HHS Public Access

Author manuscript

Magn Reson Imaging. Author manuscript; available in PMC 2018 January 01.

Published in final edited form as:

Magn Reson Imaging. 2018 January ; 45: 26–33. doi:10.1016/j.mri.2017.09.004.

Test-retest reliability of cerebral blood flow in healthy individuals using arterial spin labeling: Findings from the EMBARC study

Jorge R.C. Almeida[a,b,j,*], **Tsafrir Greenberg**[a], **Hanzhang Lu**[c], **Henry W. Chase**[a], **Jay C. Fournier**[a], **Crystal M. Cooper**[c], **Thilo Deckersbach**[d], **Phil Adams**[e], **Thomas Carmody**[e], **Maurizio Fava**[d], **Benji Kurian**[c], **Patrick J. McGrath**[f], **Melvin G. McInnis**[g], **Maria A. Oquendo**[h], **Ramin Parsey**[i], **Myrna Weissman**[f], **Madhukar Trivedi**[c], and **Mary L. Phillips**[a]

[a]Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213, USA

[b]Department of Psychiatry, Brown University School of Medicine, Providence, RI 02906, USA

[c]Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX 75235, USA

[d]Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA

[e]Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, TX 75235, USA

[f]Department of Psychiatry, Columbia University College of Physicians and Surgeons and the New York State Psychiatric Institute, New York, NY 10032, USA

[g]Department of Psychiatry, University of Michigan School of Medicine, Ann Arbor, MI 48109, USA

[h]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104-3309, USA

[i]Departments of Psychiatry & Radiology, Stony Brook University, Stony Brook, NY 11794, USA

[j]Departments of Psychiatry, Dell Medical School, University of Texas at Austin, Austin, TX 78712, USA

Abstract

Introduction—Previous investigations of test-retest reliability of cerebral blood flow (CBF) at rest measured with pseudo-continuous Arterial Spin Labeling (pCASL) demonstrated good reliability, but are limited by the use of similar scanner platforms. In the present study we examined test-retest reliability of CBF in regions implicated in emotion and the default mode network.

Material and methods—We measured absolute and relative CBF at rest in thirty-one healthy subjects in two scan sessions, one week apart, at four different sites and three different scan

[*]Corresponding author at: Dell Medical School at University of Texas at Austin, Department of Psychiatry, 1701 Trinity St., Stop Z0600 Austin, TX 78712-1873, USA, Jorge.Almeida@austin.utexas.edu (J.R.C. Almeida).

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.mri.2017.09.004.

platforms. We derived CBF from pCASL images with an automated algorithm and calculated intra-class correlation coefficients (ICCs) across sessions for regions of interest. In addition, we investigated site effects.

Results—For both absolute and relative CBF measures, ICCs were good to excellent (i.e. > 0.6) in most brain regions, with highest values observed for the subgenual anterior cingulate cortex and ventral striatum. A leave-one-site-out cross validation analysis did not show a significant effect for site on whole brain CBF and there was no proportional bias across sites. However, a significant site effect was present in the repeated measures ANOVA.

Conclusions—The high test-retest reliability of CBF measured with pCASL in a range of brain regions implicated in emotion and salience processing, emotion regulation, and the default mode network, which have been previously linked to depression symptomatology supports its use in studies that aim to identify neuroimaging biomarkers of treatment response.

Keywords

Arterial spin labeling; Reliability; fMRI; Cerebral blood flow

1. Introduction

Arterial spin-labeling (ASL) is a noninvasive magnetic resonance imaging (MRI) method that uses magnetically labeled arterial blood water as an endogenous tracer for cerebral blood flow (CBF) perfusion imaging [1]. Unlike positron emission tomography (PET) and single-photon emission computed tomography (SPECT), which require radioligands, ASL uses an endogenous marker that allows repeated use without radiation exposure [1]. Importantly, ASL shows high concordance with these more invasive neuroimaging techniques [2–4]. For example, resting CBF measures in cortical regions acquired with ASL and H_2O PET were similar in healthy subjects [3]. The non-invasive nature of ASL has facilitated its use in animal studies [5,6] and human studies with both healthy [7,8] and clinical populations [9–12], as well as pharmacological studies [13–16].

Given the promise of ASL imaging to help identify biomarkers in psychiatric disorders [1,9], it is important to demonstrate its reliability during repeated testing, both within and between scanner platforms. Several studies suggest that ASL is a reliable and reproducible measure of regional CBF over time using the same scanner [13,14,17–32], as well as across different sites using the same MRI scanner platforms [20,24,25,29,30]. Only two studies from the same group explored ASL reliability across two different scanner platforms, finding good reliability both at rest and during finger tapping in the motor cortex [33], and good reliability in mean global gray matter CBF at rest [34]. To date, only one study investigated ASL reliability across scanners of three different platforms [31], finding good reproducibility for most regions of interest studied, except posterior cingulate cortex, and concluding that differences in sequence parameters have greater impact on ASL reproducibility compared to differences in hardware/software across MRI scanner platforms. However, the study did not examine reliability across the three platforms as one unique group, but rather as two parallel investigations. Moreover, most other ASL reliability studies investigated CBF stability within a few hours [20,25,30,31] or within several days [20,24,25,29]. Overall, these studies

reported high reliability for scans completed within the same day, and a decrease in reliability after several days. Potential biomarkers of treatment-associated changes are more likely to yield clinical translation when measured after a longer interval, such as one week, rather than after few hours or days due to the time required for even the earliest treatment effects to emerge.

In the current study, we aimed to assess, for the first time, the test-retest reliability of ASL in healthy subjects scanned one week apart at four sites across three different MRI scanner platforms. This healthy cohort is part of a larger study, Establishing Moderators and Biosignatures of Antidepressant Response in Clinical Care (EMBARC), investigating treatment in major depressive disorder (MDD) [35]. We focused our analysis on brain regions previously reported to be dysfunctional in MDD, including regions implicated in emotion and salience processing, implicit and effortful emotional regulation, and regions of the default mode network [36]. Regions in these networks also show abnormalities in intrinsic (resting state) functional connectivity in MDD [36] and are therefore potential neural targets for antidepressant medications.

Based on previous findings indicating that intra-class correlation coefficients of ASL measures may differ between scans acquired a few hours apart from those acquired several days apart [18,19,21,23,25–29,32] and limited information on possible effects of data acquisition using different scans platforms, we aimed to determine the test-retest reliability of ASL measures in these networks across sites/platforms on scans acquired one week apart.

2. Material and methods

2.1. Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC) study

The EMBARC study (http://embarc.utsouthwestern.edu) [35] is a large multisite randomized controlled trial aiming to identify neuroimaging and other moderators and mediators of treatment response in depressed subjects with MDD, and also included recruitment of healthy control subjects to test reliability of the biomarkers, which is the focus of the current manuscript. All study subjects undergo neuroimaging assessment at baseline and after one week [35,37,38], which provides an ideal opportunity to acquire an ASL dataset across different sites using the same protocol.

2.2. Subjects

Forty healthy subjects were tested twice, one week apart, in one of four sites (10 different subjects per site). Subjects were between the ages of 18 and 65, fluent in English and scored lower than 8 on the Quick Inventory of Depressive Symptomatology-Self Report (QIDS-SR) -a self-report measure of depressive symptomatology [39]. Exclusion criteria included: lifetime history of Axis I mood disorder, anxiety disorder, psychotic disorder, or eating disorder; presence of epilepsy or other conditions requiring an anticonvulsant; presence of thyroid medication for hypothyroidism unless stable on thyroid medication for 3 months; history of a general medical condition that required hospitalization or deemed clinically significant; any clinically significant abnormal laboratory results and any clinically

significant neurological injury or deficit. The usual MRI exclusion criteria also applied, including being pregnant or breastfeeding, presence of ferromagnetic metal in one's body. The study was approved by the Institutional Review Board (IRB) at each site and all subjects signed informed consent.

