Received: 3 December 2022 | Revised: 6 February 2023 | Accepted: 11 February 2023

DOI: 10.1002/hbm.26244

RESEARCH ARTICLE

WILEY

Reproducibility of cerebral perfusion measurements using BOLD delay

Ahmed A. Khalil[1,2,3,4] | Ayse C. Tanritanir[1] | Ulrike Grittner[4,5] | Evgeniya Kirilina[6,7] | Arno Villringer[2,3] | Jochen B. Fiebach[1] | Ralf Mekle[1]

[1]Center for Stroke Research Berlin, Charité - Universitätsmedizin Berlin, Berlin, Germany

[2]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany

[3]Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[4]Berlin Institute of Health (BIH), Berlin, Germany

[5]Charité - Universitätsmedizin Berlin, Institute of Biometry and Clinical Epidemiology, Berlin, Germany

[6]Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[7]Center for Cognitive Neuroscience Berlin, Free University, Berlin, Germany

Correspondence
Ahmed A. Khalil, Center for Stroke Research Berlin, Charité Campus Benjamin Franklin, Hindenburgdamm 30, Berlin 12200, Germany.
Email: ahmed-abdelrahim.khalil@charite.de

Funding information
Berlin Institute of Health; Bundesministerium für Bildung und Forschung, Grant/Award Numbers: 01EO0801, 01EO01301

Abstract

BOLD delay is an emerging, noninvasive method for assessing cerebral perfusion that does not require the use of intravenous contrast agents and is thus particularly suited for longitudinal monitoring. In this study, we assess the reproducibility of BOLD delay using data from 136 subjects with normal cerebral perfusion scanned on two separate occasions with scanners, sequence parameters, and intervals between scans varying between subjects. The effects of various factors on the reproducibility of BOLD delay, defined here as the differences in BOLD delay values between the scanning sessions, were investigated using a linear mixed model. Reproducibility was additionally assessed using the intraclass correlation coefficient of BOLD delay between sessions. Reproducibility was highest in the posterior cerebral artery territory. The mean BOLD delay test–retest difference after accounting for the aforementioned factors was 1.2 s (95% CI = 1.0 to 1.4 s). Overall, BOLD delay shows good reproducibility, but care should be taken when interpreting longitudinal BOLD delay changes that are either very small or are located in certain brain regions.

KEYWORDS
blood flow, BOLD delay, perfusion, reproducibility, resting-state functional MRI

1 | INTRODUCTION

Imaging brain perfusion plays an important role in investigating the pathophysiology of cerebrovascular disorders and guiding patient management. This is typically achieved using magnetic resonance imaging (MRI) sequences that are enhanced with contrast agents such as gadolinium. Recent evidence suggests that these agents accumulate within the brain with repeated administration (Gulani et al., 2017). Despite the lack of conclusive evidence that this causes harm, these findings have led to recommendations restricting the use of these agents, particularly their repeated use in the same patient (European Medicines Agency, 2018).

Blood-oxygenation-level-dependent (BOLD) signal delay is a relatively new method that allows the assessment of brain perfusion without the need to administer intravenous contrast agents. It relies on measuring delays in the systemic low-frequency oscillations (sLFOs) of the BOLD signal in different brain regions compared to a reference signal. These BOLD signal delays are closely related to bolus arrival times derived from contrast-enhanced perfusion imaging techniques (Amemiya et al., 2014, 2016; Christen et al., 2015; Khalil et al., 2017;

Lv et al., 2013; Tong et al., 2017). Perfusion has been measured using BOLD delay in healthy individuals (Tong et al., 2017; Tong, Yao, et al., 2019) and patients with cerebrovascular disorders (Amemiya et al., 2014; Q. Chen et al., 2019; Christen et al., 2015; Khalil et al., 2017, 2020; Lv et al., 2013; Ni et al., 2017; Wu et al., 2017), Alzheimer's disease (Yan et al., 2018), epilepsy (Shah et al., 2018), and sickle cell disease (Coloigner et al., 2016). Several of these studies have validated this technique against the clinical reference standard, dynamic susceptibility contrast MRI (Amemiya et al., 2014; Q. Chen et al., 2019; Khalil et al., 2017; Lv et al., 2013; Ni et al., 2017; Tong et al., 2017), as well as noninvasive alternatives such as arterial spin labeling (Christen et al., 2015; Lv et al., 2018; Siegel et al., 2016).

Because it does not require exogenous contrast agents, one important potential use of BOLD delay is for convenient, minimally invasive repeated measurements of patients over time. One example is in acute stroke, where perfusion imaging is used to monitor disease progression and treatment outcome (Schellinger et al., 2000). Recently, studies have shown that BOLD delay detects therapeutically relevant changes in blood flow over time in stroke patients (Q. Chen et al., 2019; Khalil et al., 2020), emphasizing the potential use of BOLD delay for longitudinal imaging in acute stroke.

To further examine the value of BOLD delay in the longitudinal follow-up of patients, this study investigated the reproducibility of repeated BOLD delay measurements, defined as the variation between multiple measurements of the same individual under potentially changing conditions (Bartlett & Frost, 2008). The BOLD signal is affected by systemic and local factors that vary within and between individuals (T. T. Liu, 2016) independent of the progression of the disease being monitored. In this study, we, therefore, provide a detailed quantitative assessment of the reproducibility of BOLD delay.

2 | MATERIALS AND METHODS

2.1 | Subjects

Data from eight different sources were used in this study, six of which are publicly available. Table 1 shows an overview of the data sources and further details are available in the Supporting Information. The main inclusion criterion for this study was the availability of resting-state scans acquired from participants on two separate occasions without any explicit intervention in between that is physiologically relevant to brain perfusion (studies with placebo/sham interventions were also included).

Because we aimed to investigate the reproducibility of BOLD delay measurements, the physiological process being investigated (perfusion) should remain relatively stable during the study period. We thus included data from individuals without any known neurologic illness in all of the datasets except Dataset 2 ("LOBI"). This dataset included patients with acute ischemic stroke but no perfusion deficit on any of their BOLD delay maps (assessed visually by a stroke researcher and radiology resident with 7 years of experience in stroke perfusion imaging, Ahmed A. Khalil). We included this dataset because it represents individuals with a particular profile—older, acutely ill patients with traits, such as a greater propensity to move during MRI

TABLE 1 Overview of the data sources used in this study.

Dataset	Field strength B_0 (T)	TR (s)	TE (ms)	Flip angle (°)	Scan length (min:s)	Number of timepoints	Acceleration factor	Voxel size (mm)	Interval between scans	Number of subjects
1. CSB	3	0.4	30	43	5:40	850	MB = 6	3 × 3 × 4	1 day	12
2. LOBI	3	0.4	30	43	5:40	850	MB = 6	3 × 3 × 4	1 day	31
3. NF (Dobrushina et al., 2018)	3	2.0	30	77	5:00	150	N/A	3 × 3 × 3	30 min	26
4. Modafinil (Esposito et al., 2013)	3	1.671	35	75	4:00	145	N/A	4 × 4 × 4	6 h	13
5. MPI (Gorgolewski et al., 2015)	7	3.0	17	70	15:00	300	R = 3	1.5 × 1.5 × 1.5	1 week	23
6. MSC (Gordon et al., 2017)	3	2.2	27	90	30:00	818	NA	4 × 4 × 4	1 day	10
7. TBI (Roy et al., 2017)	3	2.0	30	90	5:00	150	NA	3 × 3 × 4	3 months	12
8. Yale (Noble et al., 2017)	3	1.0	30	55	5:00	360	R = 2 MB = 5	2 × 2 × 2	1 week	12

Note: A gradient-echo echo planar imaging (EPI) sequence was used for whole-brain rsfMRI data acquisition in all cases.
Abbreviations: MB, multiband acceleration factor; NA, information not available; R, parallel imaging reduction factor; TE, echo time; TR, repetition time.

(Andre et al., 2015; Hodgson et al., 2017; Seto et al., 2001)—that potentially lower the reproducibility of the scans.

2.2 | Image processing

As can be seen in Table 1, the imaging data included in this study spanned a range of imaging parameters, including different spatial and temporal resolutions. All data were acquired on 3 Tesla MRI scanners, except for the "MPI" dataset, which was acquired at 7 Tesla. Intervals between scanning sessions ranged from 30 min to 3 months.

