Psychiatry and Clinical Neurosciences 2015; **69**: 22–33 doi:10.1111/pcn.12209

Review Article

Near-infrared spectroscopy for examination of prefrontal activation during cognitive tasks in patients with major depressive disorder: A meta-analysis of observational studies

Huijun Zhang, BM,[1] Wentian Dong, MD,[1] Weimin Dang, MSc,[1] Wenxiang Quan, PhD,[1] Ju Tian, BS,[1] Ru Chen, BM,[2] Siyan Zhan, PhD[2] and Xin Yu, MD[1]*

[1]*Peking University Sixth Hospital, Peking University Institute of Mental Health, Key Laboratory of Mental Health, Ministry of Health (Peking University) and* [2]*Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center, Beijing, China*

Aims: Near-infrared spectroscopy has the potential for aiding the diagnosis of major depressive disorder. The purpose of this study was to systematically review the evidence from observational studies regarding the use of near-infrared spectroscopy in patients with major depressive disorder and to identify the characteristic pattern of prefrontal lobe activity in major depressive disorder.

Methods: MEDLINE, PubMed, Cochrane Library and Web of Science databases were searched in December 2013. All case–control studies were included. The quality of evidence was examined using the Newcastle–Ottawa Quality Assessment Scale. The primary outcome measures were the mean oxygenated and deoxygenated hemoglobin alterations of the cerebral cortex during cognitive activation periods. The standard mean difference for the overall pooled effects across the included studies was estimated using random or fixed effect models. The primary outcome measures were included in the meta-analysis.

Results: Fourteen studies met the inclusion criteria. Six studies ($n = 692$ participants) were included in the analysis of the mean oxygenated hemoglobin alterations; the pooled mean standardized difference was −0.74 (95% confidence interval, −0.97 to −0.52), indicating that patients with major depressive disorder were associated with attenuated increase in oxygenated hemoglobin during cognitive activation in the prefrontal regions compared to healthy controls. Five studies ($n = 668$ participants) were included in the analysis of mean deoxygenated-hemoglobin changes; the pooled standardized mean difference was 0.18 (95% confidence interval, −0.20 to 0.56).

Conclusions: Using near-infrared spectroscopy measurements, we observed that compared to healthy subjects, patients with major depressive disorder had significantly lower prefrontal activation during cognitive tasks.

Key words: major depressive disorder, meta-analysis, near-infrared spectroscopy, neuroimaging, prefrontal cortex.

MAJOR DEPRESSIVE DISORDER (MDD) is a mental disorder that is characterized by a pervasive and persistent low mood, diminished interest or pleasure in daily activities, as well as a variety of somatic, psychological and cognitive symptoms. According to epidemiological surveys, MDD is a common global disorder, associated with a high risk of lifelong chronic recurrence, increased suicide rates and numerous adverse outcomes, including reduced role functioning and financial burden.[1] The World Health Organization ranked MDD as the fourth most disabling disorder and projected that by 2020, MDD

*Correspondence: Xin Yu, MD, Peking University Sixth Hospital, Peking University Institute of Mental Health, Key Laboratory of Mental Health, Ministry of Health (Peking University), Huayuanbeilu 51, Haidian District, Beijing 100191, China.
Email: yuxin@bjmu.edu.cn
Received 28 February 2014; revised 6 May 2014; accepted 29 May 2014.

would be the second leading cause of disability worldwide.[2] Despite its prevalent and burdensome condition, MDD is significantly underdiagnosed and undertreated, particularly in the primary care environment.[3] Optimizing diagnostic approaches is essential for providing earlier diagnosis of MDD and improving patient management. The current diagnosis of MDD is based largely on a subjective examination of the symptoms and signs reported by the patients. There are no definite laboratory tests for MDD. A large body of research accumulated over recent decades has identified and characterized novel candidate biomarkers for MDD, which encompass neuroimaging, genetic/epigenetic, and neurochemical approaches.[4,5] Although several positive biological markers have been proposed, to date, none of these findings has been consistently replicated to warrant clinical use.