2.3. Imaging methodologies

2.3.1. Data acquisition—Neuroimaging data were collected at four different sites using 3 Tesla scanners: Columbia University (CU – General Electric scanner), Massachusetts General Hospital (MGH - Siemens scanner), University of Michigan (UM – Philips scanner), and the University of Texas Southwestern (UTSW – Philips scanner). Pseudo-continuous arterial spin labeling (pCASL) protocols were optimized for performance and quality while maintaining similar parameters across sites (whole brain coverage with 29 axial slices, 35 pairs, 2D EPI acquisition, labeling duration = 1516 ms, post-labeling delay = 1500 ms, sequence duration = 5 min 12 s, voxel size: $3.4 \times 3.4 \times 5$ mm^3, TR/TE = 4460/17 msec, FOV = 220 × 220 mm, matrix = 64 × 64, excitation flip angle 90°). In order to reduce variability across scanners the coil sensitivity was corrected by using a reference scan that was part of the scan protocol and the images were corrected for B1 heterogeneities. Labeling efficiency was set at 0.86 across subjects based on a group-average efficiency estimated in our previous study [40]. The labeling plane was positioned at 84 mm below the anterior-posterior line as proposed in the same study [40] and subsequently adapted by the ASL white paper [41]. Furthermore, the pCASL labeling module has been shown to be relatively robust to B0 inhomogeneities. T1-weighted Structural 3D axial images were acquired in the same session. The ASL sequence was acquired at rest with eyes open and fixated on a visual cue [42]. It was the last sequence in one-half hour scan session, which sets a high bar for reliability. However, staff received training to minimize subject head movement.

2.3.2. Arterial spin labeling preprocessing—Perfusion data were preprocessed and analyzed with the Statistical Parametric Mapping software, Version-8 (SPM8, Wellcome Department of Cognitive Neurology, UK; www.fil.ion.ucl.ac.uk/spm) implemented in Matlab7 (Math Works, Natick, MA) and a customized perfusion script to calculate CBF (https://cfn.upenn.edu/perfusion/software.htm, accessed February 17, 2017) [43]. For each subject, functional images were first realigned to correct for head movement. Next, the mean image was co-registered with the anatomical image and segmented. Perfusion weighted image series were then generated by pairwise subtraction of the label and control functional images, followed by conversion to absolute CBF image series based on a single compartment pCASL perfusion model [9,43,44]. M0 representing equilibrium magnetization was used as a global index in the perfusion model, and was determined from a center slice at the level of the thalamus. The M0 calculation was done based on a binary mask that was individually estimated from a center slice at the level of the thalamus. The binary mask was created using a flood fill algorithm that selected an area containing brain tissue and excluded skull and cerebrospinal fluid. Each mask was inspected for quality control. A mean absolute CBF (aCBF) image was generated for each subject, normalized to the Montreal Neurological Institute (MNI) template using parameters derived from the segmentation, resampled to $2 \times 2 \times 2$ mm and smoothed with full-width at half maximum 8 mm kernel.

Relative mean CBF (rCBF) was calculated by dividing the aCBF image by the global individual mean aCBF at each voxel.

2.3.3. Cerebral blood flow quantitative quality assessment—Several factors can contribute to poor data quality of the CBF image measured with ASL (for review, see (41,45)). To assess data quality, we used two quantitative measures generated by an automated calculation system to decrease subjective assessment bias: 1) signal-to-noise-ratio (SNR) between global gray and white matter; and 2) the mean frame-wise displacement calculated for each subject, based on Power's method, which provides a summary statistic of translational and rotational head movement [46]. Subjects with SNR values ≤−2.3 and ≥3.4 and mean framewise displacement > 0.319, which represented outliers for these measures as determined by descriptive analysis, were excluded from further analysis.

2.3.4. Regions of interest—Measures of mean aCBF and rCBF were extracted from regions implicated in emotion and salience processing (amygdala, striatum, insula, dorsal anterior cingulate cortex (ACC)), implicit and effortful emotional regulation (medial prefrontal cortex (mPFC), rostral ACC, ventrolateral prefrontal cortex (vlPFC), and dorsolateral prefrontal cortex (dlPFC)), and the default mode network (mPFC, posterior cingulate cortex). Cingulate regions were defined by Beckman clusters [47] and the ventral striatum was defined based on a neuroimaging meta-analysis [48,49]. Remaining regions were defined by the Wake Forest Toolbox PickAtlas Talairach Daemon template [50,51].

2.4. Analytical method

2.4.1. Intra-class correlation coefficients - across all sites—Within scanner reliability was assessed using intra-class correlation coefficients (ICCs) [52–54] in IBM SPSSv21, a measure widely used in neuroimaging research investigating arterial spin labeling reliability [14,17–23,25–30]. ICCs reflect the ratio of the between subjects variance (variance due to the fact that different subjects participated in the study) and the total variance (which includes within subjects effects and error). Thus, ICC estimates the percent of variance accounted for by between-subjects effects and reflects the consistency of measurements from session 1 to session 2. ICCs were calculated using a two-way mixed model single measure (ICC(3,1)) as in previous ASL reliability studies [14,27]. ICCs were categorized as: Poor (< 0.4), Fair (0.4 ≤ IC-C < 0.6), Good (0.6 ≤ ICC < 0.75), and Excellent (≥0.75) [53]. We calculated ICCs for both aCBF and rCBF images; however, it should be noted that previous work showed higher reliability for rCBF measures [21,55,56]. A comparison of aCBF and rCBF was considered to be significant if confidence intervals (95%) for ICCs did not overlap with each other [18].

2.4.2. Intra-class correlation coefficients - leave-one-site-out—We repeated the above analysis four times, each time leaving one site out to investigate the effect of site. Site was considered to have a significant effect if confidence intervals (95%) for ICCs for any of the sites removed did not overlap with ICCs for the other three sites [18]. We calculated ICCs for both, aCBF and rCBF.

2.4.3. Bland-Altman assessment across sites – gray matter—A Bland-Altman assessment for agreement was used to compare whole-brain gray matter aCBF at each time point and site. A range of agreement was defined as mean bias ± 2 SD [57]. Proportional bias is present when the difference between measures is a function of the mean of the measures. A multiple linear regression analysis with the difference in CBF measure as the dependent variable and the mean CBF measure and site (transformed into dummy variables with site 3 as reference) as independent variables was used to test whether the model coefficient for mean CBF equaled zero. Rejection of that null hypothesis would be consistent with the presence of proportional bias ($p < 0.05$). Gray matter aCBF in young healthy subjects is considered normal between 40 and 100 mL/100 g/min [41].

2.4.4. Assessing the effect of scanner site—To examine the effect of scanner site on whole-brain gray matter aCBF across sessions, we conducted a repeated-measures analysis of variance (ANOVA) with session as a within-subject variable and scanner site as a between-subjects variable. *P* values below 0.05 were considered significant.

3. Results

3.1. Demographic information and quantitative quality assessment

Six subjects were excluded because of low data quality (two subjects had gray/white matter SNR value > 3.4 and two subjects had values lower or equal to −2.3; and three subjects had mean framewise displacement > 0.32). One of these six subjects failed both data quality measurements. Four subjects did not complete both sessions. One subject did not complete both sessions and had low data quality. This yielded a final sample of $N = 31$ (Supplemental Fig. 1). There were no differences in age, gender, and years of education across sites (all *ps* > 0.1; Table 1). However, we found a trend towards significance for marital status ($p = 0.054$) and ethnicity ($p = 0.06$); and a significant difference for race ($p = 0.049$) across sites.