2.2.1 | Preprocessing

The first 10 s of the EPI data were removed (the exact number of volumes removed depended on the repetition time of the sequence) to allow for signal equilibration, followed by spatial realignment of the volumes to the mean of the time series and spatial smoothing with a 4 mm Full Width at Half Maximum (FWHM) Gaussian kernel. The mean framewise displacement across all volumes in a scan was used as a measure of head motion (Power et al., 2012). Framewise displacement is an estimate of motion from one volume to the next. It is calculated by taking the sum of the temporal derivatives of the three rotational parameters (roll, pitch, and yaw), after converting them to distances by calculating the arc length on the surface of a sphere of radius 50 mm, and three displacement parameters (in the x, y, and z planes). In this study, this was done using the *fsl_motion_outliers* function of the FSL library (with the "--fd" option).

2.2.2 | BOLD delay calculation

Using the software package *rapidtide* (https://github.com/bbfrederick/rapidtide), we calculated the time shift required to yield maximum cross-correlation between the BOLD signal in a given voxel and the reference time course, which was derived from the major venous sinuses (Khalil et al., 2017, 2020). The venous sinus reference was chosen because studies have shown that it correlates more strongly than the whole-brain reference with perfusion maps derived from dynamic susceptibility contrast MRI (DSC-MRI) (Christen et al., 2015; Khalil et al., 2017). We used a time shift range of −5 to +5 s for the cross-correlation. The data were temporally bandpass filtered by *rapidtide* to 0.01–0.15 Hz because the signal fluctuations of interest typically reside in this frequency range (Tong et al., 2017). BOLD delay calculation included the estimation of a spurious correlation threshold by creating a distribution of null correlation values using a Monte Carlo approach, a feature which is built into *rapidtide* and is described in more detail in (Tong, Hocke, et al., 2019).

Processing was performed with and without motion "censoring," which involves excluding volumes with a framewise displacement of >0.3 mm from the cross-correlation analysis (Siegel et al., 2017). Motion censoring was performed by providing *rapidtide* with a binary vector of volumes below (=1) and above (=0) the framewise displacement threshold of 0.3 mm using the "--tmask" option. The cross-correlation then only takes into consideration the volumes below the framewise displacement threshold. When motion censoring was performed, it was done before the other analysis steps including temporal filtering. The processing was otherwise performed with *rapidtide*'s default parameters (see processing script https://github.com/ahmedaak/BOLD_delay_repro).

The BOLD delay maps were then registered to a custom EPI template (Khalil et al., 2017), using ANTs (Tustison & Avants, 2013). The EPI template was used to enhance registration quality because some of the datasets lacked a high-resolution anatomical image (e.g., a T1-MPRAGE) (Calhoun et al., 2017).

2.2.3 | Measures of reproducibility

In this study, we assessed the systematic differences (bias) between the scanning sessions by calculating the absolute value of differences in BOLD delay between session 1 (baseline) and session 2 (follow-up) scans. These were calculated by subtracting the BOLD delay value of each voxel in session 1 from the corresponding voxel in session 2 ($|\text{Delay}_{\text{session1}} - \text{Delay}_{\text{session2}}|$) and are referred to as "BOLD delay test-retest differences" in this article (Figure 1). The absolute values of the differences in BOLD delay between the sessions were used because the aim of this study was to assess reproducibility, and we were thus not interested in the directionality of BOLD delay changes. Maps of BOLD delay test–retest differences were registered to the EPI template for further processing.

To exclude areas without blood flow from quantitative analysis, voxels within the ventricles were removed using a template. The average BOLD delay test–retest differences were extracted from a set of ROIs ($n = 400$ in total across the brain) derived from the ADHD 200 parcellation atlas (Bellec et al., 2017; Figure 1). The final units of observation in this study were the BOLD delay test–retest differences within a single ROI. Each ROI was assigned to a vascular territory (middle cerebral artery: MCA; anterior cerebral artery: ACA; posterior cerebral artery: PCA) using a template of the major vascular territories in the brain, the details of which can be found in the Supporting Information of a previous article (Khalil et al., 2017).

2.3 | Statistical analysis

All statistical analysis and data visualization was done in R version 3.1 (R Core Team, 2016). To assess the overall reproducibility of BOLD delay values, we calculated the intraclass correlation coefficient (ICC) using a two-way model of absolute agreement (Matheson, 2019) using the R package "irr" (Gamer et al., 2019).

The influence of various factors on the BOLD delay test–retest differences in the aforementioned ROIs was modeled using a generalized linear mixed model. Linear mixed models are useful for analyzing nonindependent data, such as data with repeated measures or a

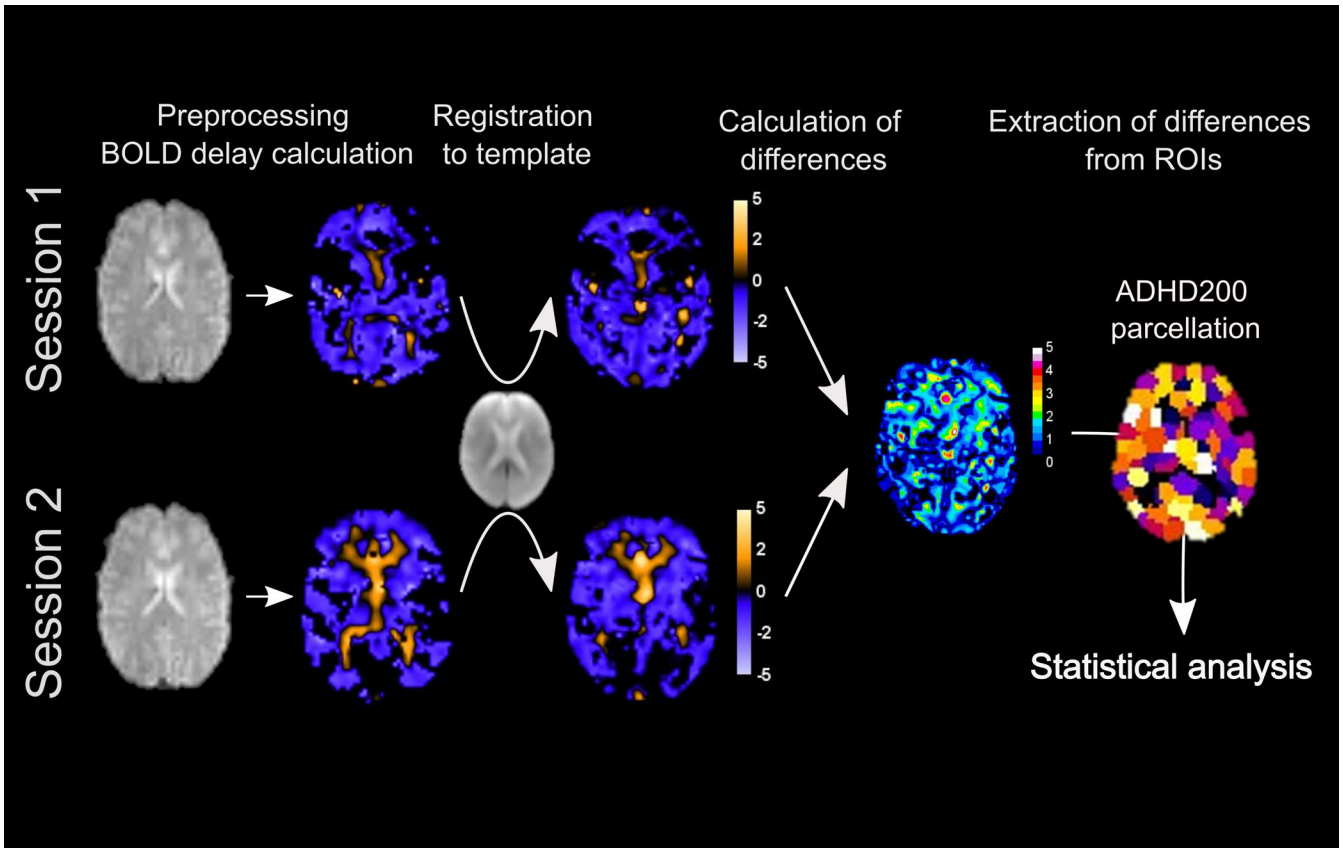

FIGURE 1 Processing pipeline of the resting-state functional MRI (rsfMRI) data. After the single-subject rsfMRI data were preprocessed and BOLD delay maps were calculated, the maps were co-registered to a custom template (using the session's rsfMRI data in an intermediate step), and the absolute differences between the session 1 and session 2 BOLD delay maps were calculated. The mean BOLD delay test–retest differences from each of a set of 400 regions of interest (ROIs) throughout the brain were extracted for each subject and used in the statistical group analysis.

hierarchical structure (Harrison et al., 2018). In the context of reproducibility studies, and unlike methods like Bland–Altman analysis, linear mixed models provide the framework to take into account both variation explained by independent variables of interest (known as "fixed" effects) and variation not explained by these independent variables of interest ("random" effects; Lai & Shiao, 2005). In this study, fixed effects were variables for which we wanted to explicitly quantify variability (e.g., vascular territory and motion), and random effects were variables whose levels could be considered to be a random sample of a population of values (e.g., subjects or datasets), as defined in (Searle et al., 1992). We thus used a three-level random intercept model (first level = measurements, second level = patients, and third level = study).

