PHA_FBN1299_0002 CAM01_PHA_FBN1392_0000 CAM01_PHA_FBN1392_0013	2	CAM01_PHA_FBN1299_0002 CAM01_PHA_FBN1392_0013	3	CAM01_PHA_FBN1299_0001 CAM01_PHA_FBN1299_0002 CAM01_PHA_FBN1392_0013
MCM	47	1	MCM01_PHA_FBN1307_0032	1	MCM01_PHA_FBN1307_0032	1	MCM01_PHA_FBN1307_0032
QNS	46	1	QNS01_PHA_FBN1305_9025	0		1	QNS01_PHA_FBN1305_9025
SBH	34	10	SBH01_PHA_FBN1381_0001 SBH01_PHA_FBN1381_0002 SBH01_PHA_FBN1381_0003 SBH01_PHA_FBN1381_0004 SBH01_PHA_FBN1381_0005 SBH01_PHA_FBN1381_0006 SBH01_PHA_FBN1381_0007 SBH01_PHA_FBN1381_0008 SBH01_PHA_FBN1381_0009 SBH01_PHA_FBN1381_0010	10	SBH01_PHA_FBN1381_0001 SBH01_PHA_FBN1381_0002 SBH01_PHA_FBN1381_0003 SBH01_PHA_FBN1381_0004 SBH01_PHA_FBN1381_0005 SBH01_PHA_FBN1381_0006 SBH01_PHA_FBN1381_0007 SBH01_PHA_FBN1381_0008 SBH01_PHA_FBN1381_0009 SBH01_PHA_FBN1381_0010	10	SBH01_PHA_FBN1381_0001 SBH01_PHA_FBN1381_0002 SBH01_PHA_FBN1381_0003 SBH01_PHA_FBN1381_0004 SBH01_PHA_FBN1381_0005 SBH01_PHA_FBN1381_0006 SBH01_PHA_FBN1381_0007 SBH01_PHA_FBN1381_0008 SBH01_PHA_FBN1381_0009 SBH01_PHA_FBN1381_0010
TBR	30	24	TBR01_PHA_FBN1319_0001 TBR01_PHA_FBN1319_0002 TBR01_PHA_FBN1319_0005 TBR01_PHA_FBN1319_0006 TBR01_PHA_FBN1319_0007 TBR01_PHA_FBN1319_0008 TBR01_PHA_FBN1319_0009	3	TBR01_PHA_FBN1319_0020 TBR01_PHA_FBN1319_0024 TBR01_PHA_FBN1319_0029	22	TBR01_PHA_FBN1319_0001 TBR01_PHA_FBN1319_0005 TBR01_PHA_FBN1319_0006 TBR01_PHA_FBN1319_0007 TBR01_PHA_FBN1319_0008 TBR01_PHA_FBN1319_0009

(continued on next page)

Table A3 (continued)