3.2. Absolute and relative cerebral blood flow reliability across all sites

The majority of brain regions had good to excellent aCBF and rCBF reliability, and six regions had fair reliability. The subgenual anterior cingulate and ventral striatum had the highest aCBF and rCBF reliability (excellent - both ICCs > 0.8) and the right amygdala had the lowest aCBF and rCBF reliability (fair - ICC = 0.4 Table 2).

The majority of brain regions had comparable aCBF and rCBF reliability. Four regions had numerically different ICCs between aCBF and rCBF, however, all had a difference lower than 0.2, and overlapping confidence intervals.

3.3. Absolute cerebral blood flow reliability - leave-one-site-out validation

The leave-one-site out analysis did not show a significant effect of site on the reliability of aCBF (Supplemental Table 1). For all regions, the confidence intervals of each of the four pair comparisons overlapped.

3.4. Relative cerebral blood flow reliability - leave-one-site-out validation

The leave-one-site out analysis did not show a significant effect of site on the reliability of rCBF (Supplemental Table 2). For all regions, the confidence intervals of each of the four pair comparisons overlapped.

However, it should be noted that the right amygdala reliability for absolute and relative CBF decreased to 0.02 when site 1 was left out of the analysis, although confidence intervals continued to overlap.

3.5. Bland-Altman assessment across sites – gray matter

Mean whole-brain gray matter aCBF across sites was 45.9 mL/100 g/min (SD = 10.8). Bland–Altman plots together with 95% confidence interval (CI) of CBF measurements from gray matter in the 31 subjects subdivided by each of the four sites are shown in Fig. 1. Gray matter CBF difference plotted against mean gray matter CBF were randomly distributed and showed no dependence on mean gray matter CBF, as can be seen in the Bland–Altman plots shown in Fig. 1. Independent variables (site and mean aCBF measure) assessed with multiple liner regression did not predict change in the dependent variable, gray matter aCBF difference. Therefore, no proportional bias was found across sites ($p > 0.2$ for gray matter aCBF).

3.6. Assessing the effect of scanner site on mean aCBF levels - repeated-measures ANOVA

We observed no significant effect of session ($F(1,27) = 0.1$, $p = 0.8$) or session by site interaction ($F(3,27) = 0.3$, $p = 0.9$) for whole-brain gray matter aCBF. However, there was a significant scanner site effect ($F(3,27) = 6.3$, $p = 0.002$) with site 3 having lower whole brain gray matter aCBF relative to site 1 ($p = 0.034$) and 4 ($p = 0.003$); and a trend for lower aCBF relative to site 2 ($p = 0.08$) (Fig. 2).

4. Discussion

Resting aCBF and rCBF measured with ASL were comparable and demonstrated good to excellent reliability between sessions for the majority of brain regions examined and were stable across sites. In addition, there was no proportional bias within each site for whole-brain gray matter aCBF and the precision of the measures was not different across sites. However, one site had significantly lower aCBF relative to the other sites. These findings support the potential utility of pCASL measures in longitudinal studies, but suggest caution in multi-site studies.

One advantage of CBF measured with ASL is that it allows for calculation of an absolute quantification of CBF. Previous studies highlighted the greater reliability of relative over absolute CBF values when using ASL [21,56]. Relative measures can factor out global intra-subject CBF variation which is a significant source of noise. In the current study, aCBF and rCBF were highly comparable, and were measured with pseudo-continuous ASL, as per the current recommendation for this type of sequence [41]. Furthermore, absolute quantification can be used as a marker for pharmacological activity [13–16], which is useful for both

longitudinal and cross-sectional studies investigating psychopathology such as MDD, which is the focus of the EMBARC study.

Our findings are consistent with previous studies reporting good to excellent test-retest reliability of CBF measures one week apart in the same scanner [13,23,27,32]. However, ASL reliability studies have typically shown decreasing reliability as time between scan sessions increases [14,18–20,23,27,30]. For example, Chen et al. reported coefficients of variability < 11% (i.e. good reliability) within one hour apart and < 15% one week apart in different cerebral artery territories [23]. Similarly, Gevers et al. found excellent ICCs (> 0.85) at whole brain level and in different cerebral artery territories during the same session acquisition and a reduction in ICCs values after one to three weeks (> 0.7) [18]. Our findings indicate that ASL measures have high test-retest reliability, as shown by image acquisition one week apart. We show that this is also the case during the four leave-one-site-out analyses. These findings suggest that variations in ASL measures over time are probably more related to errors caused by subject repositioning, data processing, and physiological fluctuations, rather than scanner instability.

Furthermore, we found that the precision of the whole gray matter aCBF measure between baseline and week 1 was not predicted by a site effect analyzed with multiple linear regression, and that the leave-one-site-out analyses revealed no effect of site on test-retest reliability of aCBF and rCBF measures, in accordance with previous ASL studies examining reliability across different sites, but using the same MRI platforms [20,24,25,29]. Petersen et al. published the largest ASL reliability study to date, with approximately one hundred ninety one subjects in 28 sites, each using a 3 Tesla Philips MRI scanner. Subjects were scanned twice, two weeks apart [20], with reasonable reproducibility across sites, but with a wide range of values (gray matter ICC ranged between 0.07 and 0.81), which the authors explained by a noticeable correlation between mean CBF within the sites and their data acquisition success rate ($r = 0.66$). Furthermore, Gevers et al. reported whole-brain pCASL perfusion varying < 20% over repeated measurements 1–3 weeks apart within the same six subjects across three 3 Tesla Philips MRI scanners [24]. Intra-scanner reliability has been shown to be typically higher than inter-scanner reliability, but the latter was still shown to be high (visual cortex ICC > 0.7 and several regions of interest ICCs between 0.4 and 0.97) across two 3 Tesla GE scanners in two studies by the same group with twelve and eight subjects examined 10–15 days apart [25,29]. Two studies from the same group explored ASL intra and inter scanner reliability across GE and Philip platforms in twenty-two subjects two weeks apart from each session. The first study found low coefficient of variation in global gray matter CBF at rest within platforms (GE: 10.8%; Philips: 11.3%) and between platforms (13.6%) [34]. The second study found good reliability at rest (GE: IC-C = 0.69, Philips: ICC = 0.77) and during finger tapping task (GE: ICC = 0.66, Philips: ICC = 0.75) in the motor cortex. Moreover, the inter scanner reliability was also good at rest (ICC = 0.69) and during motor activation (ICC = 0.61) [33]. Only one small study explored more three 3 Tesla MRI platforms – GE, Philip and Siemens [31]. The authors compared similar pCASL acquisition parameters within the same day in two parallel studies, one with eleven subjects, and a second with fourteen subjects (eleven subjects participated in both studies). The authors found high inter-scanner reliability, albeit with coefficients of variation ranging from 11.2% to 46% depending on the gray matter region, leading the authors to

conclude that it is possible to acquire comparable cerebral CBF maps on scanners of different platforms.

Our current findings add to the literature by demonstrating high reliability of ASL measures across four sites with three different 3 Tesla MRI platforms, using a larger sample and greater time interval between scans than a previous study that investigated inter scanner reliability of ASL measures across three different scanner platforms [27]. Furthermore, mean whole-brain gray matter aCBF for all subjects across sites was between 40 and 100 mL/100 g/min, which is considered to be within the normal range for healthy subjects. Despite similar precision across sites as discussed in the previous paragraph, direct comparison using repeated measures ANOVA among the four different sites and three different scanner platforms revealed significantly lower whole-brain aCBF in one of the sites (i.e. main effect of site but no effect of session nor session by site interaction), which could be explained by a within group difference rather than a scanner site difference since the subjects were not the same across sites. Furthermore, the large range of reliability values across different brain regions, underscores the need for further optimization of ASL acquisition parameters such as inversion efficiency, inversion time, post labeling delay time, background suppression, readout type, as well as SNR, and movement control [1,41,45].