Specifically, we used a penalized quasi-likelihood generalized linear mixed model using a Gamma error distribution and a log link function, implemented using the "glmmPQL" function from the "MASS" package (Venables & Ripley, 2002). The choice of this model was based on (1) the fact that examination of quantile-quantile plots revealed that the residuals were not normally distributed (hence the use of a generalized linear mixed model approach) and (2) the fact that the dependent variable (BOLD delay test–retest differences) can only take non-negative values (hence, the Gamma error distribution).

In the generalized linear mixed model, subjects and datasets were random effects and fixed effects were the session 1 BOLD delay value, the amount of head motion (defined as the mean value of the mean framewise displacement across time points, i.e., $[FD_{session1} + FD_{session2}]/2$), the study group (stroke patients vs. healthy individuals), the vascular territory, the repetition time of the scan, the length of the scan, the interval between scanning sessions, and the use of motion censoring.

In this article, the distribution of continuous variables is visualized using raincloud plots, which combine dot plots, box plots, and violin plots (Allen et al., 2019). Due to the large number of observations, relationships between BOLD delay test–retest differences and other continuous variables are visualized using heatmaps, where the XY plane of each plot is tiled with squares and the color of the square is a function of the number of observations in that part of the XY plane (Carr et al., 1987). A two-sided statistical significance level of $\alpha = 0.05$ was used in this study. No adjustment for multiple testing was applied. All *p*-values have to be interpreted with caution within an exploratory framework.

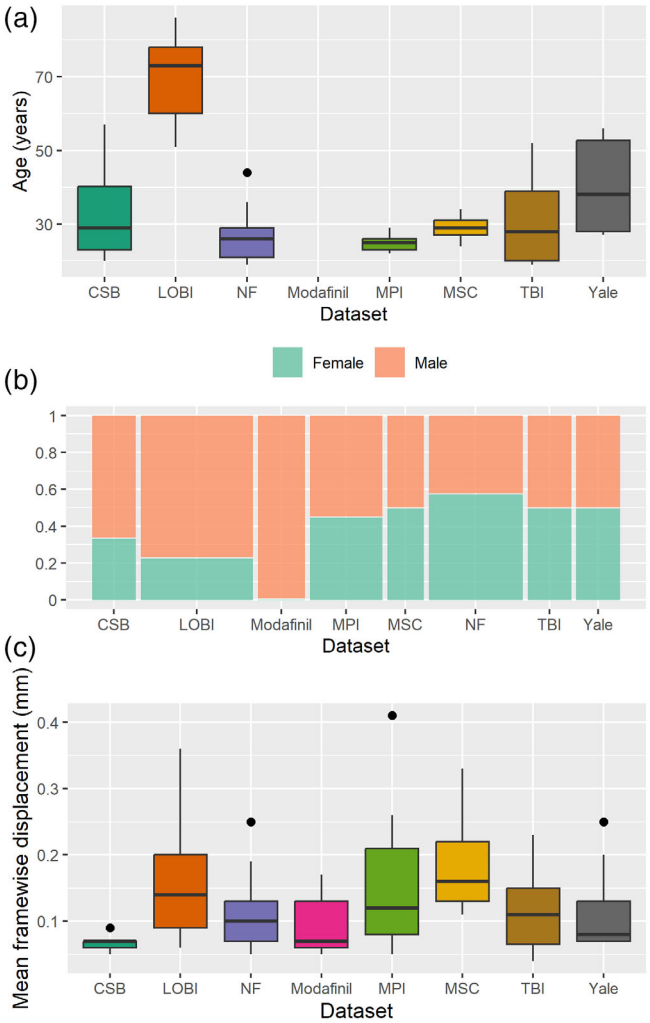

FIGURE 2 Demographics and head motion in each dataset. Note that information about the participants' ages was not available for the "Modafinil" dataset. Framewise displacement is shown as the mean across both scanning sessions. Panels a and c are Tukey boxplots, with the lower and upper hinges corresponding to the 25th and 75th percentiles respectively. The upper and lower whiskers extend from the hinges to the largest and smallest values no further than 1.5 times the interquartile range respectively. Dots represent outliers.

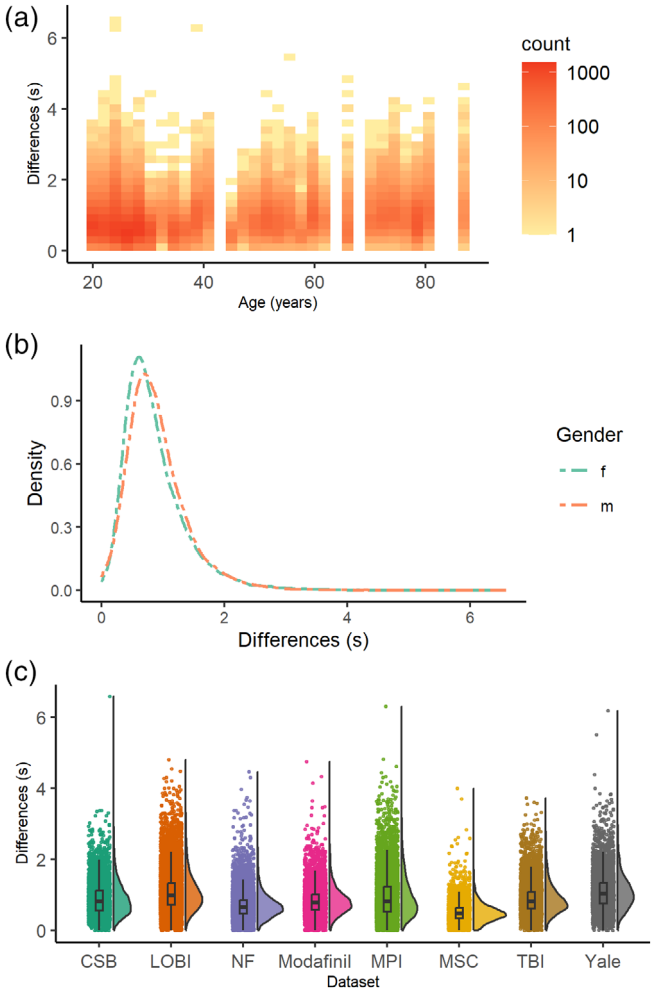

FIGURE 3 Distribution of BOLD delay test–retest differences by age, sex, and study. Heatmap showing the BOLD delay test–retest differences plotted against subject age (a). The heatmap colors represent the number of observations within the area covered by each square. Distribution of BOLD delay test–retest differences in males and females (b). Raincloud plot showing the distribution of BOLD delay test–retest differences in each of the datasets in this study (c).

3 | RESULTS

3.1 | Demographics and head motion

In total, data from 136 subjects were analyzed (see Table 1). The distribution of sex, age, and mean framewise displacement for each of the datasets is shown in Figure 2. Note that for Dataset 5 ("Modafinil"), no data on the participants' ages were available. Figure S1 shows the temporal signal-to-noise ratio (tSNR) of the rsfMRI acquisitions of each dataset (mean across subjects and sessions), calculated according to (Murphy et al., 2007).