Site	Sessions	(a) Unprocessed		(b) Slice SpikeCor		(c) Volume SpikeCor	
		Anomalies	Anomalous sessions	Anomalies	Anomalous sessions	Anomalies	Anomalous sessions
			TBR01_PHA_FBN1319_0010 TBR01_PHA_FBN1319_0011 TBR01_PHA_FBN1319_0015 TBR01_PHA_FBN1319_0016 TBR01_PHA_FBN1319_0017 TBR01_PHA_FBN1319_0018 TBR01_PHA_FBN1319_0019 TBR01_PHA_FBN1319_0020 TBR01_PHA_FBN1319_0021 TBR01_PHA_FBN1319_0023 TBR01_PHA_FBN1319_0024 TBR01_PHA_FBN1319_0025 TBR01_PHA_FBN1319_0026 TBR01_PHA_FBN1319_0027 TBR01_PHA_FBN1319_0028 TBR01_PHA_FBN1319_0029 TBR01_PHA_FBN1319_0030				TBR01_PHA_FBN1319_0010 TBR01_PHA_FBN1319_0011 TBR01_PHA_FBN1319_0015 TBR01_PHA_FBN1319_0016 TBR01_PHA_FBN1319_0017 TBR01_PHA_FBN1319_0018 TBR01_PHA_FBN1319_0019 TBR01_PHA_FBN1319_0021 TBR01_PHA_FBN1319_0023 TBR01_PHA_FBN1319_0024 TBR01_PHA_FBN1319_0025 TBR01_PHA_FBN1319_0026 TBR01_PHA_FBN1319_0027 TBR01_PHA_FBN1319_0028 TBR01_PHA_FBN1319_0029 TBR01_PHA_FBN1319_0030
SMH	48	2	SMH01_PHA_FBN1395_0020 SMH01_PHA_FBN1395_0023	1	SMH01_PHA_FBN1395_0020	1	SMH01_PHA_FBN1395_0020
TOH	33	1	TOH01_PHA_FBN1384_0039	0		0	
TWH	40	7	TWH01_PHA_FBN1304_0001 TWH01_PHA_FBN1304_0002 TWH01_PHA_FBN1304_0003 TWH01_PHA_FBN1304_0004 TWH01_PHA_FBN1304_0010 TWH01_PHA_FBN1304_0024 TWH01_PHA_FBN1304_0031	5	TWH01_PHA_FBN1304_0001 TWH01_PHA_FBN1304_0002 TWH01_PHA_FBN1304_0003 TWH01_PHA_FBN1304_0004 TWH01_PHA_FBN1304_0031	7	TWH01_PHA_FBN1304_0001 TWH01_PHA_FBN1304_0002 TWH01_PHA_FBN1304_0003 TWH01_PHA_FBN1304_0004 TWH01_PHA_FBN1304_0014 TWH01_PHA_FBN1304_0024 TWH01_PHA_FBN1304_0031
UBC	38	34	UBC01_PHA_FBN0001_0008 UBC01_PHA_FBN0001_0009 UBC01_PHA_FBN0001_0010 UBC01_PHA_FBN0001_0012 UBC01_PHA_FBN0001_0013 UBC01_PHA_FBN0001_0014 UBC01_PHA_FBN0001_0015 UBC01_PHA_FBN0001_0016 UBC01_PHA_FBN0001_0017 UBC01_PHA_FBN0001_0018 UBC01_PHA_FBN0001_0020 UBC01_PHA_FBN0001_0021 UBC01_PHA_FBN0001_0022 UBC01_PHA_FBN0001_0023 UBC01_PHA_FBN0001_0024 UBC01_PHA_FBN0001_0025 UBC01_PHA_FBN0001_0026 UBC01_PHA_FBN0001_0027 UBC01_PHA_FBN0001_0028 UBC01_PHA_FBN0001_0029 UBC01_PHA_FBN0001_0030 UBC01_PHA_FBN0001_0031 UBC01_PHA_FBN0001_0032 UBC01_PHA_FBN0001_0033 UBC01_PHA_FBN0001_0034 