The subgenual ACC and ventral striatum had the highest reliability values. These regions are implicated in the psychopathology of depression [36]. For example, a previous ASL study characterizing CBF in unipolar depression relative to healthy controls found decreased CBF in the subgenual ACC [12]. Moreover, CBF measures with ASL in subgenual ACC have been shown to differentiate unipolar and bipolar depression using pattern recognition analysis [9], and elevated ventral striatal activation has been implicated in bipolar disorder [58,59]. Differentiating between unipolar and bipolar depression is of key importance because it determines the choice of treatment [60]. Studies to ascertain the utility of pCASL for differentiating between these two disorders would thus be instructive.

SNR is a widely used measure to assess data imaging quality (for a review of advancements in ASL imaging parameters designed to optimize signal see (1,41,45)). The pCASL employed in this study has been shown to yield superior SNR compared with pulsed ASL, and is now becoming the sequence of choice to measure CBF [23,24,41,45]. In addition, earlier studies of BOLD sequences showed that movement and other imaging artifacts can significantly impact data quality [61,62]. We assessed data quality using two quantitative measures: gray/white matter SNR and framewise displacement.

Our sample size was relatively small relative to reliability studies of BOLD sequences; however, it is the largest test-retest reliability study of ASL across MRI platforms.

5. Conclusions

This is the first study to examine the test-retest reliability of ASL measures in healthy subjects across multiple sites and MRI platforms. We demonstrated good to excellent test-retest reliability of CBF measured with pCASL in a range of brain regions implicated in emotion and salience processing, emotion regulation, and the default mode network, which

have been previously linked to MDD symptomatology. Given its high reliability, ASL is a promising neuroimaging technique for longitudinal studies of psychiatric disorders and a useful tool to identify biomarkers reflecting neuropathophysiologic processes and neural predictors of treatment response.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under award numbers U01MH092221 and U01MH092250. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work was supported by the EMBARC National Coordinating Center at UT Southwestern Medical Center and Data Center at Columbia University.

Conflict of interest statements

Dr. Almeida is part funded by NIMH (1R25 MH101076). Dr. Deckersbach's research has been funded by NIH, NIMH, NARSAD, TSA, IOCDF, Tufts University, DBDAT, Otsuka Pharmaceuticals and Cogito, Inc. He has received honoraria, consultation fees and/or royalties from the MGH Psychiatry Academy, BrainCells Inc., Clintara, LLC., Systems Research and Applications Corporation, Boston University, the Catalan Agency for Health Technology Assessment and Research, the National Association of Social Workers Massachusetts, the Massachusetts Medical Society, Tufts University, NIDA, NIMH, and Oxford University Press. He has also participated in research funded by DARPA, NIH, NIMH, NIA, AHRQ, PCORI, Janssen Pharmaceuticals, The Forest Research Institute, Shire Development Inc., Medtronic, Cyberonics, Northstar, and Takeda. Dr. Madhukar H. Trivedi, is or has been an advisor/consultant and received fee from (lifetime disclosure): Abbott Laboratories, Inc., Abdi Ibrahim, Akzo (Organon Pharmaceuticals Inc.), Alkermes, AstraZeneca, Axon Advisors, Bristol-Myers Squibb Company, Cephalon, Inc., Cerecor, CME Institute of Physicians,Concert Pharmaceuticals, Inc., Eli Lilly & Company, Evotec, Fabre Kramer Pharmaceuticals, Inc., Forest Pharmaceuticals, GlaxoSmithKline, Janssen Global Services, LLC, Janssen Pharmaceutica Products, LP, Johnson & Johnson PRD, Libby, Lundbeck, Meade Johnson, MedAvante, Medtronic, Merck, Mitsubishi Tanabe Pharma Development America, Inc., Naurex, Neuronetics, Otsuka Pharmaceuticals, Pamlab, Parke-Davis Pharmaceuticals, Inc., Pfizer Inc., PgxHealth, Phoenix Marketing Solutions, Rexahn Pharmaceuticals, Ridge Diagnostics, Roche Products Ltd., Sepracor, SHIRE Development, Sierra, SK Life and Science, Sunovion, Takeda, Tal Medical/Puretech Venture, Targacept, Transcept, VantagePoint, Vivus, and Wyeth-Ayerst Laboratories. In addition, he has received grants/research support from: Agency for Healthcare Research and Quality (AHRQ), Cyberonics, Inc., National Alliance for Research in Schizophrenia and Depression, National Institute of Mental Health and National Institute on Drug Abuse. Dr. Pat McGrath has received funding from the National Institute of Mental Health, New York State Department of Mental Hygiene, Research Foundation for Mental Hygiene (New York State), Forest Research Laboratories, Sunovion Pharmaceuticals, and Naurex Pharmaceuticals (now Allergan). Dr. Fava has received research support from Abbot

Laboratories; Alkermes, Inc.; American Cyanamid;Aspect Medical Systems; AstraZeneca; Avanir Pharmaceuticals; BioResearch; BrainCells Inc.; Bristol-Myers Squibb; CeNeRx BioPharma; Cephalon; Clintara, LLC; Cerecor; Covance; Covidien; Eli Lilly and Company; EnVivo Pharmaceuticals, Inc.; Euthymics Bioscience, Inc.; Forest Pharmaceuticals, Inc.; Ganeden Biotech, Inc.; GlaxoSmithKline; Harvard Clinical Research Institute; Hoffman-LaRoche; Icon Clinical Research; i3 Innovus/Ingenix; Janssen R & D, LLC; Jed Foundation; Johnson & Johnson Pharmaceutical Research & Development; Lichtwer Pharma GmbH; Lorex Pharmaceuticals; Lundbeck Inc.; MedAvante; Methylation Sciences Inc.; National Alliance for Research on Schizophrenia & Depression (NARSAD); National Center for Complementary and Alternative Medicine (NCCAM); National Institute of Drug Abuse (NIDA); National Institute of Mental Health (NIMH); Neuralstem, Inc.; Novartis AG; Organon Pharmaceuticals; PamLab, LLC.; Pfizer Inc.; Pharmacia-Upjohn; Pharmaceutical Research Associates., Inc.; Pharmavite® LLC; PharmoRx Therapeutics; Photothera; Reckitt Benckiser; Roche Pharmaceuticals; RCT Logic, LLC (formerly Clinical Trials Solutions, LLC); Sanofi-Aventis US LLC; Shire; Solvay Pharmaceuticals, Inc.; Stanley Medical Research Institute (SMRI); Synthelabo; Tal Medical; Wyeth-Ayerst Laboratories; he has served as advisor or consultant to Abbott Laboratories; Acadia; Affectis Pharmaceuticals AG; Alkermes, Inc.; Amarin Pharma Inc.; Aspect Medical Systems; AstraZeneca; Auspex Pharmaceuticals; Avanir Pharmaceuticals; AXSOME Therapeutics; Bayer AG; Best Practice Project Management, Inc.; Biogen; BioMarin Pharmaceuticals, Inc.; Biovail Corporation; BrainCells Inc.; Bristol-Myers Squibb; CeNeRx BioPharma; Cephalon, Inc.; Cerecor; CNS Response, Inc.; Compellis Pharmaceuticals; Cypress Pharmaceutical, Inc.; DiagnoSearch Life Sciences (P) Ltd.; Dinippon Sumitomo Pharma Co. Inc.; Dov Pharmaceuticals, Inc.; Edgemont Pharmaceuticals, Inc.; Eisai Inc.; Eli Lilly and Company; EnVivo Pharmaceuticals, Inc.; ePharmaSolutions; EPIX Pharmaceuticals, Inc.; Euthymics Bioscience, Inc.; Fabre-Kramer Pharmaceuticals, Inc.; Forest Pharmaceuticals, Inc.; Forum Pharmaceuticals; GenOmind, LLC; GlaxoSmithKline; Grunenthal GmbH; i3 Innovus/ Ingenis; Intracellular; Janssen Pharmaceutica; Jazz Pharmaceuticals, Inc.; Johnson & Johnson Pharmaceutical Research & Development, LLC; Knoll Pharmaceuticals Corp.; Labopharm Inc.; Lorex Pharmaceuticals; Lundbeck Inc.; MedAvante, Inc.; Merck & Co., Inc.; MSI Methylation Sciences, Inc.; Naurex, Inc.; Nestle Health Sciences; Neuralstem, Inc.; Neuronetics, Inc.; NextWave Pharmaceuticals; Novartis AG; Nutrition 21; Orexigen Therapeutics, Inc.; Organon Pharmaceuticals; Osmotica; Otsuka Pharmaceuticals; Pamlab, LLC.; Pfizer Inc.; PharmaStar; Pharmavite® LLC.; PharmoRx Therapeutics; Precision Human Biolaboratory; Prexa Pharmaceuticals, Inc.; Puretech Ventures; PsychoGenics; Psylin Neurosciences, Inc.; RCT Logic, LLC Formerly Clinical Trials Solutions, LLC); Rexahn Pharmaceuticals, Inc.; Ridge Diagnostics, Inc.; Roche; Sanofi-Aventis US LLC.; Sepracor Inc.; Servier Laboratories; Schering-Plough Corporation; Solvay Pharmaceuticals, Inc.; Somaxon Pharmaceuticals, Inc.; Somerset Pharmaceuticals, Inc.; Sunovion Pharmaceuticals; Supernus Pharmaceuticals, Inc.; Synthelabo; Taisho Pharmaceutical; Takeda Pharmaceutical Company Limited; Tal Medical, Inc.; Tetragenex Pharmaceuticals, Inc.; TransForm Pharmaceuticals, Inc.; Transcept Pharmaceuticals, Inc.; Vanda Pharmaceuticals, Inc.; VistaGen; he has received speaking or publishing fees from Adamed, Co; Advanced Meeting Partners; American Psychiatric Association; American Society of Clinical Psychopharmacology; AstraZeneca; Belvoir Media Group; Boehringer Ingelheim