Similar to previous studies that showed that head motion is a stable, trait-like property with low between-session variability (Van Dijk et al., 2012), we found that framewise displacement values between the two scanning sessions were highly concordant (Lin's concordance correlation coefficient = 0.70, 95% CI = 0.61–0.78). To simplify the linear mixed model, the mean framewise displacement across both scanning sessions for each subject was used ($[FD_{session1} + FD_{session2}]/2$).

3.2 | BOLD delay reproducibility

The distribution of BOLD delay values for each vascular territory and scanning session is shown in Figures S2 and S3. Figure 3 shows the distribution of BOLD delay test–retest differences by age, gender, and dataset. Figure 4 shows the distribution of BOLD delay test–retest differences across each of the fixed effects in the mixed model.

Table S1 shows the intraclass correlation coefficients (ICCs) between the session 1 and session 2 BOLD delay values. For both the data with and without motion censoring, the highest ICCs were

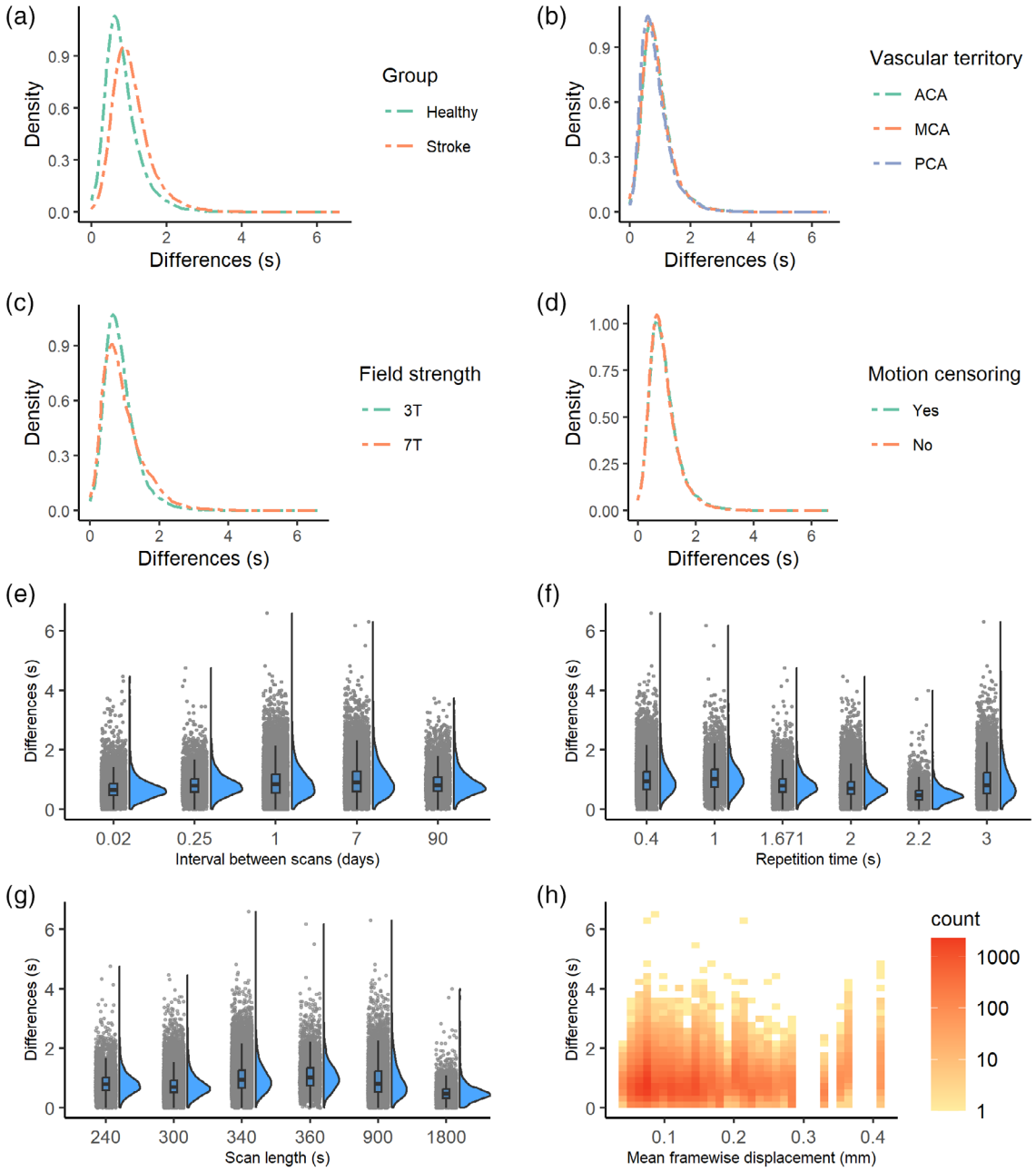

FIGURE 4 Distribution of BOLD delay test–retest differences based on variables in the mixed models. Density plots showing the distribution of BOLD delay test–retest differences according to the study group (a), vascular territory (b), magnetic field strength (c), and the use of motion censoring to mitigate the effects of head motion (d). The raincloud plots show the distribution of BOLD delay test–retest differences according to the interscan interval (e), repetition time of the resting-state functional MRI sequence (f), and resting-state functional MRI scan length (g). The heatmap (h) shows the BOLD delay test–retest differences plotted against the mean framewise displacement across scanning sessions. The heatmap colors represent the number of observations within the area covered by each square.

observed for the posterior cerebral artery (PCA) territory (0.55/0.57), and the lowest ICC for MCA (0.43/0.46).

The B coefficient of the intercept in the model in Table 2 corresponds to the mean of the BOLD delay test–retest differences (on the log scale) after accounting for the fixed and random effects. Table 2 shows that the mean BOLD delay test–retest difference after accounting for the fixed effects was exp(0.16) = 1.2 s (95% CI = 1.0–1.4 s). Figure 5 shows the spatial distribution of the BOLD delay test–retest differences in the sample. The highest BOLD delay differences (exceeding 1 s) were observed in the ventricles, in the frontobasal and temporobasal regions, and around the major brain arteries (Video S1).

3.3 | Effect of covariates on BOLD delay reproducibility

Table 2 summarizes the results of the generalized linear mixed model. BOLD delay test–retest differences were larger in ROIs with higher

TABLE 2 Results of the mixed model of the association between different variables and the BOLD delay test–retest differences.

Variable	B	95% confidence interval	*p*-value
Intercept	0.16	−0.02 to 0.34	.0860
BOLD delay value in session 1	**0.06**	**0.05** to **0.06**	**<.0001**
Vascular territory (MCA)			
PCA	**−0.03**	**−0.04** to **−0.02**	**<.0001**
ACA	**0.03**	**0.02** to **0.03**	**<.0001**
Group (healthy)			
Stroke	0.03	−0.42 to 0.48	.7798
Mean FD	0.04	−0.02 to 0.10	.157
TR	−0.19	−0.47 to 0.08	.0916
Field strength (3T)			
7T	0.51	−0.003 to 1.03	.0506
Interval	0.02	−0.03 to 0.07	.2558
Scan length	−0.02	−0.04 to 0.003	.0667
Motion censoring (no)			
Yes	**0.02**	**0.01** to **0.02**	**<.0001**

Note: For the categorical predictors in the model, the reference category is shown in parentheses. Note that, the mean framewise displacement (FD) variable, in millimeters, was multiplied by 10 to enhance the interpretability of the model's results. Bold values are the ones with a *p*-value <0.0001
Abbreviations: ACA, anterior cerebral artery; B, regression coefficient; FD, framewise displacement; MCA, middle cerebral artery; PCA, posterior cerebral artery; TR, repetition time.