UBC01_PHA_FBN0001_0036 UBC01_PHA_FBN0001_0037 UBC01_PHA_FBN0001_0038 UBC01_PHA_FBN0001_0039 UBC01_PHA_FBN0001_0041 UBC01_PHA_FBN0001_0042 UBC01_PHA_FBN0001_0043 UBC01_PHA_FBN0001_0044 UBC01_PHA_FBN0001_0045	28	UBC01_PHA_FBN0001_0008 UBC01_PHA_FBN0001_0009 UBC01_PHA_FBN0001_0010 UBC01_PHA_FBN0001_0012 UBC01_PHA_FBN0001_0013 UBC01_PHA_FBN0001_0014 UBC01_PHA_FBN0001_0015 UBC01_PHA_FBN0001_0016 UBC01_PHA_FBN0001_0017 UBC01_PHA_FBN0001_0021 UBC01_PHA_FBN0001_0022 UBC01_PHA_FBN0001_0023 UBC01_PHA_FBN0001_0025 UBC01_PHA_FBN0001_0026 UBC01_PHA_FBN0001_0027 UBC01_PHA_FBN0001_0028 UBC01_PHA_FBN0001_0029 UBC01_PHA_FBN0001_0030 UBC01_PHA_FBN0001_0031 UBC01_PHA_FBN0001_0032 UBC01_PHA_FBN0001_0033 UBC01_PHA_FBN0001_0034 UBC01_PHA_FBN0001_0043 UBC01_PHA_FBN0001_0044 UBC01_PHA_FBN0001_0045	35	UBC01_PHA_FBN0001_0008 UBC01_PHA_FBN0001_0009 UBC01_PHA_FBN0001_0010 UBC01_PHA_FBN0001_0012 UBC01_PHA_FBN0001_0013 UBC01_PHA_FBN0001_0014 UBC01_PHA_FBN0001_0015 UBC01_PHA_FBN0001_0016 UBC01_PHA_FBN0001_0017 UBC01_PHA_FBN0001_0018 UBC01_PHA_FBN0001_0020 UBC01_PHA_FBN0001_0021 UBC01_PHA_FBN0001_0022 UBC01_PHA_FBN0001_0023 UBC01_PHA_FBN0001_0025 UBC01_PHA_FBN0001_0026 UBC01_PHA_FBN0001_0027 UBC01_PHA_FBN0001_0028 UBC01_PHA_FBN0001_0029 UBC01_PHA_FBN0001_0030 UBC01_PHA_FBN0001_0031 UBC01_PHA_FBN0001_0032 UBC01_PHA_FBN0001_0033 UBC01_PHA_FBN0001_0034 UBC01_PHA_FBN0001_0035 UBC01_PHA_FBN0001_0036 UBC01_PHA_FBN0001_0037 UBC01_PHA_FBN0001_0038 UBC01_PHA_FBN0001_0039 UBC01_PHA_FBN0001_0041 UBC01_PHA_FBN0001_0042 UBC01_PHA_FBN0001_0043 UBC01_PHA_FBN0001_0044 UBC01_PHA_FBN0001_0045
UCA	47	7	UCA01_PHA_FBN1245_0009 UCA01_PHA_FBN1245_0020 UCA01_PHA_FBN1245_0030 UCA01_PHA_FBN1245_0031 UCA01_PHA_FBN1245_0035 UCA01_PHA_FBN1245_0036 UCA01_PHA_FBN1245_0037	7	UCA01_PHA_FBN1245_0009 UCA01_PHA_FBN1245_0020 UCA01_PHA_FBN1245_0030 UCA01_PHA_FBN1245_0031 UCA01_PHA_FBN1245_0035 UCA01_PHA_FBN1245_0036 UCA01_PHA_FBN1245_0037	7	UCA01_PHA_FBN1245_0009 UCA01_PHA_FBN1245_0020 UCA01_PHA_FBN1245_0030 UCA01_PHA_FBN1245_0031 UCA01_PHA_FBN1245_0035 UCA01_PHA_FBN1245_0036 UCA01_PHA_FBN1245_0037
UTO	36	0		0		0	
WEU	41	0		0		0	