GmbH; Bristol-Myers Squibb; Cephalon, Inc.; CME Institute/Physicians Postgraduate Press, Inc.; Eli Lilly and Company; Forest Pharmaceuticals, Inc.; GlaxoSmithKline; Imedex, LLC; MGH Psychiatry Academy/Primedia; MGH Psychiatry Academy/Reed Elsevier; Novartis AG; Organon Pharmaceuticals; Pfizer Inc.; PharmaStar; United BioSource,Corp.; Wyeth-Ayerst Laboratories; he has equity holdings in Compellis and PsyBrain, Inc.; he has a patent for Sequential Parallel Comparison Design (SPCD), which are licensed by MGH to Pharmaceutical Product Development, LLC (PPD); and patent application for a combination of Ketamine plus Scopolamine in Major Depressive Disorder (MDD), licensed by MGH to Biohaven; and he receives copyright royalties for the MGH Cognitive & Physical Functioning Questionnaire (CPFQ), Sexual Functioning Inventory (SFI), Antidepressant Treatment Response Questionnaire (ATRQ), Discontinuation-Emergent Signs & Symptoms (DESS), Symptoms of Depression Questionnaire (SDQ), and SAFER; Lippincott, Williams & Wilkins; Wolkers Kluwer; World Scientific Publishing Co. Pte.Ltd. Dr. Ramin Parsey reports no relevant or material financial interests that relate to the research described in this paper. Dr. Benji Kurian has received research grant support from the following organizations: Targacept, Inc., Pfizer, Inc., Johnson & Johnson, Evotec, Rexahn, Naurex, Forest Pharmaceuticals and the National Institute of Mental Health (NIMH). Mary L. Phillips has received funding from NIMH; the Emmerling-Pittsburgh Foundation; and Roche Pharmaceuticals. Dr. Oquendo receives royalties for use of the Columbia Suicide Severity Rating Scale. Her family owns stock in Bristol Myers Squibb. Drs. Jorge Almeida, Crystal Cooper, Phil Adams, Jay Fournier, Tsafrir Greenberg, Hanzhang Lu, Thomas Carmody, and Henry Chase do not report any conflicting interests. Dr. Melvin McInnis has no conflicts of interest with respect to this paper. In the past two years, Dr. Myrna Weissman received funding from the National Institute of Mental Health (NIMH), the National Institute on Drug Abuse (NIDA), the National Alliance for Research on Schizophrenia and Depression (NARSAD), the Sackler Foundation, the Templeton Foundation; and receives royalties from the Oxford University Press, Perseus Press, the American Psychiatric Association Press, and MultiHealth Systems.

References

1. Detre JA, Rao H, Wang DJ, Chen YF, Wang Z. Applications of arterial spin labeled MRI in the brain. J Magn Reson Imaging. 2012; 35(5):1026–37. [PubMed: 22246782]
2. Xu G, Rowley HA, Wu G, et al. Reliability and precision of pseudo-continuous arterial spin labeling perfusion MRI on 3.0 T and comparison with (15)O-water PET in elderly subjects at risk for Alzheimer's disease. NMR Biomed. 2009
3. Ye FQ, Berman KF, Ellmore T, et al. H(2)(15)O PET validation of steady-state arterial spin tagging cerebral blood flow measurements in humans. Magn Reson Med. 2000; 44(3):450–6. [PubMed: 10975898]
4. Fan AP, Jahanian H, Holdsworth SJ, Zaharchuk G. Comparison of cerebral blood flow measurement with [15O]-water positron emission tomography and arterial spin labeling magnetic resonance imaging: a systematic review. J Cereb Blood Flow Metab. 2016
5. Bruns A, Kunnecke B, Risterucci C, Moreau JL, von Kienlin M. Validation of cerebral blood perfusion imaging as a modality for quantitative pharmacological MRI in rats. Magn Reson Med. 2009; 61(6):1451–8. [PubMed: 19358231]
6. Andersen JB, Henning WS, Lindberg U, et al. Positron emission tomography/magnetic resonance hybrid scanner imaging of cerebral blood flow using (15)O-water positron emission tomography

and arterial spin labeling magnetic resonance imaging in newborn piglets. J Cereb Blood Flow Metab. 2015; 35(11):1703–10. [PubMed: 26058699]