Near-infrared spectroscopy (NIRS), which is a newly developed optical neuroimaging technology, has been recently proposed as a potential diagnostic method for detecting MDD.[6] NIRS can noninvasively and continuously monitor alterations in tissue chromophore concentrations, such as oxygenated [oxyHb] and deoxygenated [deoxyHb] hemoglobin in micro blood vessels, by using near-infrared light. Jöbsis was the first investigator to propose the principle of NIRS.[7] NIRS is based on the fact that light in the NIR spectrum (650–950 nm) penetrates cerebral tissues and is preferentially absorbed by oxyHb and deoxyHb. Because oxyHb and deoxyHb have different optical properties (light absorbance and extinction coefficients), hemoglobin concentration changes can be calculated from the intensity changes of detected NIR light in tissue at two or more wavelengths using the modified Lambert–Beer law.[8,9] Assuming that hematocrit is constant, the increases of oxyHb and decreases of deoxyHb detected in NIRS are correlated with regional cerebral blood flow (rCBF) and may indicate brain activation as shown in simultaneous measurements with other methodologies.[10,11]

Within the last 2 decades, NIRS has been increasingly used to investigate abnormal cerebral hemodynamic changes during functional activation studies in psychiatric disorders,[12] focusing particularly on affective disorders. Findings obtained by several original NIRS studies identified the characteristic pattern of MDD patients concerning the time course of prefrontal oxygenation during verbal fluency task relative to controls,[13] based on which NIRS was approved as advanced medical technology to facilitate depression diagnosis by the Japanese Ministry of Health, Labor, and Welfare in 2009. This novel technology has not been clinically validated. Critics claim that this technology should not be offered without supporting evidence.[14] Because most of the supporting studies have small sample sizes, a meta-analysis, which is more powerful in estimating effect size, may present more generalizable evidence. Within this context, we used NIRS measurements and conducted a systemic review of NIRS studies of depression to explore whether a consistent, reproducible and typical prefrontal activation pattern exists in patients with MDD compared to normal people.

METHODS

Search strategy

Searches of both English-language and Chinese-language literature concerning the use of NIRS for investigating brain activation in MDD were undertaken from 1980 to December 2013 with MEDLINE, PubMed, Cochrane Library, Web of Science, ScienceDirect Onsite, John Wiley, PsychiatryOnline, Proquest, Ovid, and CBMDISC databases. The searches were performed using the terms 'Near-infrared Spectroscopy', 'Infrared Spectrophotometry', 'Optical Imaging' along with additional terms, such as 'Depression', 'Depressive Disorder', 'Major Depressive Disorder', 'Mood Disorders', 'Affective Disorders', and all possible combinations. Relevant journals, conference proceedings and bibliographies of retrieved papers were manually searched. Google Scholar and trial registries (ClinicalTrial.gov, Chinese Clinical Trial Register) were also searched for published or unpublished studies.

Inclusion criteria

Studies were included if they employed NIRS to assess cerebral hemodynamic changes in response to cognitive paradigms in patients with MDD compared to healthy controls. The patients with MDD were diagnosed according to DSM-IV or ICD-10 criteria; the patients included in the study had no history of manic, mixed or hypomanic episodes. Additionally, a primary end-point, including the mean changes in oxyHb and deoxyHb concentrations (all of the data were expressed as the mean and standard deviation), was presented.

Data collection

Two independent reviewers scanned all of the titles and abstracts of the studies retrieved. Full-text articles were retrieved for additional assessment when appropriate. Regarding studies that met the inclusion criteria, two reviewers independently extracted their study characteristics and results. Any discrepancies were resolved by discussion. The data review form comprised the following categories: (i) authors and year of publication; (ii) study design; (iii) sample characteristics (sample size, demographic and clinical features); (iv) NIRS device; (v) cognitive paradigm; and (vi) study outcomes.

The original data were obtained as much as possible from the articles; data that could not be obtained were calculated based on the provided data in the articles. In the absence of data in a table or text, GetaData Graph Digitizer 2.22 (http://getdata-graph-digitizer.com/) was employed to convert data in the figure into values. We attempted to contact the authors if the key data required for analysis were not provided.

Quality assessment

The methodological quality of non-randomized studies was assessed, adapting the Newcastle–Ottawa Quality Assessment Scale (NOS).[15] The scale does not have an overall scoring or threshold for a 'good' or 'poor' study. The NOS for case–control studies has eight items within three domains: selection (representativeness), comparability (attributed to design or analysis), and outcomes (assessment and follow-up). A study can receive one star for meeting each criterion, except comparability, for which a study can receive a maximum of two stars. In this review, for one star under comparability, the study controlled for medication status. For two stars under comparability, the study also controlled for important variables, such as age, sex, and handedness. Pairs of reviewers worked independently, and discrepancies were resolved by discussion.

