Feasibility of ASL-based phMRI with a single dose of oral citalopram for repeated assessment of serotonin function

A. Klomp a,b,⁎, M.W.A. Caan a,b, D. Denys a,c, A.J. Nederveen a,b, L. Reneman a,b

a Brain Imaging Centre, Academic Medical Centre, Amsterdam, The Netherlands
b Department of Radiology, Academic Medical Centre, Amsterdam, The Netherlands
c Department of Psychiatry, Academic Medical Centre, Amsterdam, The Netherlands

A R T I C L E   I N F O

Article history:
Accepted 18 July 2012
Available online 27 July 2012

Keywords:
Pharmacological MRI
Arterial spin labeling
5-HT
SSRIs
Repeatability

A B S T R A C T

Assessment of cerebral serotonin (5-HT) function with arterial spin labeling (ASL)-based pharmacological magnetic resonance imaging (phMRI) could be a highly useful tool in clinical psychiatric research. The goal of this study was to verify the reliability of ASL-based phMRI after an oral challenge of a selective serotonin reuptake inhibitor (SSRI) in repeated assessment of cerebral 5-HT function. In a placebo-controlled, within-subject crossover study we investigated the effect of a single oral dose of citalopram on brain cerebral blood flow (CBF) using a pulsed ASL sequence (PASL) in twelve female healthy volunteers. The within-session repeatability of the PASL signal was good for all regions tested (wsCV<15%). Both ROI- and voxel-based analyses revealed small but significant effects of a citalopram challenge on CBF values in 5-HT rich brain regions, among which the frontal gyrus and thalamus. These effects could however not be replicated between sessions, most probably due to the small effect size of the oral citalopram challenge on cerebral blood flow. We therefore conclude that the test–retest reliability of PASL phMRI with an oral citalopram challenge is low, limiting the technique's sensitivity to time-dependent changes and consequently its use as a (clinical) research tool.

© 2012 Elsevier Inc. All rights reserved.

Introduction

Impaired serotonin (5-HT) neurotransmission is hypothesized to be involved in a number of psychiatric disorders, such as mood-, anxiety- and impulse control disorders. In vivo assessment of the human brain 5-HT function may help to improve our understanding of the underlying mechanism of these disorders. Although changes in the 5-HT system are typically visualized using positron emission tomography (PET) (Aznavour et al., 2006) and single photon emission computed tomography (SPECT) (Hwang et al., 2007), these techniques have shortcomings. First, they are not ideal research tools because of the radiation exposure involved. Second, PET and SPECT suffer from relatively low spatial resolution. Besides these practical limitations, until now studies using 5-HT radioligands have been less successful and are (consequently) less abundant than studies with for example dopaminergic radioligands (Paterson et al., 2011; Saulin et al., 2012; Smith et al., 2002). Thus, there is a clear need for novel, non-invasive methods to assess 5-HT function in humans and especially in psychiatric patients. A relatively new and promising approach for determining neurotransmitter function is the use of pharmacological MRI (phMRI) (Leslie and James, 2000). With phMRI, neurotransmitter function is modulated pharmacologically using a psychotropic drug while brain functional MRI is simultaneously collected. This is a non-invasive method, which offers a good spatial and temporal resolution allowing longitudinal studies to follow disease progression and/or treatment efficacy without using harmful ionizing radiations. It has already been proven to be a useful method for assessing 5-HT function in preclinical studies (Anderson et al., 2008; Klomp et al., 2012). Up to now, most phMRI studies examined drug effects using intravenous administration or following a chronic drug pre-treatment. Obtaining the same information using a single oral dosage of the challenging drug would be of interest since it is less bothersome to the patient.