7. Gillihan SJ, Rao H, Wang J, et al. Serotonin transporter genotype modulates amygdala activity during mood regulation. Soc Cogn Affect Neurosci. 2010; 5(1):1–10. [PubMed: 19858108]
8. Lim J, Wu W, Wang J, Detre JA, Dinges DF, Rao H. Imaging brain fatigue from sustained mental workload: an ASL perfusion study of the time-on-task effect. Neuroimage. 2010; 49(4):3426–35. [PubMed: 19925871]
9. Almeida JR, Mourao-Miranda J, Aizenstein HJ, et al. Pattern recognition analysis of anterior cingulate cortex blood flow to classify depression polarity. Br J Psychiatry. 2013; 203(3):310–1. [PubMed: 23969484]
10. Clark CP, Brown GG, Archibald SL, et al. Does amygdalar perfusion correlate with antidepressant response to partial sleep deprivation in major depression? Psychiatry Res. 2006; 146(1):43–51. [PubMed: 16380239]
11. Chao LLP, Buckley STBA, Kornak JP, et al. ASL perfusion MRI predicts cognitive decline and conversion from MCI to dementia. Alzheimer Dis Assoc Disord. 2010; 24(1):19–27. [Article]. [PubMed: 20220321]
12. Ota M, Noda T, Sato N, et al. Characteristic distributions of regional cerebral blood flow changes in major depressive disorder patients: a pseudo-continuous arterial spin labeling (pCASL) study. J Affect Disord. 2014; 165:59–63. [PubMed: 24882178]
13. Chen Y, Wan HI, O'Reardon JP, et al. Quantification of cerebral blood flow as biomarker of drug effect: arterial spin labeling phMRI after a single dose of oral citalopram. Clin Pharmacol Ther. 2011; 89(2):251–8. [PubMed: 21191380]
14. Klomp A, Caan MW, Denys D, Nederveen AJ, Reneman L. Feasibility of ASL-based phMRI with a single dose of oral citalopram for repeated assessment of serotonin function. Neuroimage. 2012; 63(3):1695–700. [PubMed: 22842212]
15. Chen Y, Pressman P, Simuni T, Parrish TB, Gitelman DR. Effects of acute levodopa challenge on resting cerebral blood flow in Parkinson's disease patients assessed using pseudo-continuous arterial spin labeling. PeerJ. 2015; 3:e1381. [PubMed: 26734502]
16. Kaichi Y, Okada G, Takamura M, et al. Changes in the regional cerebral blood flow detected by arterial spin labeling after 6-week escitalopram treatment for major depressive disorder. J Affect Disord. 2016; 194:135–43. [PubMed: 26826533]
17. Jahng GH, Song E, Zhu XP, Matson GB, Weiner MW, Schuff N. Human brain: reliability and reproducibility of pulsed arterial spin-labeling perfusion MR imaging. Radiology. 2005; 234(3): 909–16. [PubMed: 15734942]
18. Gevers S, Majoie CB, van den Tweel XW, Lavini C, Nederveen AJ. Acquisition time and reproducibility of continuous arterial spin-labeling perfusion imaging at 3T. AJNR Am J Neuroradiol. 2009; 30(5):968–71. [PubMed: 19193760]
19. Jiang L, Kim M, Chodkowski B, et al. Reliability and reproducibility of perfusion MRI in cognitively normal subjects. Magn Reson Imaging. 2010; 28(9):1283–9. [PubMed: 20573464]
20. Petersen ET, Mouridsen K, Golay X. The QUASAR reproducibility study, part II: results from a multi-center arterial spin labeling test-retest study. Neuroimage. 2010; 49(1):104–13. [PubMed: 19660557]
21. Pfefferbaum A, Chanraud S, Pitel AL, et al. Volumetric cerebral perfusion imaging in healthy adults: regional distribution, laterality, and repeatability of pulsed continuous arterial spin labeling (PCASL). Psychiatry Res. 2010; 182(3):266–73. [PubMed: 20488671]
22. Xu G, Rowley HA, Wu G, et al. Reliability and precision of pseudo-continuous arterial spin labeling perfusion MRI on 3.0 T and comparison with 15O-water PET in elderly subjects at risk for Alzheimer's disease. NMR Biomed. 2010; 23(3):286–93. [PubMed: 19953503]
23. Chen Y, Wang DJ, Detre JA. Test-retest reliability of arterial spin labeling with common labeling strategies. J Magn Reson Imaging. 2011; 33(4):940–9. [PubMed: 21448961]
24. Gevers S, van Osch MJ, Bokkers RP, et al. Intra- and multicenter reproducibility of pulsed, continuous and pseudo-continuous arterial spin labeling methods for measuring cerebral perfusion. J Cereb Blood Flow Metab. 2011; 31(8):1706–15. [PubMed: 21304555]

25. Huang D, Wu B, Shi K, Ma L, Cai Y, Lou X. Reliability of three-dimensional pseudo-continuous arterial spin labeling MR imaging for measuring visual cortex perfusion on two 3T scanners. PLoS One. 2013; 8(11):e79471. [PubMed: 24278137]
26. Kilroy E, Apostolova L, Liu C, Yan L, Ringman J, Wang DJ. Reliability of two-dimensional and three-dimensional pseudo-continuous arterial spin labeling perfusion MRI in elderly populations: comparison with 15O-water positron emission tomography. J Magn Reson Imaging. 2014; 39(4): 931–9. [PubMed: 24038544]
27. Mezue M, Segerdahl AR, Okell TW, Chappell MA, Kelly ME, Tracey I. Optimization and reliability of multiple postlabeling delay pseudo-continuous arterial spin labeling during rest and stimulus-induced functional task activation. J Cereb Blood Flow Metab. 2014; 34(12):1919–27. [PubMed: 25269517]
28. Tatewaki Y, Higano S, Taki Y, et al. Regional reliability of quantitative signal targeting with alternating radiofrequency (STAR) labeling of arterial regions (QUASAR). J Neuroimaging. 2014; 24(6):554–61. [PubMed: 25370338]
29. Wu B, Lou X, Wu X, Ma L. Intra- and interscanner reliability and reproducibility of 3D whole-brain pseudo-continuous arterial spin-labeling MR perfusion at 3T. J Magn Reson Imaging. 2014; 39(2):402–9. [PubMed: 23723043]
30. Jann K, Gee DG, Kilroy E, et al. Functional connectivity in BOLD and CBF data: similarity and reliability of resting brain networks. Neuroimage. 2015; 106:111–22. [PubMed: 25463468]
31. Mutsaerts HJ, van Osch MJ, Zelaya FO, et al. Multi-vendor reliability of arterial spin labeling perfusion MRI using a near-identical sequence: implications for multi-center studies. Neuroimage. 2015; 113:143–52. [PubMed: 25818685]
32. Sousa I, Vilela P, Figueiredo P. Reproducibility of the quantification of arterial and tissue contributions in multiple postlabeling delay arterial spin labeling. J Magn Reson Imaging. 2014; 40(6):1453–62. [PubMed: 24227019]
33. Steketee RM, Mutsaerts HJ, Bron EE, et al. Quantitative functional arterial spin labeling (fASL) MRI–sensitivity and reproducibility of regional CBF changes using pseudo-continuous ASL product sequences. PLoS One. 2015; 10(7):e0132929. [PubMed: 26172381]
34. Mutsaerts HJ, Steketee RM, Heijtel DF, et al. Inter-vendor reproducibility of pseudo-continuous arterial spin labeling at 3 tesla. PLoS One. 2014; 9(8):e104108. [PubMed: 25090654]
35. Trivedi MH, McGrath PJ, Fava M, et al. Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): rationale and design. J Psychiatr Res. 2016; 78:11–23. [PubMed: 27038550]
36. Phillips ML, Chase HW, Sheline YI, et al. Identifying predictors, moderators, and mediators of antidepressant response in major depressive disorder: neuroimaging approaches. Am J Psychiatry. 2015; 172(2):124–38. [PubMed: 25640931]
37. Chase HW, Fournier JC, Greenberg T, et al. Accounting for dynamic fluctuations across time when examining fMRI test-retest reliability: analysis of a reward paradigm in the EMBARC study. PLoS One. 2015; 10(5):e0126326. [PubMed: 25961712]
38. Greenberg, T., Chase, HW., Almeida, JR., et al. Moderation of the relationship between reward expectancy and prediction error-related ventral striatal reactivity by anhedonia in unmedicated major depressive disorder: findings from the EMBARC study. Am J Psychiatry. 2015. (Epub ahead of print)http://dx.doi.org/10.1176/appi.ajp.2015.14050594
39. Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry. 2003; 54(5):573–83. [PubMed: 12946886]
40. Aslan S, Xu F, Wang PL, et al. Estimation of labeling efficiency in pseudocontinuous arterial spin labeling. Magn Reson Med. 2010; 63(3):765–71. [PubMed: 20187183]
41. Alsop DC, Detre JA, Golay X, et al. Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: a consensus of the ISMRM perfusion study group and the European consortium for ASL in dementia. Magn Reson Med. 2015; 73(1):102–16. [PubMed: 24715426]