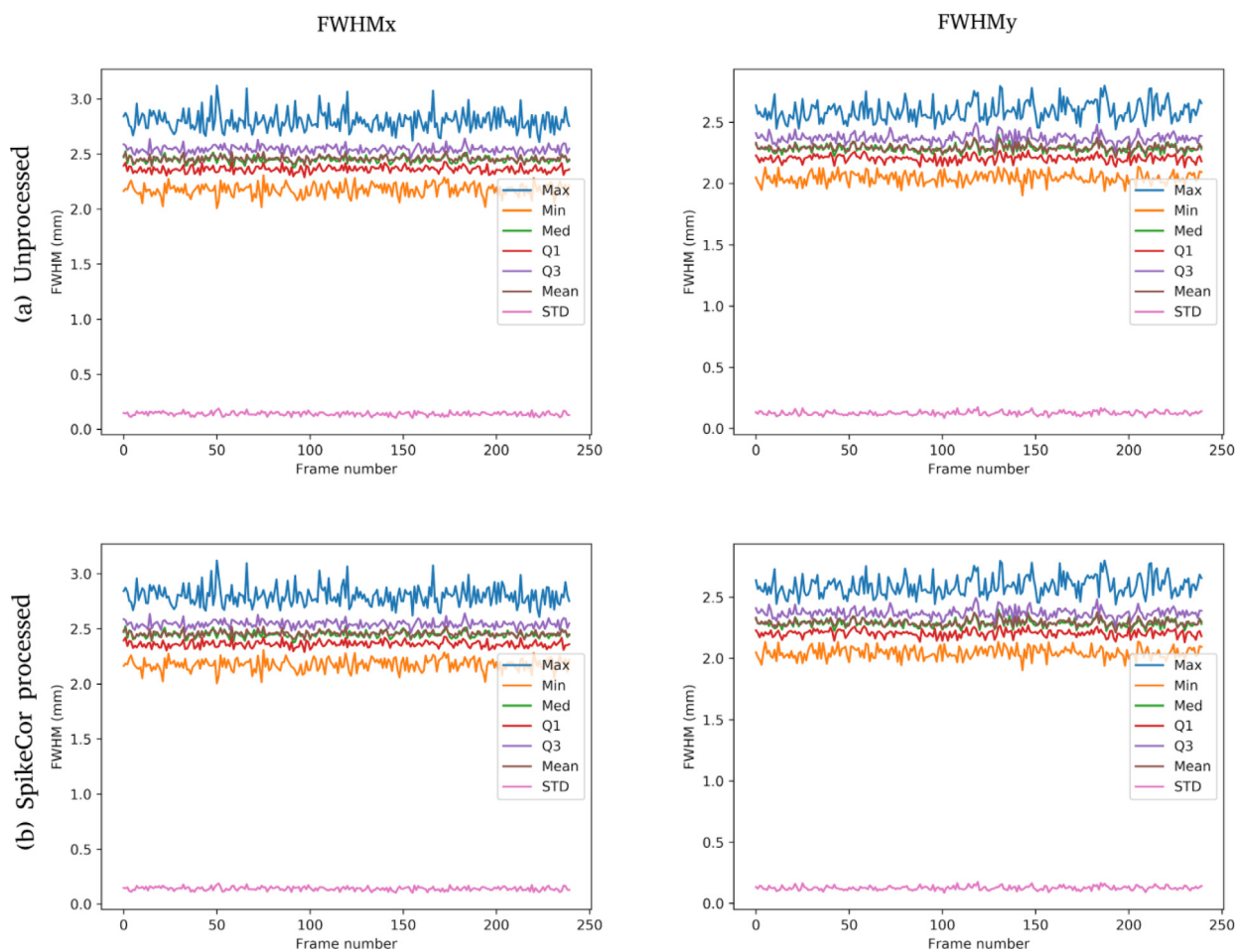
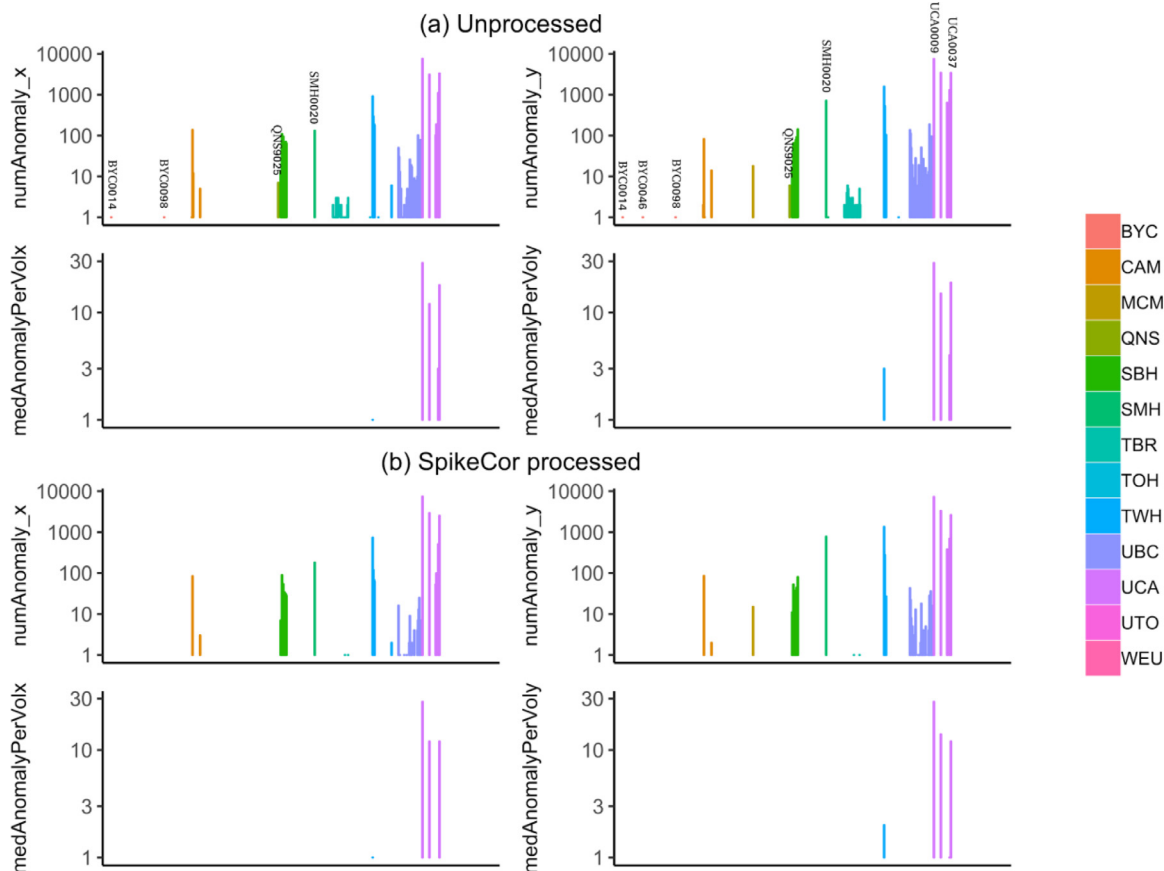