BOLD (blood oxygen-level dependent) contrast is the most frequently applied acquisition method in functional MRI because of its ease of data acquisition (Zhang et al., 2006). It is nevertheless not a quantitative measure and known to vary over time (Aguirre et al., 1998). In order to measure a reliable BOLD signal change, the challenging drug needs to enter the brain in rapid and consistent manner during a single scan session (Anderson et al., 2008). Administering a drug orally and subsequently measuring changes in (resting-state) MRI BOLD signal is less feasible considering the long mean time to maximum concentration (Tmax) values oral psychotropic drugs typically have (in the order of hours). This problem may be overcome by using CBF- (cerebral blood flow) or CBV- (cerebral blood volume) based techniques, such as arterial spin labeling (ASL) (Dijkhuizen and Nicolay, 2003; Shah and Marsden, 2004; Wang et al., 2011a; Wolf and...
oral citalopram solution is 2 h after ingestion (Wang et al., 2011a). Recently, ASL has become the most common method to measure perfusion and is applied more often in phMRI studies (Chen et al., 2011a; Luo et al., 2009; Schouw et al., 2011). A recent paper by Wang et al. (2011a) states that ASL as an imaging marker for drug actions has many advantages above available BOLD-based techniques. Recently, a study by Chen et al. (2011a) suggested that ASL-based phMRI with a single oral SSRI challenge is a feasible method to detect 5-HT activity in the brain. To evaluate the test–retest abilities of this technique, so that it may be used reliably on multiple time points, the reproducibility of ASL-based phMRI following a single oral challenge needs to be verified. If reliable, this technique could ultimately become a fast and non-invasive tool to assess serotonergic activity on different time points, which is also suited for vulnerable patient populations, such as psychiatric patients and children.

Methods

Participants

Twelve healthy female volunteers were recruited through advertisement. Inclusion criteria were as follows: right-handedness, age between 20 and 30 years, and currently taking oral contraceptives. Exclusion criteria included medical conditions that might interfere with the interpretation of results, contraindications to MR imaging, current or past known brain disease, psychiatric or neurological disorders, presently taking prescribed medication that could affect 5-HT function, habitual drug use, and excessive consumption of alcohol (>21 U/week), caffeine (>8 cups of coffee/day), or nicotine (>10 cigarettes/day). The absence of psychiatric or neurological disorders and drug abuse was checked with the well-validated Mini-International Neuropsychiatric Interview (M.I.N.I.) Plus (van Vliet and de Beurs, 2007). All subjects were instructed to withhold the consumption of substances that could directly influence CBF (e.g. caffeine, alcohol, nicotine; (Mathew and Wilson, 1991)) before each scan session and between scans. Also, subjects were asked to refrain from any form of drug use during the entire study. The study was approved by a local research ethics committee and written informed consent was obtained from each volunteer.

Experimental design

All participants underwent two MRI scan sessions on three different days, so six scan sessions in total. The first scan session (baseline) was immediately followed by intake of the oral challenge on a fixed time point, consisting of either a 16 mg citalopram solution (equivalent to one 20 mg tablet; Lundbeck, Amsterdam) dissolved in lemonade or a placebo solution (lemonade only). A second scan session (challenge), which was identical to the first, was made exactly 2 h later (Tmax of 5-HT challenge was 2 h after ingestion1). On two out of three assessment days, a high resolution 3D T1-weighted structural scan was made for registration and segmentation purposes. The imaging parameters for the perfusion-weighted scan were as follows: a pulsed ASL sequence (PASL), based on the PULSAR sequence developed by Golay et al. (2005) was used. The labeling plane was positioned parallel to the imaging volume with a labeling gap between the imaging volume and the labeling volume of 25 mm. ASL parameters were TR, 3000 ms; TE, 14 ms; FOV, 240×240 mm²; matrix size, 80×79; 17 slices with slice thickness of 7 mm and no gap; gradient echo single shot EPI; SENSE 2.0; post-labeling delay of 1.2 to 2 s, depending on the slice number; number of dynamics, 50. One dynamic consisted of a control and a labeled image. For the structural scan, parameter values were as follows: high-resolution 3D T1-weighted anatomical image; TR, 9.8 ms; TE, 4.6 ms; 120 contiguous transversal slices.