42. Zou Q, Miao X, Liu D, Wang DJ, Zhuo Y, Gao JH. Reliability comparison of spontaneous brain activities between BOLD and CBF contrasts in eyes-open and eyes-closed resting states. Neuroimage. 2015; 121:91–105. [PubMed: 26226087]
43. Wang Z, Aguirre GK, Rao H, et al. Empirical optimization of ASL data analysis using an ASL data processing toolbox: ASLtbx. Magn Reson Imaging. 2008; 26(2):261–9. [PubMed: 17826940]
44. Liu P, Uh J, Lu H. Determination of spin compartment in arterial spin labeling MRI. Magn Reson Med. 2011; 65(1):120–7. [PubMed: 20740655]
45. Amukotuwa SA, Yu C, Zaharchuk G. 3D Pseudocontinuous arterial spin labeling in routine clinical practice: a review of clinically significant artifacts. J Magn Reson Imaging. 2016; 43(1):11–27. [PubMed: 25857715]
46. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage. 2012; 59(3):2142–54. [PubMed: 22019881]
47. Beckmann M, Johansen-Berg H, Rushworth MF. Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. J Neurosci. 2009; 29(4):1175–90. [PubMed: 19176826]
48. Di Martino A, Scheres A, Margulies DS, et al. Functional connectivity of human striatum: a resting state FMRI study. Cereb Cortex. 2008; 18(12):2735–47. [PubMed: 18400794]
49. Postuma RB, Dagher A. Basal ganglia functional connectivity based on a meta-analysis of 126 positron emission tomography and functional magnetic resonance imaging publications. Cereb Cortex. 2006; 16(10):1508–21. [PubMed: 16373457]
50. Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage. 2003; 19(3):1233–9. [PubMed: 12880848]
51. Lancaster JL, Woldorff MG, Parsons LM, et al. Automated Talairach atlas labels for functional brain mapping. Hum Brain Mapp. 2000; 10(3):120–31. [PubMed: 10912591]
52. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979; 86(2):420–8. [PubMed: 18839484]
53. Fleiss, JL.Levin, B.Paik, MC., Hoboken, NJ., editors. Statistical methods for rates and proportions. Vol. xxvii. J Wiley; 2003. p. 760
54. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods. 1996; 1(1):30–46.
55. Aslan S, Xu F, Wang PL, et al. Estimation of labeling efficiency in pseudocontinuous arterial spin labeling. Magn Reson Med. 2010; 63(3):765–71. [PubMed: 20187183]
56. Aslan S, Lu H. On the sensitivity of ASL MRI in detecting regional differences in cerebral blood flow. Magn Reson Imaging. 2010; 28(7):928–35. [PubMed: 20423754]
57. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. J Biopharm Stat. 2007; 17(4):571–82. [PubMed: 17613642]
58. Caseras X, Lawrence NS, Murphy K, Wise RG, Phillips ML. Ventral striatum activity in response to reward: differences between bipolar I and II disorders. Am J Psychiatry. 2013; 170(5):533–41. [PubMed: 23558337]
59. Nusslock R, Almeida JR, Forbes EE, et al. Waiting to win: elevated striatal and orbitofrontal cortical activity during reward anticipation in euthymic bipolar disorder adults. Bipolar Disord. 2012; 14(3):249–60. [PubMed: 22548898]
60. Almeida JR, Phillips ML. Distinguishing between unipolar depression and bipolar depression: current and future clinical and neuroimaging perspectives. Biol Psychiatry. 2013; 73(2):111–8. [PubMed: 22784485]
61. Van Dijk KR, Sabuncu MR, Buckner RL. The influence of head motion on intrinsic functional connectivity MRI. Neuroimage. 2012; 59(1):431–8. [PubMed: 21810475]
62. Chen JJ, Jann K, Wang DJ. Characterizing resting-state brain function using arterial spin labeling. Brain Connect. 2015; 5(9):527–42. [PubMed: 26106930]

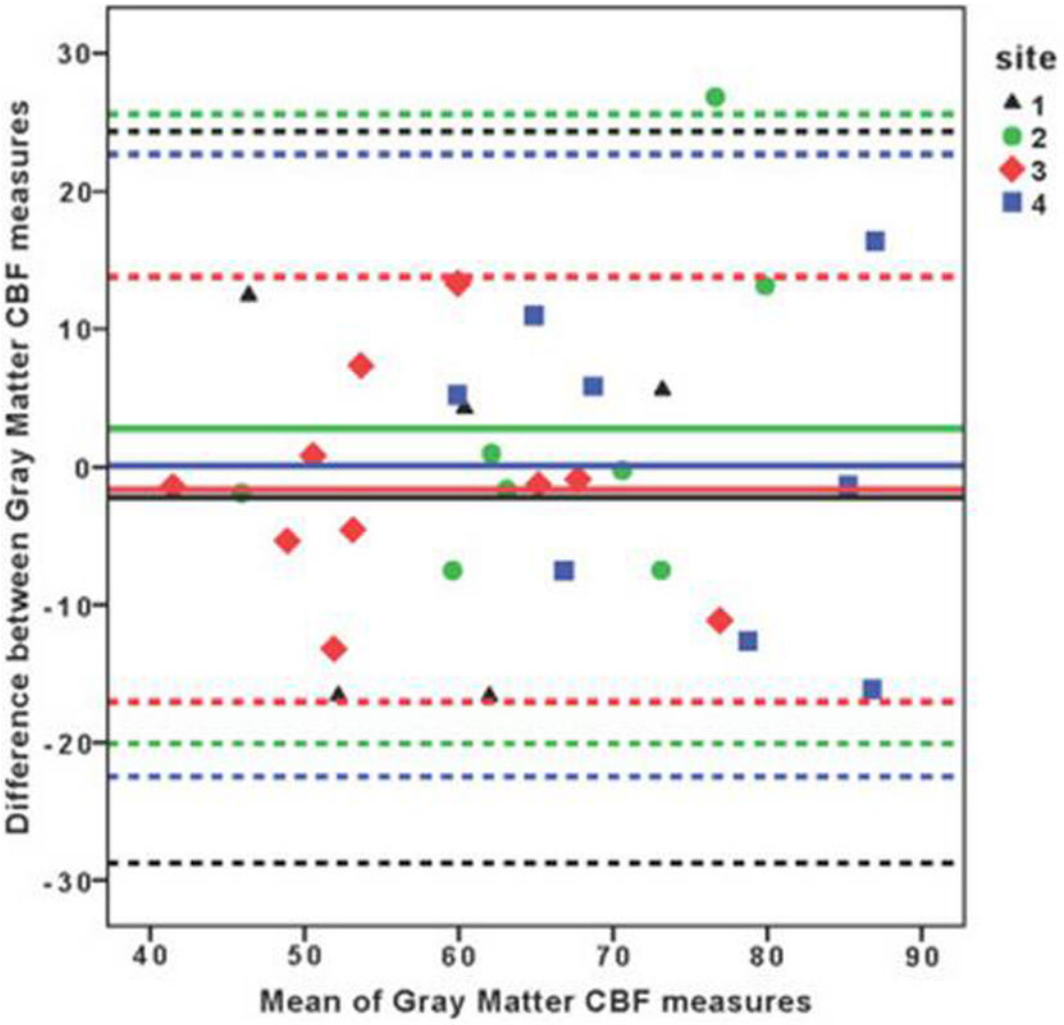

Fig. 1.