Fig. A2. Summary of slice ACF measures of FWHM for SMH0019 (normal session). Plotted are max, min, median, first quartile (Q1), third quartile (Q3), mean, and standard deviation (STD) values of ACF FWHM for each volume in the time series in the x and y directions.

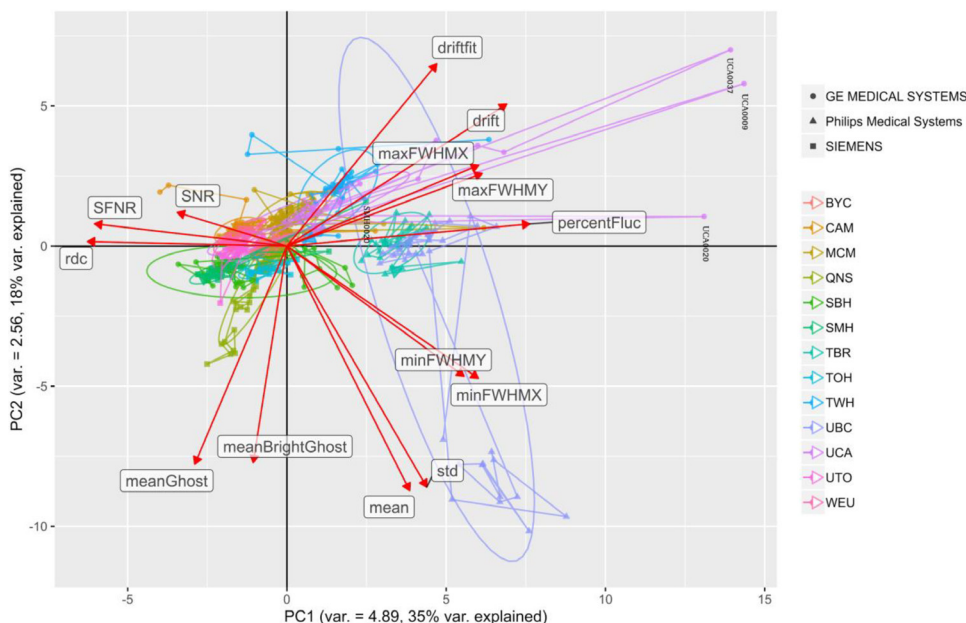
Table A4

Variance in the first principal component (PC1) within different sites and different manufacturers. PCA was performed on fBIRN QA data of all scan sessions (including anomalous sessions).

Manufacturer	PC1 Variance	Site	PC1 Variance
GE	4.8	CAM	0.76
		MCM	0.69
		SBH	0.98
		TWH	1.8
		UCA	14
Siemens	0.27	BYC	0.066
		WEU	0.068
		QNS	0.59
		TOH	0.057
		UTO	0.14
Philips	0.56	SMH	0.34
		TBR	0.18
		UBC	0.34



**Fig. A3.** Prevalence of slice anomalies per session with the anomaly threshold conservatively set to  $\theta = 10$  for (a) unprocessed data, and (b) SpikeCor processed data. For each scan session, total number of anomalies and the median of the number of anomalies per volume is shown. The first column shows anomalous FWHM values in the x direction and the second one shows those values in the y direction.



**Fig. A4.** Principal component analysis (PCA) of the fBIRN QA variables after spatial smoothing. Shown is the scatter plot of the first two principal components (PC1 and PC2) for all the imaging sessions. Each imaging session is color-coded by scan site and shaped-coded by scanner manufacturer. The imaging sessions within each site are connected in temporal order of acquisition. The ellipses show 95% normal probabilities. The original fBIRN QA variables are also plotted using red arrows on the PCA plane with lengths proportional to the variable loadings. An interactive version of this plot can be viewed at [https://kayvanrad.github.io/phantomQA/#spatially\\_smoothed\\_pca](https://kayvanrad.github.io/phantomQA/#spatially_smoothed_pca).

**Table A5**

Variance in the first principal component (PC1) within different sites and different manufacturers. PCA was performed on fBIRN QA data with anomalous sessions excluded using a conservative anomaly threshold of  $\theta = 10$ .

Manufacturer	PC1 Variance	Site	PC1 Variance
GE	1.0	CAM	0.12
		MCM	0.53
		SBH	0.34
		TWH	0.22
		UCA	0.76
Siemens	0.33	BYC	0.078
		WEU	0.058
		QNS	0.91
		TOH	0.097
		UTO	0.15
		SMH	0.10
Philips	1.3	TBR	0.68
		UBC	0.73

**Table A6**

Mean and 95%-confidence intervals for the slice ACF minFWHM values for all the imaging sessions at each site (including anomalous sessions).