Data preprocessing

Matlab (The MathWorks Inc., Natick, USA; http://www.mathworks.com), and the Statistical Parametric Mapping (SPM8) toolbox (http://www.fil.ion.ucl.ac.uk/spm) were used for offline data processing. Fifty pairs of labeled and control images were subtracted and averaged to yield whole brain perfusion weighted images. The mean equilibrium magnetization (M0) of arterial blood per subject was calculated, from which the absolute CBF was computed (Chalela et al., 2000). To allow for a voxel-based analysis, the structural images were non-rigidly normalized to a population-based average using diffeomorphic anatomical registration through exponentiated lie algebra (DARTEL) (Ashburner, 2007). The non-rigid transformations were then applied to the ASL volumes that were previously registered to gray matter masks, such that all dynamics of all subjects resided in one common frame of reference. Next, all images were spatially normalized into the Talairach and Tournoux stereotactic space (Talairach and Tournoux, 1988) using Montreal Neurological Institute (MINI) templates to facilitate intersubject averaging. Finally, a Gaussian smoothing with FWHM = 6 mm was applied to all volumes. Predefined regions-of-interest (ROIs) were determined based on the Harvard Oxford structural atlas (provided within FSL 4.1: FMRI-Software-Library, Functional Magnetic Resonance Imaging of the Brain Centre, Dept. of Clinical Neurology, University of Oxford, Oxford, UK, http://www.fmrib.ox.ac.uk). Based upon previous literature on the effects of citalopram on brain function (Anderson et al., 2008; McKie et al., 2005; Schwarz et al., 2007), the following ROIs were included: orbitofrontal cortex (OFC), inferior frontal gyrus (IFG), medial frontal gyrus (MFG), superior frontal gyrus (SFG), anterior cingulated gyrus (ACG), amygdala, thalamus, hippocampus, putamen and caudate and gray matter (GM) as a control region. Within these ROIs, the median CBF-value of all voxels within a specific ROI was calculated and used for statistical analysis.

Reproducibility of PASL signal

In order to establish the stability of the CBF measures, the between-session repeatability was assessed for the placebo session and its baseline (reproducibility within 1 day) and for all three baseline sessions (reproducibility over several weeks) by calculating three commonly used reproducibility indices per ROI: the within-subject coefficient of variation (wsCV), coefficient of repeatability (CR) and intraclass correlation coefficient (ICC) (Chen et al., 2011b; Gevers et al., 2011; Wang

1 http://www.medicines.org.uk/EMC/medicine/6153/SPC.
between CBV-values of the challenged scan and its accompanying baseline scan in that ROI. Based upon the estimated general effect size and expected within-day signal variation (standard deviation of paired difference within the placebo session), the sample size needed to detect significant differences with 90% power was calculated per ROI.

Results

Subjects

Citalopram administration was well tolerated in all subjects (mean age 22.9±3.0 years) and no clinical side effects were reported during or after scanning. The average time between the baseline scan session and the second scan session, i.e. time between challenge intake and the subsequent scan session, was 2 h and 3 min±4.2 min. The average time between the assessment days was 15±2.5 days. Due to technical problems and no-show, data from either one or two scan sessions of three subjects were missing. These included two placebo sessions and one of each citalopram sessions.

Between-session repeatability of PASL

On both the voxel-based level and ROI-based level, there were no significant differences between average CBV-values of the placebo session and its preceding baseline session. The placebo sessions were therefore used to assess the typical between-session variation of CBV measures within 1 day. Between-session reproducibility within 1 day ranged from good (wsCV>10%) to excellent (wsCV<10%) in all ROIs and could also be considered good to very good on the voxel-based level (a vast majority of the GM voxels showed ICC-values above 0.70). The wsCV, CR and ICC values for all ROIs can be found in Table 1. The voxel-based ICC map is depicted in Fig. 2A. The between-session variation within 2–4 weeks was determined using the three separate baseline scan sessions. On the ROI-based level, reproducibility of the PASL signal within 2–4 weeks was considerably lower than the within-day reproducibility. It was nevertheless still considered acceptable to good (10%<wsCV<20%) for most regions (see Table 1). On the voxel-based level, the ICC map also showed lower consistency between the scan sessions several weeks

Effect of citalopram challenge

Within each ROI, the effect of the challenge compared to baseline values was assessed using a Wilcoxon signed-rank test in SPSS for Windows (v.16.0, SPSS Inc., Chicago, IL) on the average CBV-values per ROI. On the voxel-based level, a higher level statistical analysis was carried out in SPM8 using the preprocessed mean CBF images. A paired t-test analysis was used to compare the challenge scans and their accompanying baseline scans on a group level. The resulting statistical parametric maps were thresholded at a T-value>4.14, corresponding to a p-value>0.001, uncorrected. Only clusters of a size bigger than 30 contiguous significant voxels were reported.