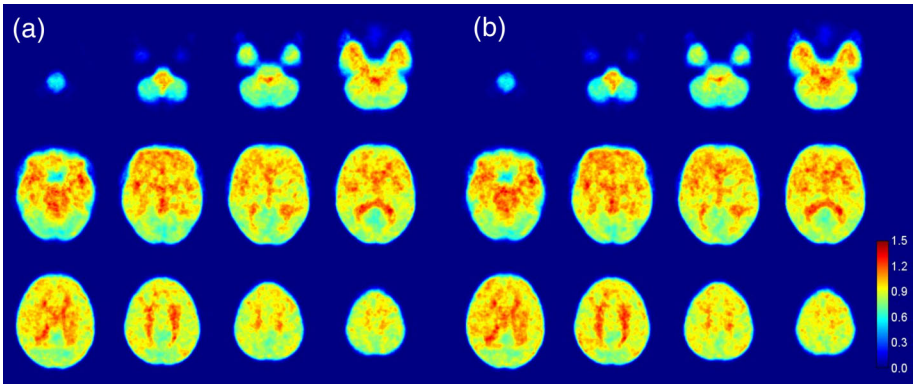

FIGURE 5 The BOLD delay test–retest differences (in seconds; mean across subjects) between scanning sessions. This is shown for the dataset with motion censoring to mitigate the effects of head motion (b) and without motion censoring (a). The lowest reproducibility (highest BOLD delay test–retest differences) is seen in the ventricles, in the frontobasal and temporobasal regions, and in the vicinity of major blood vessels (particularly in the circle of Willis and lateral and central insular sulci).

session 1 BOLD delay values and were smaller in the posterior cerebral artery territory than in the middle cerebral artery vascular territory. In the anterior cerebral artery territory, BOLD delay test–retest differences were larger than in the middle cerebral artery territory. In addition, the use of motion censoring was associated with slightly larger BOLD delay test–retest differences compared with not using motion censoring. Repetition time, interscan interval, magnetic field strength, and scan length showed no substantial association with BOLD delay test–retest differences.

The variability explained by the fixed effects in the model (i.e., the marginal R^2) was 0.12 and the variability explained by both the fixed and random effects in the model (i.e., the conditional R^2) was 0.26.

4 | DISCUSSION

In this study, we evaluated the reproducibility of an emerging, noninvasive method for assessing perfusion based on resting-state functional MRI (BOLD delay). In a large and heterogeneous cohort of subjects with normal cerebral perfusion, we found that the mean BOLD delay test–retest difference, after accounting for several

different variables, was 1.2 s (95% CI = 1.0–1.4 s). Moreover, BOLD delay's reproducibility differed depending on the vascular territory and whether or not motion censoring was used to mitigate the effects of head motion.

4.1 | Overall reproducibility

Few studies have thus far investigated the reproducibility of BOLD delay. In a study of 10 healthy volunteers, Aso et al. found that BOLD delay maps (termed "lag mapping" in their study) showed moderate agreement across scanning sessions (at least 6 days apart) (Aso et al., 2017). In 50 young adults scanned on two consecutive days as part of the Human Connectome Project, (Amemiya et al., 2022) found a median ICC of just under 0.5, comparable to the ICC in our study (0.43–0.57 depending on vascular territory and analysis strategy). The spatial distribution of the test–retest differences in the study by (Amemiya et al., 2022) was also very similar to the current study, being highest in the frontobasal, temporobasal, and periventricular regions.

Yao et al. investigated the within-subject variability of the cerebral circulation time, a measure derived from resting-state fMRI and corresponding to the BOLD delay value across the entire brain, from the brain's arterial input to its venous output (Yao et al., 2019). They found that the standard deviation across three to four scans of the same subject at the 5% confidence level was 2.23 s, and concluded that the variability of this measure is relatively small and comparable to other methods of measuring cerebral circulation time (Hoffmann et al., 2000; X. Liu et al., 2014).

The BOLD delay test–retest differences in this study (mean = 1.2 s, 95% CI = 1.0–1.4) should be interpreted in the context of the magnitude of longitudinal changes in BOLD delay seen in conditions of disturbed perfusion. In a group of acute ischemic stroke patients with baseline perfusion deficits associated with large-vessel occlusions, BOLD delay decreased by a mean of 4.4 s following vessel recanalization (Khalil et al., 2020). Another study found a median BOLD delay decrease of 1.9 s following operative revascularization in patients with chronic steno-occlusive disease (Amemiya et al., 2022). These studies provide further evidence that the BOLD delay method is suitable for the longitudinal monitoring of acute stroke patients with baseline perfusion deficits. For future studies using BOLD delay to monitor other diseases, this study provides a benchmark to aid the interpretation of longitudinal changes.

4.2 | Factors affecting reproducibility

We found differences in the reproducibility of BOLD delay values in different vascular territories. Reproducibility was lowest in the anterior cerebral artery (ACA) territory, followed by the middle cerebral artery (MCA) and posterior cerebral artery (PCA) territories. The similarity between the BOLD signal fluctuations in the brain parenchyma and the venous sinuses (as measured by the maximum cross-correlation coefficient) was highest in the PCA territory in our study (Figure S4). The higher reproducibility of BOLD delay in the PCA territory may be related to this, as larger correlation coefficients are likely to be more robust to noise. Substantial differences exist in the blood flow dynamics of the anterior and posterior circulations, with the latter showing lower flow rates (Dunås et al., 2019) and delayed blood arrival times (Artzi et al., 2013; Goldman-Yassen et al., 2021). In addition, studies of arterial spin labeling mostly show lower reproducibility in the posterior circulation compared to the anterior (Y. Chen et al., 2011; Lin et al., 2020). This may be due to suboptimal angles between the tagging planes and the posterior cerebral arteries, differences in arterial transit time (and thus optimal postlabeling delay) between the anterior and posterior circulations, or higher measurement error caused by lower cerebral blood flow in the posterior circulation (Lin et al., 2020; Mutsaerts et al., 2015). Overall, this indicates that the reproducibility of perfusion values in different vascular territories is dependent on the method used.

In addition, BOLD signal fluctuations related to cardiac and respiratory activity are most prominent in posterior circulation regions (Bianciardi et al., 2009; Brooks et al., 2013) and are highly stable over time within individuals (Birn et al., 2014). The higher proportion of these highly reproducible fluctuations in the posterior compared to the anterior circulation may have led to a higher reproducibility of BOLD delay values, as BOLD delay is calculated by cross-correlating the BOLD fluctuations with a reference signal. Importantly, although a high ICC may be driven by either high between-subject variability or low within-subject variability (Liljequist et al., 2019), Figure S5 suggests that the former is not the case in the PCA territory, where the standard deviation across study subjects is relatively low. In addition, the other measure of reproducibility in this study, the test–retest difference, is not affected by between-subject variability and was also lower in the PCA territory.

The use of motion censoring (the exclusion of high-motion EPI volumes from the BOLD delay calculation) was associated with larger BOLD delay test–retest differences in the current study. This difference was very small and is not readily appreciable on visual inspection of test–retest difference maps averaged across subjects (Figure 5). While the observation that a measure aimed at reducing the effects of motion reduces reproducibility may seem counter-intuitive, several potential explanations exist. First, motion censoring may impair the temporal continuity of the BOLD time courses, which may adversely affect the reproducibility of the BOLD delay calculation, which is based on cross-correlation. Second, particularly in high temporal resolution fMRI, motion calculated using framewise displacement may actually represent "respiratory pseudomotion" of the brain caused by the lungs expanding (Power et al., 2019). Importantly, this pseudomotion partially occurs in the low frequencies (<0.15 Hz), the very same frequencies of interest in studies of functional connectivity and BOLD delay (Power et al., 2019). Respiration may play an important role in driving the sLFOs that are the basis of the BOLD delay calculation, through changes in oxygen saturation leading to variations in the level of deoxyhemoglobin in the vasculature (Aso et al., 2019). Thus, removing volumes with high apparent motion may remove part of the

signal-of-interest driving BOLD delay calculation, particularly when using fast repetition times.

Several studies provide empirical support for the explanation that BOLD delay reproducibility may be impaired by denoising. These showed that the reproducibility of BOLD delay decreased substantially when fMRI data were denoised by regressing out high-frequency, slice-dependent, and white matter components (Amemiya et al., 2022; Aso et al., 2017). Regression of physiological noise (including motion) has also been found to reduce the test–retest reliability of estimates of functional connectivity (Birn et al., 2014). Overall, our results support the notion that denoising prior to BOLD delay calculation should be done with caution, and that motion censoring may be an inappropriate method for dealing with head motion in this context. Further studies are needed to evaluate the influence of other motion correction strategies on BOLD delay assessment.

Interestingly, our study showed only a weak negative association between fMRI sampling frequency (repetition time) and BOLD delay test–retest differences. This is in contrast to the study by Aso et al., in which the artificial decimation of the sampling interval (from 0.5 to 5 s in intervals of 0.5 s) resulted in reduced reproducibility of the BOLD delay maps. This discrepancy may be because artificially decimated data and data acquired with different repetition times will not be identical, or because the range of repetition times used in our study was narrower (from 0.4 to 3.0 s) than that used by Aso et al. This observation also suggests that the signal driving the BOLD delay was likely sufficiently sampled with the range of TRs used in this study.