Manufacturer	Site	(a) unprocessed data		(b) spatially smoothed data	
		minFWHMx (95%-CI)	minFWHM <sub>y</sub> (95%-CI)	minFWHMx (95%-CI)	minFWHM <sub>y</sub> (95%-CI)
GE	CAM	2.74 (2.65,2.83)	2.99 (2.91,3.08)	7.74 (7.29,8.19)	7.76 (7.44,8.08)
	MCM	2.65 (2.59,2.71)	3.08 (3.00,3.16)	7.98 (7.77,8.19)	7.99 (7.86,8.13)
	SBH	2.82 (2.77,2.88)	3.35 (3.29,3.41)	8.14 (7.99,8.3)	8.06 (7.88,8.23)
	TWH	2.81 (2.69,2.92)	3.36 (3.27,3.45)	6.35 (5.90,6.80)	7.17 (6.97,7.38)
	UCA	2.81 (2.76,2.85)	2.99 (2.91,3.06)	7.84 (7.59,8.09)	7.84 (7.60,8.07)
Siemens	BYC	1.95 (1.94,1.95)	1.88 (1.87,1.88)	8.25 (8.22,8.28)	8.28 (8.17,8.39)
	WEU	2.00 (1.98,2.01)	1.90 (1.88,1.91)	8.17 (8.11,8.22)	8.22 (8.15,8.28)
	QNS	1.97 (1.95,2.00)	1.91 (1.88,1.93)	7.79 (7.49,8.08)	8.11 (7.91,8.31)
	TOH	2.00 (1.99,2.02)	1.91 (1.90,1.92)	8.54 (8.51,8.58)	8.58 (8.53,8.63)
	UTO	1.99 (1.98,2.01)	1.89 (1.87,1.90)	8.01 (7.95,8.07)	8.01 (7.89,8.12)
	SMH	1.97 (1.94,2.00)	1.91 (1.09,1.92)	8.46 (8.38,8.54)	8.62 (8.57,8.67)
Philips	TBR	3.02 (2.96,3.09)	3.12 (3.01,3.23)	7.44 (7.29,7.60)	7.66 (7.55,7.77)
	UBC	3.14 (3.01,3.27)	3.25 (3.11,3.39)	7.21 (7.08,7.34)	7.18 (7.02,7.34)