Reproducibility of citalopram effect

The effect of the challenge in the first citalopram session (EffCit1) was compared to the effect of the challenge in the second citalopram session (EffCit2) using again the CR on an ROI-based level and an ICC map on a voxel-based level as measures of reliability. In each ROI where a significant effect of citalopram was detected, the effect size of this challenge effect was determined as the standardized mean difference

Table 1

<table>
<thead>
<tr>
<th>Within 1 day</th>
<th>CR (ml/100 g/min)</th>
<th>wsCV (%)</th>
<th>ICC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM</td>
<td>6.40</td>
<td>5.37</td>
<td>0.923 (0.72−0.98)</td>
</tr>
<tr>
<td>SFG</td>
<td>6.57</td>
<td>5.38</td>
<td>0.950 (0.81−0.99)</td>
</tr>
<tr>
<td>IFG</td>
<td>8.90</td>
<td>6.53</td>
<td>0.951 (0.82−0.99)</td>
</tr>
<tr>
<td>MFG</td>
<td>11.99</td>
<td>6.75</td>
<td>0.937 (0.77−0.98)</td>
</tr>
<tr>
<td>OFC</td>
<td>17.18</td>
<td>13.31</td>
<td>0.813 (0.46−0.96)</td>
</tr>
<tr>
<td>ACG</td>
<td>22.59</td>
<td>9.98</td>
<td>0.589 (0.59−0.97)</td>
</tr>
<tr>
<td>Amygdala</td>
<td>15.08</td>
<td>13.79</td>
<td>0.825 (0.44−0.95)</td>
</tr>
<tr>
<td>Thalamus</td>
<td>10.02</td>
<td>9.64</td>
<td>0.948 (0.81−0.99)</td>
</tr>
<tr>
<td>HC</td>
<td>24.00</td>
<td>15.45</td>
<td>(−0.30−0.80)</td>
</tr>
<tr>
<td>Caudate</td>
<td>15.60</td>
<td>8.95</td>
<td>0.870 (0.56−0.97)</td>
</tr>
<tr>
<td>Putamen</td>
<td>11.27</td>
<td>6.64</td>
<td>0.930 (0.75−0.98)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Within 2–4 weeks</th>
<th>CR (ml/100 g/min)</th>
<th>wsCV (%)</th>
<th>ICC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM</td>
<td>16.26</td>
<td>17.34</td>
<td>0.499 (0.08−0.84)</td>
</tr>
<tr>
<td>SFG</td>
<td>20.24</td>
<td>18.96</td>
<td>0.589 (0.19−0.87)</td>
</tr>
<tr>
<td>IFG</td>
<td>21.10</td>
<td>20.01</td>
<td>0.677 (0.31−0.91)</td>
</tr>
<tr>
<td>MFG</td>
<td>24.88</td>
<td>22.00</td>
<td>0.668 (0.29−0.90)</td>
</tr>
<tr>
<td>OFC</td>
<td>21.22</td>
<td>16.82</td>
<td>0.575 (0.17−0.87)</td>
</tr>
<tr>
<td>ACG</td>
<td>40.85</td>
<td>19.75</td>
<td>0.658 (0.12−0.85)</td>
</tr>
<tr>
<td>Amygdala</td>
<td>16.40</td>
<td>17.97</td>
<td>0.535 (0.34−0.91)</td>
</tr>
<tr>
<td>Thalamus</td>
<td>26.90</td>
<td>18.14</td>
<td>0.647 (0.27−0.90)</td>
</tr>
<tr>
<td>HC</td>
<td>21.41</td>
<td>22.47</td>
<td>0.647 (0.07−0.83)</td>
</tr>
<tr>
<td>Caudate</td>
<td>21.20</td>
<td>20.01</td>
<td>0.581 (0.18−0.87)</td>
</tr>
<tr>
<td>Putamen</td>
<td>22.36</td>
<td>17.81</td>
<td>0.537 (0.13−0.85)</td>
</tr>
</tbody>
</table>

wsCV = within subject coefficient of variation; CR = coefficient of repeatability; GM = gray matter; HC = hippocampus; ICC = intraclass correlation coefficient; IFG = inferior frontal gyrus; MFG = medial frontal gyrus; OFC = orbital frontal gyrus; SFG = superior frontal gyrus. wsCV values below 10% (bold) are considered good to excellent, values between 10% and 20% (italic) are considered acceptable to good. ICC values above 0.80 (bold) are considered excellent, values between 0.80 and 0.70 (italic) are considered good.