Although this was not reflected in the results of the mixed model, Figure 4g shows that test–retest differences are smaller in the dataset with the longest rsfMRI scan time (1800 s). This is likely due to the fact that there is a higher chance of reliably detecting cross-correlation peaks with longer scans. The optimal scan length for the calculation of BOLD delay maps will, however, also depend on factors other than reproducibility such as the specific use case (Amemiya et al., 2022; Tanrıtanır et al., 2020).

This study is the largest detailed evaluation of the reproducibility of this emerging, noninvasive method of assessing brain perfusion to date. Its strengths include a relatively large sample size, heterogeneous data from multiple scanners, the inclusion of data from stroke patients (one of BOLD delay's most promising use cases), and the use of a statistical model accounting for multiple factors potentially affecting the method's reproducibility. We nonetheless acknowledge several limitations. These include the fact that the reproducibility depends on the quality of between-session registration of the rsfMRI data. We attempted to mitigate this effect by taking the mean of the BOLD delay test–retest differences across regions-of-interest, instead of relying on voxelwise values, which would be more sensitive to registration errors. Additionally, in the cohort of stroke patients included in this study (the "LOBI" dataset), we cannot completely exclude the presence of small stroke-related changes in perfusion and consequently BOLD delay (i.e., a change in perfusion over time due to the natural history of the disease, or therapeutic interventions having nothing to do with the reproducibility of the measure). The exclusion of subjects with BOLD delay lesions based on visual screening of the maps aimed to minimize this. However, Figure 4a shows that the BOLD delay test–retest differences were indeed larger in the stroke cohort than in healthy controls. We believe including a stroke cohort is justified despite this potential confounder, as they tend to have unique, often more extreme, head motion patterns than healthy individuals or nonacutely ill patients (Seto et al., 2001). They are also older than the rest of the sample, have a higher incidence of general stroke risk factors such as vessel wall stiffness, atherosclerosis, and diabetes, and thus reflect the spectrum of comorbidities seen in cerebrovascular disease patients.

We could not account for all factors potentially influencing longitudinal changes in BOLD delay. Only 26% of the variance in BOLD delay test–retest differences was explained by the fixed and random effects included in our generalized linear mixed model. Factors such as caffeine intake, for example, are known to affect cerebral perfusion (Clement et al., 2018; Pelligrino et al., 2010) and have been shown to influence BOLD delay directly (Yang et al., 2018). Such information was not available in the datasets used in this study, so they could not be included in the models. Future studies accounting for physiological factors influencing longitudinal cerebral perfusion changes may shed more light on this issue. Processing parameters of the rsfMRI data (e.g., the degree of spatial smoothing and temporal filtering, the reference time series, and the time shift range) might also influence the calculated BOLD delay values and their effect on BOLD delay reproducibility should also be investigated in future studies.

Cerebral perfusion measured using BOLD delay shows good reproducibility that is highest in the posterior cerebral artery territory and lowest in the anterior cerebral artery territory. Overall, longitudinal changes in BOLD delay of up to 1.2 s are not necessarily pathological and might be the result of physiological factors that require further investigation. Longitudinal BOLD delay changes in stroke are typically more than three times larger than this, thus BOLD delay is likely appropriate for monitoring perfusion deficits in cerebrovascular disorders.

AUTHOR CONTRIBUTIONS

Ahmed A. Khalil: Conceptualization, methodology, software, formal analysis, investigation, data curation, writing - original draft, writing - review and editing, visualization, project administration. **Ayse C. Tanritanir:** Methodology, investigation, resources, writing - review and editing. **Ulrike Grittner:** Formal analysis, writing - review and editing. **Evgeniya Kirilina:** Resources, writing - review and editing. **Arno Villringer:** Methodology, supervision, writing - review and editing. **Jochen B. Fiebach:** Methodology, supervision, resources, funding acquisition, writing - review and editing. **Ralf Mekle:** Conceptualization, methodology, investigation, supervision, writing - original draft, writing - review and editing.

ACKNOWLEDGMENTS

The authors thank the University of Minnesota Center for Magnetic Resonance Research (CMRR) for providing the multiband echo planar imaging sequence used in this study (as part of the "CSB" and "LOBI" datasets).

FUNDING INFORMATION

During the time this research was conducted, Dr. Ahmed Khalil was a participant in the BIH-Charité Junior Clinician Scientist Program funded by the Charité–Universitätsmedizin Berlin and the Berlin Institute of Health. Prof. Jochen Fiebach was funded by the German Federal Ministry of Education and Research (01EO0801, 01EO01301). The funding sources played no role in this study's design, in data collection, in the analysis or interpretation of the data, in writing this article, or in the decision to submit this article for publication.

CONFLICT OF INTEREST STATEMENT

Dr. Ahmed Khalil and Prof. Jochen Fiebach are co-inventors of a method for automated, user-independent delineation of perfusion lesions (now in the public domain). Dr. Ahmed Khalil has received personal fees from Bayer AG outside the submitted work. Prof. Jochen Fiebach has received personal fees from Abbvie, AC Immune, Artemida, Biogen, BMS, Brainomix, Cerevast, Clario, Daiichi-Sankyo, Eisai, Eli Lilly, Guerbet, Ionis Pharmaceuticals, Julius clinical, Jung diagnostics, Merck, Nicolab, and Tau Rx, outside the submitted work.

DATA AVAILABILITY STATEMENT

The code and data used for performing the analysis and visualizing the results are available here: https://github.com/ahmedaak/BOLD_delay_repro.

ORCID

Ahmed A. Khalil https://orcid.org/0000-0003-1752-4305
Arno Villringer https://orcid.org/0000-0003-2604-2404

REFERENCES

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. A. (2019). Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Research*, *4*, 63.

Amemiya, S., Kunimatsu, A., Saito, N., & Ohtomo, K. (2014). Cerebral hemodynamic impairment: Assessment with resting-state functional MR imaging. *Radiology*, *270*(2), 548–555.

Amemiya, S., Takao, H., Hanaoka, S., & Ohtomo, K. (2016). Global and structured waves of rs-fMRI signal identified as putative propagation of spontaneous neural activity. *NeuroImage*, *133*, 331–340.

Amemiya, S., Takao, H., Watanabe, Y., Miyawaki, S., Koizumi, S., Saito, N., & Abe, O. (2022). Reliability and sensitivity to alterered hemodynamics measured with resting-state fMRI metrics: Comparison with 123I-IMP SPECT. *NeuroImage*, *263*, 119654.

Andre, J. B., Bresnahan, B. W., Mossa-Basha, M., Hoff, M. N., Smith, C. P., Anzai, Y., & Cohen, W. A. (2015). Toward quantifying the prevalence, severity, and cost associated with patient motion during clinical MR examinations. *Journal of the American College of Radiology*, *12*(7), 689–695.

Artzi, M., Aizenstein, O., Abramovitch, R., & Bashat, D. B. (2013). MRI multiparametric hemodynamic characterization of the normal brain. *Neuroscience*, *240*, 269–276.

Aso, T., Jiang, G., Urayama, S. I., & Fukuyama, H. (2017). A resilient, non-neuronal source of the spatiotemporal lag structure detected by bold signal-based blood flow tracking. *Frontiers in Neuroscience*, *11*(MAY), 256.

Aso, T., Urayama, S., Fukuyama, H., & Murai, T. (2019). Axial variation of deoxyhemoglobin density as a source of the low-frequency time lag structure in blood oxygenation level-dependent signals. *PLoS One*, *14*(9), e0222787.

Bartlett, J. W., & Frost, C. (2008). Reliability, repeatability and reproducibility: Analysis of measurement errors in continuous variables. *Ultrasound in Obstetrics & Gynecology*, *31*(4), 466–475.

Bellec, P., Chu, C., Chouinard-Decorte, F., Benhajali, Y., Margulies, D. S., & Craddock, R. C. (2017). The Neuro Bureau ADHD-200 Preprocessed repository. *NeuroImage*, *144*(Pt B), 275–286.