## References

- Caballero-Gaudes, C., Reynolds, R.C., 2017. Methods for cleaning the BOLD fMRI signal. *NeuroImage* 154, 128–149. doi:10.1016/j.neuroimage.2016.12.018.
- Campbell, K.L., Grigg, O., Saverino, C., Churchill, N., Grady, C.L., 2013. Age differences in the intrinsic functional connectivity of default network subsystems. *Front. Aging Neurosci.* 5. doi:10.3389/fnagi.2013.00073.
- Carbonell, F., Bellec, P., Shmuel, A., 2011. Global and system-specific resting-state fMRI fluctuations are uncorrelated: principal component analysis reveals anti-correlated networks. *Brain Connect.* 1, 496–510. doi:10.1089/brain.2011.0065.
- Churchill, N.W., Oder, A., Abdi, H., Tam, F., Lee, W., Thomas, C., Ween, J.E., Graham, S.J., Strother, S.C., 2012a. Optimizing Preprocessing and Analysis Pipelines for Single-Subject fMRI. I. Standard Temporal Motion and Physiological Noise Correction Methods. *Hum. Brain Mapp* 33, 609–627. doi:10.1002/hbm.21238.
- Churchill, N.W., Raamana, P., Spring, R., Strother, S.C., 2017. Optimizing fMRI preprocessing pipelines for block-design tasks as a function of age. *NeuroImage*, Cleaning up the fMRI time series: mitigating noise with advanced acquisition and correction strategies 154, 240–254. doi:10.1016/j.neuroimage.2017.02.028.
- Churchill, N.W., Spring, R., Afshin-Pour, B., Dong, F., Strother, S.C., 2015. An Automated, Adaptive Framework for Optimizing Preprocessing Pipelines in Task-Based Functional MRI. *PLoS ONE* 10, e0131520. doi:10.1371/journal.pone.0131520.
- Churchill, N.W., Strother, S.C., 2013. PHYCAA+: an optimized, adaptive procedure for measuring and controlling physiological noise in BOLD fMRI. *NeuroImage* 82, 306–325. doi:10.1016/j.neuroimage.2013.05.102.
- Churchill, N.W., Yourganov, G., Oder, A., Tam, F., Graham, S.J., Strother, S.C., 2012b. Optimizing preprocessing and analysis pipelines for single-subject fMRI: 2. Interactions with ICA, PCA, task contrast and inter-subject heterogeneity. *PLoS ONE* 7, e31147. doi:10.1371/journal.pone.0031147.
- Duchesne, S., Chouinard, I., Potvin, O., Fonov, V.S., Khademi, A., Bartha, R., Bellec, P., Collins, D.L., Descoteaux, M., Hoge, R., McCreary, C.R., Ramirez, J., Scott, C.J.M., Smith, E.E., Strother, S.C., Black, S.E., 2019. The Canadian Dementia Imaging Protocol: harmonizing National Cohorts. *J. Magn. Reson. Imaging* 49, 456–465. doi:10.1002/jmri.26197.
- Farhan, S.M.K., Bartha, R., Black, S.E., Corbett, D., Finger, E., Freedman, M., Greenberg, B., Grimes, D.A., Hegele, R.A., Hudson, C., Kleinstiver, P.W., Lang, A.E., Masellis, M., McIlroy, W.E., McLaughlin, P.M., Montero-Odasso, M., Munoz, D.G., Munoz, D.P., Strother, S., Swartz, R.H., Symons, S., Tartaglia, M.C., Zinman, L., Investigators, O., Strong, M.J., 2017. The Ontario Neurodegenerative Disease Research Initiative (ON-DRI). *Can. J. Neurol. Sci.* 44, 196–202. doi:10.1017/cjn.2016.415.
- Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.* 33, 636–647. doi:10.1002/mrm.1910330508.
- Friedman, L., Glover, G.H., 2006. Report on a multicenter fMRI quality assurance protocol. *J. Magn. Reson. Imaging* 23, 827–839. doi:10.1002/jmri.20583.
- Friedman, L., Glover, G.H., Krenz, D., Magnotta, V., FIRST BIRN, 2006. Reducing inter-scanner variability of activation in a multicenter fMRI study: role of smoothness equalization. *NeuroImage* 32, 1656–1668. doi:10.1016/j.neuroimage.2006.03.062.
- Garrett, D.D., Samanez-Larkin, G.R., MacDonald, S.W.S., Lindenberger, U., McIntosh, A.R., Grady, C.L., 2013. Moment-to-moment brain signal variability: a next frontier in human brain mapping? *Neurosci. Biobehav. Rev.* 37, 610–624. doi:10.1016/j.neubiorev.2013.02.015.
- Glover, G.H., Mueller, B.A., Turner, J.A., van Erp, T.G.M., Liu, T.T., Greve, D.N., Voyvodic, J.T., Rasmussen, J., Brown, G.G., Keator, D.B., Calhoun, V.D., Lee, H.J., Ford, J.M., Mathalon, D.H., Diaz, M., O'Leary, D.S., Gadge, S., Preda, A., Lim, K.O., Wible, C.G., Stern, H.S., Belger, A., McCarthy, G., Ozyurt, B., Potkin, S.G., 2012. Function biomedical informatics research network recommendations for prospective multi-center functional magnetic resonance imaging studies. *J. Magn. Reson. Imaging* 36, 39–54. doi:10.1002/jmri.23572.
- Griswold, M.A., Jakob, P.M., Heidemann, R.M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., Haase, A., 2002. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn. Reson. Med.* 47, 1202–1210. doi:10.1002/mrm.10171.
- Haase, A., Frahm, J., Matthaei, D., Hancic, W., Merboldt, K.-D., 1986. FLASH imaging. Rapid NMR imaging using low flip-angle pulses. *J. Magn. Reson.* 67, 258–266. doi:10.1016/0022-2364(86)90433-6.