Fig. 1. Within-subject cross-over study design. Oral intake of the challenge medication took place right after the baseline scan. A placebo scan was randomly assigned to day 1 or day 2 and subjects were blind to the type of challenge received.

et al., 2011b). The wsCV was calculated as the SD of the difference between two CBF measurements given as a percentage of the mean of the two measurements (Bland and Altman, 1996). The CR is defined as the largest likely size of difference between 2 repeated measurements (within a 95% confidence limit). It is given by the following equation: CR = 1.96 × SD_SMB, with SD_SMB being the SD of the difference between repeated measurements. The ICC measures the contribution of between-subject variances to the total variance (Shrout and Fleiss, 1979). ICC values can range from zero (no agreement) to one (perfect agreement) (McGraw and Wong, 1996). Two-way mixed model ICCs (ICC(3,1)) were calculated using SPSS for Windows (v.16.0, SPSS Inc., Chicago, IL). An ICC value above 0.70 was considered as adequate repeatability of the used method. On a voxel-based level, ICC maps were generated to assess the reliability per voxel using the ICC Toolbox for SPM5 (http://brainmap.co.uk). Statistical methods behind this tool are described in detail in Caceres et al. (2009).
apart than between the scan sessions within 1 day (see Fig. 2B). However, most voxels still showed ICC values above 0.70.

**Effect of citalopram**

An ROI-based analysis showed a significant effect of the citalopram challenge in the SFG (effect size $+4.94 \text{ ml/100 g/min, } p\text{-value } 0.041$ (uncorrected)) and a marginally significant effect in the thalamus (effect size $+4.72 \text{ ml/100 g/min, } p\text{-value } 0.050$ (uncorrected)), both only during session 2 (see Table 2). In the same session, there was a trend towards significant effect of the challenge in the MFG and IFG (effect size $+5.12$ and $+3.48 \text{ ml/100 g/min respectively, } p\text{-value } 0.075$ in both cases). During session 1, in none of the examined regions a significant effect of the citalopram challenge was found.

Voxel-based analysis showed small but significantly activated clusters in both citalopram sessions compared to the baseline sessions (see Fig. 3). The effect of the citalopram challenge during session 1 (CitEff1) was only seen in the frontal cerebral white matter (WM). During session 2 (CitEff2), significantly activated clusters were found in the right thalamus, MFG, SFG and the right cerebellum, which greatly overlaps with the ROI-based findings. The exact location and size of these clusters are given in Table 3. As can be clearly seen in Fig. 3, the activated clusters did not overlap between the two different citalopram sessions.

**Test–retest reliability of citalopram effect**

A low reproducibility of the citalopram effect was confirmed by the relatively high CR values (Table 2 under ‘reliability of CitEff’) in comparison to the within-day CR values of the PASL signal itself (Table 1), which are on average $\pm 10 \text{ ml/100 g/min lower.}$ On a voxel-based level, the ICC map of CitEff1 and CitEff2 showed a very low number of voxels (<5%) above the threshold of 0.70 (Fig. 2C). In addition, voxels in which the effect of citalopram could be considered repeatable (ICC>0.70) were not located in the same regions in which actual effects of the citalopram were seen during one of the two sessions.

The effect sizes of the (near) significant effects of citalopram obtained with the ROI analysis lay around $+5 \text{ ml/100 g/min (see under ‘effect size’ in Table 2).}$ The sample sizes needed to detect such an effect size with 90% power, considering the previous determined within-day variability in PASL signal per ROI, are given in Table 2 (under ‘sample size’). This power analysis indicated that sample sizes should consist of 5 to 15 subjects in the GM, thalamus, SFG and IFG. Large sample sizes of above 50 subjects are needed in the ACG and hippocampus.

**Discussion**

In this placebo-controlled, within-subject cross-over study we assessed the feasibility and reliability of ASL-based phMRI as a non-invasive tool for assessing brain 5-HT function.