Bianciardi, M., Fukunaga, M., van Gelderen, P., Horovitz, S. G., de Zwart, J. A., Shmueli, K., & Duyn, J. H. (2009). Sources of functional magnetic resonance imaging signal fluctuations in the human brain at rest: A 7 T study. *Magnetic Resonance Imaging*, *27*(8), 1019–1029.

Birn, R. M., Cornejo, M. D., Molloy, E. K., Patriat, R., Meier, T. B., Kirk, G. R., Nair, V. A., Meyerand, M. E., & Prabhakaran, V. (2014). The influence of physiological noise correction on test-retest reliability of resting-state functional connectivity. *Brain Connectivity*, *4*(7), 511–522.

Brooks, J. C. W. P., Faull, O. K., Pattinson, K. T. S. D. F., & Jenkinson, M. P. (2013). Physiological noise in brainstem fMRI. *Frontiers in Human Neuroscience*, *7*(October), 623.

Calhoun, V. D., Wager, T. D., Krishnan, A., Rosch, K. S., Seymour, K. E., Nebel, M. B., Mostofsky, S. H., Nyalakanai, P., & Kiehl, K. (2017). The impact of T1 versus EPI spatial normalization templates for fMRI data analyses. *Human Brain Mapping*, *38*, 5331–5342. https://doi.org/10.1002/hbm.23737

Carr, D. B., Littlefield, R. J., Nicholson, W. L., & Littlefield, J. S. (1987). Scatterplot matrix techniques for large N. *Journal of the American Statistical Association*, *82*(398), 424–436.

Chen, Q., Zhou, J., Zhang, H., Chen, Y., Mao, C., Chen, X., Ni, L., Zhuo, Z., Zhang, Y., Geng, W., Yin, X., & Lv, Y. (2019). One-step analysis of brain perfusion and function for acute stroke patients after reperfusion: A resting-state fMRI study. *Journal of Magnetic Resonance Imaging*, *50*(1), 221–229.

Chen, Y., Wang, D. J. J., & Detre, J. A. (2011). Test-retest reliability of arterial spin labeling with common labeling strategies. *Journal of Magnetic Resonance Imaging*, *33*(4), 940–949.

Christen, T., Jahanian, H., Ni, W. W., Qiu, D., Moseley, M. E., & Zaharchuk, G. (2015). Noncontrast mapping of arterial delay and functional connectivity using resting-state functional MRI: A study in Moyamoya patients. *Journal of Magnetic Resonance Imaging*, *41*(2), 424–430.

Clement, P., Mutsaerts, H.-J., Václavů, L., Ghariq, E., Pizzini, F. B., Smits, M., Acou, M., Jovicich, J., Vanninen, R., Kononen, M., Wiest, R., Rostrup, E., Bastos-Leite, A. J., Larsson, E.-M., & Achten, E. (2018). Variability of physiological brain perfusion in healthy subjects—A systematic review of modifiers. Considerations for multi-center ASL studies. *Journal of Cerebral Blood Flow and Metabolism*, *38*(9), 1418–1437.

Cologner, J., Vu, C., Bush, A., Borzage, M., Rajagopalan, V., Lepore, N., & Wood, J. (2016). BOLD delay times using group delay in sickle cell disease. In M. A. Styner & E. D. Angelini (Eds.), *SPIE medical imaging* (p. 97843M). International Society for Optics and Photonics.

Dobrushina, O. R., Pechenkova, E. V., Vlasova, R. M., Rumshiskaya, A. D., Litvinova, L. D., Mershina, E. A., & Sinitsyn, V. E. (2018). Exploring the brain contour of implicit infra-low frequency EEG neurofeedback: A resting state fMRI study. *International Journal of Psychophysiology*, *131*, S76.

Dunås, T., Wåhlin, A., Zarrinkoob, L., Malm, J., & Eklund, A. (2019). 4D flow MRI: Automatic assessment of blood flow in cerebral arteries. *Biomedical Physics & Engineering Express*, *5*(1). https://doi.org/10.1088/2057-1976/aae8d1

Esposito, R., Cilli, F., Pieramico, V., Ferretti, A., Macchia, A., Tommasi, M., Saggino, A., Ciavardelli, D., Manna, A., Navarra, R., Cieri, F., Stuppia, L., Tartaro, A., & Sensi, S. L. (2013). Acute effects of modafinil on brain resting state networks in young healthy subjects. *PLoS One*, *8*(7), e69224.

European Medicines Agency. (2018). *Gadolinium-containing contrast agents*. Https://www.ema.europa.eu/. https://www.ema.europa.eu/en/medicines/human/referrals/gadolinium-containing-contrast-agents

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2019). irr: Various coefficients of interrater reliability and agreement. https://CRAN.R-project.org/package=irr

Goldman-Yassen, A. E., Straka, M., Uhouse, M., & Dehkharghani, S. (2021). Normative distribution of posterior circulation tissue time-to-maximum: Effects of anatomic variation, tracer kinetics, and implications for patient selection in posterior circulation ischemic stroke. *Journal of Cerebral Blood Flow and Metabolism*, *41*(8):1912–1923.

Gordon, E. M., Laumann, T. O., Gilmore, A. W., Newbold, D. J., Greene, D. J., Berg, J. J., Ortega, M., Hoyt-Drazen, C., Gratton, C., Sun, H., Hampton, J. M., Coalson, R. S., Nguyen, A. L., McDermott, K. B., Shimony, J. S., Snyder, A. Z., Schlaggar, B. L., Petersen, S. E., Nelson, S. M., & Dosenbach, N. U. F. (2017). Precision functional mapping of individual human brains. *Neuron*, *95*(4), 791–807.e7.

Gorgolewski, K. J., Mendes, N., Wilfling, D., Wladimirow, E., Gauthier, C. J., Bonnen, T., Ruby, F. J. M., Trampel, R., Bazin, P.-L., Cozatl, R., Smallwood, J., & Margulies, D. S. (2015). A high resolution 7-Tesla resting-state fMRI test-retest dataset with cognitive and physiological measures. *Scientific Data*, *2*, 140054.

Gulani, V., Calamante, F., Shellock, F. G., Kanal, E., & Reeder, S. B. (2017). Gadolinium deposition in the brain: Summary of evidence and recommendations. *Lancet Neurology*, *16*(7), 564–570.

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794.

Hodgson, K., Poldrack, R. A., Curran, J. E., Knowles, E. E., Mathias, S., Göring, H. H. H., Yao, N., Olvera, R. L., Fox, P. T., Almasy, L., Duggirala, R., Barch, D. M., Blangero, J., & Glahn, D. C. (2017). Shared genetic factors influence head motion during MRI and Body Mass Index. *Cerebral Cortex*, *27*(12), 5539–5546.

Hoffmann, O., Weih, M., Schreiber, S., Einhäupl, K. M., & Valdueza, J. M. (2000). Measurement of cerebral circulation time by contrast-enhanced Doppler sonography. *Cerebrovascular Diseases*, *10*(2), 142–146.

Khalil, A. A., Ostwaldt, A.-C., Nierhaus, T., Ganeshan, R., Audebert, H. J., Villringer, K., Villringer, A., & Fiebach, J. B. (2017). Relationship between changes in the temporal dynamics of the blood-oxygen-level-dependent signal and hypoperfusion in acute ischemic stroke. *Stroke*, *48*(4), 925–931.

Khalil, A. A., Villringer, K., Filleböck, V., Hu, J.-Y., Rocco, A., Fiebach, J. B., & Villringer, A. (2020). Non-invasive monitoring of longitudinal changes in cerebral hemodynamics in acute ischemic stroke using BOLD signal delay. *Journal of Cerebral Blood Flow and Metabolism*, *40*(1), 23–34.

Lai, D., & Shiao, S.-Y. P. K. (2005). Comparing two clinical measurements: A linear mixed model approach. *Journal of Applied Statistics*, *32*(8), 855–860.

Liljequist, D., Elfving, B., & Skavberg Roaldsen, K. (2019). Intraclass correlation—A discussion and demonstration of basic features. *PLoS One*, *14*(7), e0219854.

Lin, T., Qu, J., Zuo, Z., Fan, X., You, H., & Feng, F. (2020). Test-retest reliability and reproducibility of long-label pseudo-continuous arterial spin labeling. *Magnetic Resonance Imaging*, *73*, 111–117.

Liu, T. T. (2016). Noise contributions to the fMRI signal: An overview. *NeuroImage*, *143*, 141–151.