- Jovicich, J., Minati, L., Marizzoni, M., Marchitelli, R., Sala-Llonch, R., Bartrés-Faz, D., Arnold, J., Benninghoff, J., Fiedler, U., Roccatagliata, L., Picco, A., Nobili, F., Blin, O., Bombois, S., Lopes, R., Bordet, R., Sein, J., Ranjeva, J.-P., Didic, M., Gros-Dagnac, H., Payoux, P., Zoccatelli, G., Alessandrini, F., Beltramello, A., Bargalló, N., Ferretti, A., Caulo, M., Aiello, M., Cavaliere, C., Soricelli, A., Parnetti, L., Tarducci, R., Floridi, P., Tsolaki, M., Constantinidis, M., Drevelegas, A., Rossini, P.M., Marra, C., Schönknecht, P., Hensch, T., Hoffmann, K.-T., Kuijter, J.P., Visser, P.J., Barkhof, F., Frisoni, G.B., 2016. Longitudinal reproducibility of default-mode network connectivity in healthy elderly participants: a multicentric resting-state fMRI study. *NeuroImage* 124, 442–454. doi:10.1016/j.neuroimage.2015.07.010.
- Karch, J.D., Filevich, E., Wenger, E., Lisofsky, N., Becker, M., Butler, O., Mårtensson, J., Lindenberger, U., Brandmaier, A.M., Kühn, S., 2019. Identifying predictors of within-person variance in MRI-based brain volume estimates. *NeuroImage* 200, 575–589. doi:10.1016/j.neuroimage.2019.05.030.
- Lam, R.W., Milev, R., Rotzinger, S., Andrezza, A.C., Blier, P., Brenner, C., Daskalakis, Z.J., Dharsee, M., Downar, J., Evans, K.R., Farzan, F., Foster, J.A., Frey, B.N., Geraci, J., Giacobbe, P., Feiltoer, H.E., Hall, G.B., Harkness, K.L., Hassel, S., Ismail, Z., Leri, F., Liotti, M., MacQueen, G.M., McAndrews, M.P., Minuzzi, L., Müller, D.J., Parikh, S.V., Placenza, F.M., Quilty, L.C., Ravindran, A.V., Salomons, T.V., Soares, C.N., Strother, S.C., Turecki, G., Vaccarino, A.L., Vila-Rodriguez, F., Kennedy, S.H. on behalf of the CAN-BIND Investigator Team, 2019. Discovering biomarkers for antidepressant response: protocol from the Canadian biomarker integration network in depression (CAN-BIND) and clinical characteristics of the first patient cohort. *BMC Psychiatry* 16, 105. doi:10.1186/s12888-016-0785-x.
- Larkman, D.J., Nunes, R.G., 2007. Parallel magnetic resonance imaging. *Phys. Med. Biol.* 52, R15. doi:10.1088/0031-9155/52/7/R01.
- Lustig, M., Donoho, D., Pauly, J.M., 2007. Sparse MRI: the application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.* 58, 1182–1195. doi:10.1002/mrm.21391.
- Mansfield, P., 1977. Multi-planar image formation using NMR spin echoes. *J. Phys. C* 10, L55. doi:10.1088/0022-3719/10/3/004.
- Margosian, P., Schmitt, F., Purdy, D., 1986. Faster MR imaging: imaging with half the data. *Health Care Instrum.* 1, 195.
- Nakamura, K., Brown, R.A., Narayanan, S., Collins, D.L., Arnold, D.L., 2015. Diurnal fluctuations in brain volume: statistical analyses of MRI from large populations. *NeuroImage* 118, 126–132. doi:10.1016/j.neuroimage.2015.05.077.
- Power, J.D., Schlaggar, B.L., Petersen, S.E., 2015. Recent progress and outstanding issues in motion correction in resting state fMRI. *NeuroImage* 105, 536–551. doi:10.1016/j.neuroimage.2014.10.044.
- Price, R.R., Axel, L., Morgan, T., Newman, R., Perman, W., Schneiders, N., Selikson, M., Wood, M., Thomas, S.R., 1990. Quality assurance methods and phantoms for magnetic resonance imaging: report of AAPM nuclear magnetic resonance Task Group No. 1. *Med. Phys.* 17, 287–295. doi:10.1118/1.596566.
- Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P., 1999. SENSE: sensitivity encoding for fast MRI. *Magn. Reson. Med.* 42, 952–962 https://doi.org/10.1002/(SICI)1522-2594(199911)42:5<952::AID-MRM16>3.0.CO;2-S.
- Samsonov, A.A., Kholmovski, E.G., Parker, D.L., Johnson, C.R., 2004. POCSSENSE: POCS-based reconstruction for sensitivity encoded magnetic resonance imaging. *Magn. Reson. Med.* 52, 1397–1406. doi:10.1002/mrm.20285.
- Sodickson, D.K., Manning, W.J., 1997. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. *Magn. Reson. Med.* 38, 591–603. doi:10.1002/mrm.1910380414.
- Strother, S.C., 2006. Evaluating fMRI preprocessing pipelines. *IEEE Eng. Med. Biol. Mag.* 25, 27–41. doi:10.1109/EMEM.2006.1607667.
- Strother, S.C., Anderson, J., Hansen, L.K., Kjems, U., Kustra, R., Sidtis, J., Frutiger, S., Muley, S., LaConte, S., Rottenberg, D., 2002. The quantitative evaluation of functional neuroimaging experiments: the NPAIRS data analysis framework. *NeuroImage* 15, 747–771. doi:10.1006/nimg.2001.1034.
- Wijk, G. van der, Harris, J.K., Hassel, S., Davis, A.D., Zamyadi, M., Arnott, S.R., Milev, R., Lam, R.W., Frey, B.N., Hall, G.B., Müller, D.J., Rotzinger, S., Kennedy, S.H., Strother, S.C., MacQueen, G.M., Protzner, A.B., 2021. Baseline functional connectivity in resting state networks associated with depression and remission status after 16 weeks of pharmacotherapy: A CAN-BIND Report. *medRxiv* 2021.01.27.21250490. https://doi.org/10.1101/2021.01.27.21250490
- Yan, C.-G., Craddock, R.C., Zuo, X.-N., Zang, Y.-F., Milham, M.P., 2013. Standardizing the intrinsic brain: towards robust measurement of inter-individual variation in 1000 functional connectomes. *NeuroImage* 80, 246–262. doi:10.1016/j.neuroimage.2013.04.081.