Bland-Altman plot of repeated CBF measurements (baseline and week 1) across four sites in absolute whole-brain gray matter CBF (mL/100 g/min). The Bland-Altman plot represents intrascan aCBF differences plotted against mean aCBF. Intrascan aCBF differences are randomly distributed and are not dependent on mean aCBF. Multiple linear regression demonstrated no similarity difference between sites. The solid middle line denotes the bias from zero, which represents absolute no difference between baseline and week 1 measures. The dotted top and bottom lines represents limits of agreement (i.e. bias ± 1.96 × standard deviation). Different colors and symbols represent each of the four sites. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

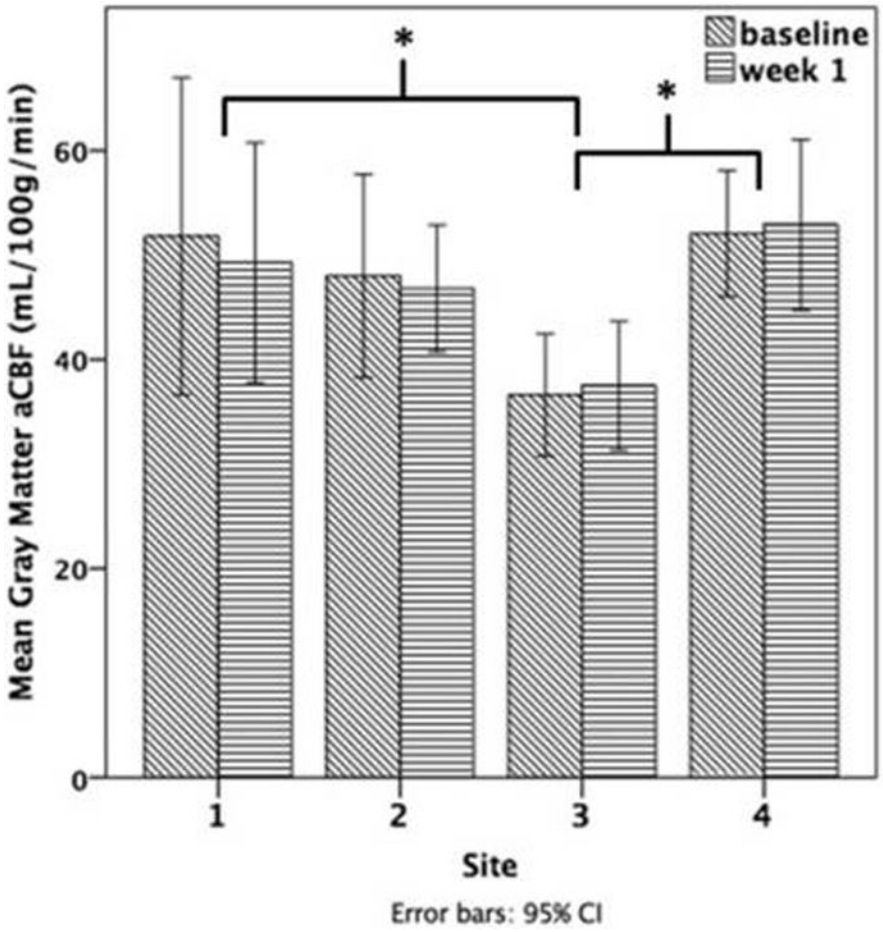

Fig. 2.

Bar graph representing mean absolute whole-brain gray matter CBF (mL/100 g/min) at baseline and week 1 for each site. Each cluster represents a specific site. Each bar constitutes either baseline (diagonal lines) or week 1 (horizontal lines) mean whole-brain gray matter aCBF. Asterisks denote a significant difference ($p < 0.05$) between sites resulting from a mean effect of site in the repeated measures ANOVA.

Table 1

Demographic information across sites.

Site										
		1 (N = 5)		2 (*N* = 8)		3 (*N* = 10)		4 (N = 8)		Total (N = 31)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age		30.0	8.9	45.6	19.1	31.9	11.0	43.4	16.8	38.1	15.6
Year of education		13.6	3.0	17.8	8.7	14.9	2.4	16.4	2.3	15.8	4.9
		N = 5	%	N = 8	%	N = 10	%	N = 8	%	N = 31	%
Gender	Male	1	20.0%	1	12.5%	5	50.0%	3	25.0%	9	29.0%
	Female	4	80.0%	7	87.5%	5	50.0%	6	75.0%	22	71.0%
Race	White	2	40.0%	4	50.0%	7	70.0%	6	75.0%	19	61.3%
	African American	0	0.0%	2	25.0%	3	30.0%	2	25.0%	7	22.6%
	Asian	1	20.0%	2	25.0%	0	0.0%	0	0.0%	3	9.7%
	Other	2	40.0%	0	0.0%	0	0.0%	0	0.0%	2	6.5%
Hispanic	No	3	60.0%	7	87.5%	10	100.0%	8	100.0%	28	90.3%
	Yes	2	40.0%	1	12.5%	0	0.0%	0	0.0%	3	9.7%
Marital Status	Single	5	100.0%	4	50.0%	8	80.0%	3	37.5%	20	64.5%
	Married	0	0.0%	1	12.5%	2	20.0%	4	50.0%	7	22.6%
	Divorced	0	0.0%	3	37.5%	0	0.0%	1	12.5%	4	12.9%

Age F = 2.1; $p = 0.1$; Years of Education F = 0.9; $p = 0.5$; Gender $x^2 = 3.5$; $p = 0.3$; Race $x^2 = 17.$; p = 0.05; Hispanic $x^2 = 7.3$; p = 0.06; Marital Status $x^2 = 12.4$; $p = 0.054$.

Table 2

Absolute and Relative Cerebral Blood Flow Reliability Across All Sites.

	Absolute CBF			Relative CBF		
	ICC	95% CI		ICC	95% CI	
		Lower bound	Upper bound		Lower bound	Upper bound
sgACC	0.81	0.64	0.90	0.81	0.64	0.90
rACC	0.76	0.56	0.88	0.76	0.56	0.88
dACC	0.78	0.59	0.89	0.78	0.58	0.89
mCC	0.69	0.45	0.84	0.69	0.45	0.84
pCC	0.67	0.42	0.83	0.67	0.42	0.83
Left amygdala	0.59	0.30	0.78	0.59	0.30	0.78
Right amygdala	0.42	0.09	0.67	0.42	0.09	0.67
Left dlPFC	0.68	0.43	0.83	0.68	0.43	0.83
Right dlPFC	0.69	0.45	0.84	0.69	0.45	0.84
Left vlPFC	0.57	0.28	0.77	0.57	0.28	0.77
Right vlPFC	0.54	0.24	0.75	0.54	0.23	0.75
Left caudate body	0.66	0.41	0.82	0.46	0.14	0.70
Right caudate body	0.49	0.18	0.72	0.68	0.44	0.83
Left insula	0.68	0.43	0.83	0.68	0.43	0.83
Right insula	0.59	0.30	0.78	0.59	0.30	0.78
Left mPFC	0.71	0.47	0.85	0.71	0.47	0.85
Right mPFC	0.69	0.45	0.84	0.69	0.45	0.84
Left putamen	0.64	0.37	0.81	0.64	0.37	0.81
Right putamen	0.64	0.37	0.81	0.65	0.39	0.81
Left VS	0.80	0.63	0.90	0.72	0.50	0.86
Right VS	0.72	0.50	0.86	0.80	0.63	0.90

ICC: intraclass correlation coefficient; CI: confidence interval; no color = poor reliability (< 0.40); red color = fair reliability (0.40 < ICC < 0.60); blue color = good reliability (0.60 < ICC < 0.75); green color = excellent reliability (> 0.75); ACC: anterior cingulate cortex; sgACC: subgenual ACC; rACC: rostral ACC; dACC: dorsal ACC; CC: cingulate cortex; mCC: middle CC; pCC: posterior CC; PFC: prefrontal cortex; dlPFC: dorsolateral PFC; vlPFC: ventrolateral PFC; mPFC: medial PFC; VS: ventral striatum.