As anticipated, the within-day between-session repeatability of CBF values using a PASL sequence was good to excellent for most ROIs. Within-subject variation (wsCV) was in the order of 5–15% and the largest expected difference between two repeated measurements in one subject (coefficient of repeatability; CR) was on average $\pm 13 \text{ ml/100 g/min.}$ These findings agree with earlier reported values of reproducibility of PASL on 3 T (Chen et al., 2011b). The variation in CBF-values between scans acquired 2 to 4 weeks apart was considerably higher. This suggests that in order to keep variation as low as possible and thus retain power to detect CBF differences caused by

![Fig. 2. Voxel-based ICC maps overlayed on population averaged perfusion map (transversal view; n=11). A) ICC(3,1) values between-session within 1 day (baseline vs. placebo; n=10), B) between-session ICC within 2–4 weeks (all three baseline scans; n=9), C) reliability of citalopram effect (CitEff1 vs. CitEff2) (n=10).](image-url)
the time of peak plasma concentration of citalopram (Dalgaard and
Larsen, 1999; Mendoza et al., 2005). These effects of individual vari-
ation will nevertheless for a great part be overcome with larger sam-
ple sizes. Our results confirm recent findings by Chen et al. (2011a), who
reported problems in detecting drug-induced CBF changes with PASL
using the same sample size. Due to its better signal-to-noise ratio, the re-
cently developed pCASL sequence measures CBF even more reliably than
the PASL sequence does (Chen et al., 2011b; Gevers et al., 2011) and
might therefore be better suited for detecting signal changes of this
small size. However, the reliability of this specific ASL technique in
phMRI studies still needs to be confirmed.

In summary, our study illustrates the requirements and limita-
tions of (PASL-based) phMRI in assessing brain 5-HT function on
multiple time points. Despite the good reproducibility of the PASL sig-
nal, the between-session variation remained too high to reliably
detect the modest effects on CBF induced by a single oral dose of
citalopram with this speci-
fi
1.956 3.48 5.12 0.075 0.075 0.091 0.091

c technique. We therefore conclude that
the PASL sequence does (Chen et al., 2011b; Gevers et al., 2011) and
might therefore be better suited for detecting signal changes of this
small size. However, the reliability of this specific ASL technique in
phMRI studies still needs to be confirmed.

In summary, our study illustrates the requirements and limita-
tions of (PASL-based) phMRI in assessing brain 5-HT function on
multiple time points. Despite the good reproducibility of the PASL sig-
nal, the between-session variation remained too high to reliably
detect the modest effects on CBF induced by a single oral dose of
citalopram with this specific technique. We therefore conclude that
the test–retest reliability of PASL phMRI with an oral citalopram chal-

enge is low, limiting the technique's sensitivity to time-depen-
dent changes.

The relatively short post-labeling delay of PASL results in large sig-
nal variations near blood vessels (Wolf and Detre, 2007). To detect
citalopram effects (with an effect size of 5 ml/100 g/min) in these areas,
such as the ACG and hippocampus, large sample sizes (above 50 sub-
jects) are needed. Other areas that are sufficiently large or dis-
tant from vascular components, among which are the GM, thalamus,
SFG and IFG, show an effect in much smaller groups of 5

The challenging drug, it is essential to compare the challenged scan
with a baseline scan made on the same day.

Despite the excellent within-day reproducibility of the PASL signal,
the effect of a single oral dose of citalopram on CBF values appeared to
be small and could not be replicated between sessions. Although during
session two an increase in CBF of about 5 ml/100 g/min was seen in the
5-HT rich thalamus, SFG and MFG, no effects were reported during
session one in GM regions where the CBF can be reliably measured
(Wolf and Detre, 2007). To detect the modest effects on CBF induced by

citalopram with this speci-
fi
1.956 3.48 5.12 0.075 0.075 0.091 0.091

c technique. We therefore conclude that
the PASL sequence does (Chen et al., 2011b; Gevers et al., 2011) and
might therefore be better suited for detecting signal changes of this
small size. However, the reliability of this specific ASL technique in
phMRI studies still needs to be confirmed.