Liu, X., Yang, Y. L., Sun, S. G., Yang, R. J., Wang, J., Li, Y., Zhang, L., & Duan, Y. Y. (2014). A new method of measurement of cerebral circulation time: Contrast-enhanced ultrasonography in healthy adults and patients with intracranial shunts. *Ultrasound in Medicine & Biology*, *40*(10), 2372–2378. https://doi.org/10.1016/j.ultrasmedbio.2014.03.025

Lv, Y., Margulies, D. S., Cameron Craddock, R., Long, X., Winter, B., Gierhake, D., Endres, M., Villringer, K., Fiebach, J., & Villringer, A. (2013). Identifying the perfusion deficit in acute stroke with resting-state functional magnetic resonance imaging. *Annals of Neurology*, *73*(1), 136–140.

Lv, Y., Wei, W., Song, Y., Han, Y., Zhou, C., Zhou, D., Zhang, F., Xue, Q., Liu, J., Zhao, L., Zhang, C., Li, L., Zang, Y. F., & Han, X. (2018). Non-invasive evaluation of cerebral perfusion in patients with transient ischemic attack: An fMRI study. *Journal of Neurology*, *266*(1), 157–164.

Matheson, G. J. (2019). We need to talk about reliability: Making better use of test-retest studies for study design and interpretation. *PeerJ*, *7*, e6918.

Murphy, K., Bodurka, J., & Bandettini, P. A. (2007). How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. *NeuroImage*, *34*(2), 565–574.

Mutsaerts, H. J. M. M., van Dalen, J. W., Heijtel, D. F. R., Groot, P. F. C., Majoie, C. B. L. M., Petersen, E. T., Richard, E., & Nederveen, A. J. (2015). Cerebral perfusion measurements in elderly with hypertension using arterial spin labeling. *PLoS One*, *10*(8), e0133717.

Ni, L., Li, J., Li, W., Zhou, F., Wang, F., Schwarz, C. G., Liu, R., Zhao, H., Wu, W., Zhang, X., Li, M., Yu, H., Zhu, B., Villringer, A., Zang, Y., Zhang, B., Lv, Y., & Xu, Y. (2017). The value of resting-state functional MRI in subacute ischemic stroke: Comparison with dynamic susceptibility contrast-enhanced perfusion MRI. *Scientific Reports*, *7*, 41586.

Noble, S., Spann, M. N., Tokoglu, F., Shen, X., Constable, R. T., & Scheinost, D. (2017). Influences on the test–retest reliability of functional connectivity MRI and its relationship with behavioral utility. *Cerebral Cortex*, *27*(11), 5415–5429.

Pelligrino, D. A., Xu, H.-L., & Vetri, F. (2010). Caffeine and the control of cerebral hemodynamics. *Journal of Alzheimer's Disease*, *20*(suppl 1), S51–S62.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, *59*(3), 2142–2154.

Power, J. D., Lynch, C. J., Silver, B. M., Dubin, M. J., Martin, A., & Jones, R. M. (2019). Distinctions among real and apparent respiratory motions in human fMRI data. *NeuroImage*, *201*, 116041.

R Core Team. (2016). R: A language and environment for statistical computing. In *R foundation for statistical computing*. R Foundation for Statistical Computing.

Roy, A., Bernier, R. A., Wang, J., Benson, M., French, J. J., Good, D. C., & Hillary, F. G. (2017). The evolution of cost-efficiency in neural networks during recovery from traumatic brain injury. *PLoS One*, *12*(4), e0170541.

Schellinger, P. D., Jansen, O., Fiebach, J. B., Heiland, S., Steiner, T., Schwab, S., Pohlers, O., Ryssel, H., Sartor, K., & Hacke, W. (2000). Monitoring intravenous recombinant tissue plasminogen activator thrombolysis for acute ischemic stroke with diffusion and perfusion MRI. *Stroke*, *31*(6), 1318–1328.

Searle, S. R., Casella, G., & McCulloch, C. E. (Eds.). (1992). Introduction. In *Variance components* (pp. 1–18). John Wiley & Sons, Inc.

Seto, E., Sela, G., McIlroy, W. E., Black, S. E., Staines, W. R., Bronskill, M. J., McIntosh, A. R., & Graham, S. J. (2001). Quantifying head motion associated with motor tasks used in fMRI. *NeuroImage*, *14*(2), 284–297.

Shah, M. N., Mitra, A., Goyal, M. S., Snyder, A. Z., Zhang, J., Shimony, J. S., Limbrick, D. D., Raichle, M. E., & Smyth, M. D. (2018). Resting state signal latency predicts laterality in pediatric medically refractory temporal lobe epilepsy. *Child's Nervous System*, *34*(5), 901–910.

Siegel, J. S., Shulman, G. L., & Corbetta, M. (2017). Measuring functional connectivity in stroke: Approaches and considerations. *Journal of Cerebral Blood Flow and Metabolism*, *37*(8), 2665–2678.

Siegel, J. S., Snyder, A. Z., Ramsey, L., Shulman, G. L., & Corbetta, M. (2016). The effects of hemodynamic lag on functional connectivity and behavior after stroke. *Journal of Cerebral Blood Flow and Metabolism*, *36*(12), 2162–2176.

Tanrıtanır, A. C., Villringer, K., Galinovic, I., Grittner, U., Kirilina, E., Fiebach, J. B., Villringer, A., & Khalil, A. A. (2020). The effect of scan length on the assessment of BOLD delay in ischemic stroke. *Frontiers in Neurology*, *11*, 381.

Tong, Y., Hocke, L. M., & Frederick, B. B. (2019). Low frequency systemic hemodynamic "noise" in resting state BOLD fMRI: Characteristics, causes, implications, mitigation strategies, and applications. *Frontiers in Neuroscience*, *13*, 787.

Tong, Y., Lindsey, K. P., Hocke, L. M., Vitaliano, G., Mintzopoulos, D., & Frederick, B. D. (2017). Perfusion information extracted from resting state functional magnetic resonance imaging. *Journal of Cerebral Blood Flow and Metabolism*, *37*(2), 564–576.

Tong, Y., Yao, J. F., Chen, J. J., & Frederick, B. D. (2019). The resting-state fMRI arterial signal predicts differential blood transit time through the brain. *Journal of Cerebral Blood Flow and Metabolism*, *39*(6), 1148–1160.

Tustison, N. J., & Avants, B. B. (2013). Explicit B-spline regularization in diffeomorphic image registration. *Frontiers in Neuroinformatics*, *7*, 39.

Van Dijk, K. R. A., Sabuncu, M. R., & Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *NeuroImage*, *59*(1), 431–438.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer http://www.stats.ox.ac.uk/pub/MASS4

Wu, J., Dehkharghani, S., Nahab, F., Allen, J., & Qiu, D. (2017). The effects of acetazolamide on the evaluation of cerebral hemodynamics and functional connectivity using blood oxygen level-dependent MR imaging in patients with chronic steno-occlusive disease of the anterior circulation. *American Journal of Neuroradiology*, *38*(1), 139–145.

Yan, S., Qi, Z., An, Y., Zhang, M., Qian, T., & Lu, J. (2018). Detecting perfusion deficit in Alzheimer's disease and mild cognitive impairment patients by resting-state fMRI. *Journal of Magnetic Resonance Imaging*, *49*, 1099–1104. https://doi.org/10.1002/jmri.26283

Yang, H.-C. S., Liang, Z., Yao, J. F., Shen, X., Frederick, B. D., & Tong, Y. (2018). Vascular effects of caffeine found in BOLD fMRI. *Journal of Neuroscience Research*, *97*, 456–466. https://doi.org/10.1002/jnr.24360

Yao, J. F., Wang, J. H., Yang, H. S., Liang, Z., Cohen-Gadol, A. A., Rayz, V. L., & Tong, Y. (2019). Cerebral circulation time derived from fMRI signals in large blood vessels. *Journal of Magnetic Resonance Imaging*, *50*, 1504–1513.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Khalil, A. A., Tanritanir, A. C., Grittner, U., Kirilina, E., Villringer, A., Fiebach, J. B., & Mekle, R. (2023). Reproducibility of cerebral perfusion measurements using BOLD delay. *Human Brain Mapping*, 1–12. https://doi.org/10.1002/hbm.26244