In summary, our study illustrates the requirements and limita-
tions of (PASL-based) phMRI in assessing brain 5-HT function on
multiple time points. Despite the good reproducibility of the PASL sig-
nal, the between-session variation remained too high to reliably
detect the modest effects on CBF induced by a single oral dose of
citalopram with this specific technique. We therefore conclude that
the test–retest reliability of PASL phMRI with an oral citalopram chal-

enge is low, limiting the technique's sensitivity to time-depen-
dent changes.

The relatively short post-labeling delay of PASL results in large sig-
nal variations near blood vessels (Wolf and Detre, 2007). To detect
citalopram effects (with an effect size of 5 ml/100 g/min) in these areas,
such as the ACG and hippocampus, large sample sizes (above 50 sub-
jects) are needed. Other areas that are sufficiently large or dis-
tant from vascular components, among which are the GM, thalamus,
SFG and IFG, show an effect in much smaller groups of 5–15 subjects.
Our data thus indicate that it is advisable to confine data analysis to
these brain regions. Indeed, effects of the challenge were seen in the
thalamus and SFG during one of the assessments, as well as a trend
in IFG. Note that the GM was used as a reference region since no
systemic effect of the low dose citalopram on the general CBF value
was expected (Smith et al., 2002). Physiological variations form a
other limitation to measuring the modest effect of oral citalopram
on CBF values. Perfusion values may already vary up to 20% over
repeated imaging sessions within subjects (Gevers et al., 2011).
Furthermore, individual physiological differences in responsiveness
of the 5-HT system to citalopram influence the change in CBF
(Hu et al., 2007; Kronenberg et al., 2007; Lotrich et al., 2001) and
the time of peak plasma concentration of citalopram (Dalgaard and

Table 2
Citalopram effect per ROI and per session (Wilcoxon signed-rank test) and its reliability over sessions.

<table>
<thead>
<tr>
<th>Effect of citalopram</th>
<th>GM</th>
<th>SFG</th>
<th>IFG</th>
<th>MFG</th>
<th>OFC</th>
<th>ACG</th>
<th>Amygdala</th>
<th>Thalamus</th>
<th>HC</th>
<th>Caudate</th>
<th>Putamen</th>
</tr>
</thead>
<tbody>
<tr>
<td>CitEff1 (n=11)</td>
<td>Z</td>
<td>0.03</td>
<td>0.53</td>
<td>0.18</td>
<td>0.53</td>
<td>0.80</td>
<td>0.35</td>
<td>0.889</td>
<td>0.711</td>
<td>0.089</td>
<td>0.711</td>
</tr>
<tr>
<td>CitEff2 (n=11)</td>
<td>p</td>
<td>0.06</td>
<td>0.54</td>
<td>0.54</td>
<td>0.85</td>
<td>0.59</td>
<td>0.42</td>
<td>0.722</td>
<td>0.374</td>
<td>0.477</td>
<td>0.929</td>
</tr>
<tr>
<td>Effect sizea</td>
<td></td>
<td>4.55</td>
<td>3.48</td>
<td>4.52</td>
<td>5.12</td>
<td>4.72</td>
<td>4.39</td>
<td>3.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample sizeb</td>
<td></td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>16</td>
<td>32</td>
<td>56</td>
<td>25</td>
<td>11</td>
<td>63</td>
<td>27</td>
</tr>
</tbody>
</table>

CR (ml/100 g/min) 17.61 18.90 14.18 23.52 26.46 35.37 24.33 23.89 30.72 20.75 17.91

ACG = anterior cingulate gyrus; CitEff1 = effect of citalopram; CitEff2 = effect of citalopram on CBF values. Perfusion values may already vary up to 20% over
retest reliability of PASL phMRI with an oral citalopram challenge is low, limiting the technique's sensitivity to time-depen-
dent changes.

The relatively short post-labeling delay of PASL results in large sig-

Fig. 3. Statistical parametric T-maps (p<0.001, uncorrected; cluster size>30 voxels) of citalopram effect. A) Significant effects of citalopram during session 1 (CitEff1) and B) session 2 (CitEff2), on top of the population averaged anatomical map (coronal view; n=11).
Voxels thresholded at $T > 4.14$. Clusters with a size of $> 30$ voxels were reported. Size is the number of voxels above threshold in the cluster; $T$-value is the $T$-value of the most activated voxel in the corresponding cluster; $X$, $Y$, and $Z$ represent the MNI coordinates (mm).