Statistical analysis

The mean changes and standard differences in oxyHb and deoxyHb concentrations during the performance of cognitive tasks were compared between patients with MDD and healthy controls. Due to the technological progress of NIRS instrumentation (from single-location continuous wave NIRS measurements of the early years progressing to multi-channel functional cortical near-infrared topography), and the various NIRS prototypes that were developed by different companies,[16] NIRS instrumentation prototypes used in the studies have undergone changes over time. For the statistical analysis of mean oxyHb and deoxyHb changes, one difficulty is that the number of channels and the size of coverage area of NIRS devices were markedly varied across studies. We presented a new method in which the pooled effects of all channels for each group were used in the analysis. Another difficulty is that for several multi-channel NIRS machines, measurement of areas of cerebral cortex covered a broader range of frontal-temporal regions. We extracted the signal data of channels of the prefrontal area if the correspondence of NIRS channels and measurement points to the cerebral cortex were confirmed in the studies. Otherwise, we used roughly the pooled effects of all channels for analysis, limited by the characteristic of poor spatial resolution of NIRS.

Analyses were conducted using RevMan 5.1 (The Nordic Cochrane Centre, The Cochrane Collaboration, 2011, Copenhagen, Denmark). The standardized mean difference (SMD) was used as a summary statistic in the meta-analysis to standardize the results of the studies to a uniform scale before they could be combined, when the studies all assess the same outcome but measure the outcome in various ways. In the present meta-analysis, SMD was calculated to measure the difference between the mean value of the two groups in a study because of the different NIRS instrumentation and cognitive tasks used in the retrieved studies. The data were pooled using the generic inverse variance method. A random-effects model was used if the studies were heterogeneous on the basis of the Q statistic for heterogeneity and if the reasons for the heterogeneity could not be identified. Otherwise, a fixed-effect model was used to estimate the pooled effect. Furthermore, we used the I^2 to quantify heterogeneity. Values of 25% correspond to low heterogeneity, values of 50% correspond to moderate heterogeneity and values of 75% correspond to high heterogeneity. A subgroup meta-analysis was performed using mood states and cognitive tasks to investigate the effect of mood state in the patients and the effect of the cognitive task paradigms, respectively. In the first subgroup analysis, studies were further divided into either remitting or non-remitting MDD subgroups. Remitting MDD was classified as

having a Hamilton Depression Rating Scale (HAMD) score less than 8 points. A funnel plot of effect sizes versus standard errors was used as an outcomes test to detect publishing or reporting bias through visual inspection if the number of included studies was larger than ten. If the funnel plot showed asymmetry, Egger's test was performed.

RESULTS

Selection of studies

A total of 2260 citations were identified within our database search. After reviewing the titles and abstracts and removing duplicate publications, 19 full-text articles[17–35] were reviewed. Of these 19 articles, 14 studies[22–35] met the inclusion criteria (Fig. 1).

Study characteristics

The primary characteristics of the included studies are summarized in Table 1. All 14 studies finally selected were case–control studies. Eleven studies were conducted in Japan, two studies were conducted in Germany, and one study was conducted in China. The pooled sample across the 14 studies included 1093 individuals (307 MDD patients; 786 healthy controls). All of the patients met the DSM-IV or ICD-10 diagnostic criteria for MDD. Of 1093 participants, 46% were male. The age range of patients with MDD and the control groups was 18–79 years and 20–77 years, respectively; five studies[24,26,28,29,32] enrolled patients with late-onset depression or geriatric depression. Baseline depression severity differed across studies, with the mean score of HAMD ranging from 0.9 to 29.5. With the exception of two studies[28,33] that included medication-free patients, patients received antidepressants of different types and dosages with or without antipsychotics and benzodiazepines. The majority of the studies adopted the verbal fluency task (VFT) as a short cognitive stimulation, whereas three studies[30,31,33] assessed the working memory task (WM). The NIRS instruments that were used varied markedly across studies. One study used

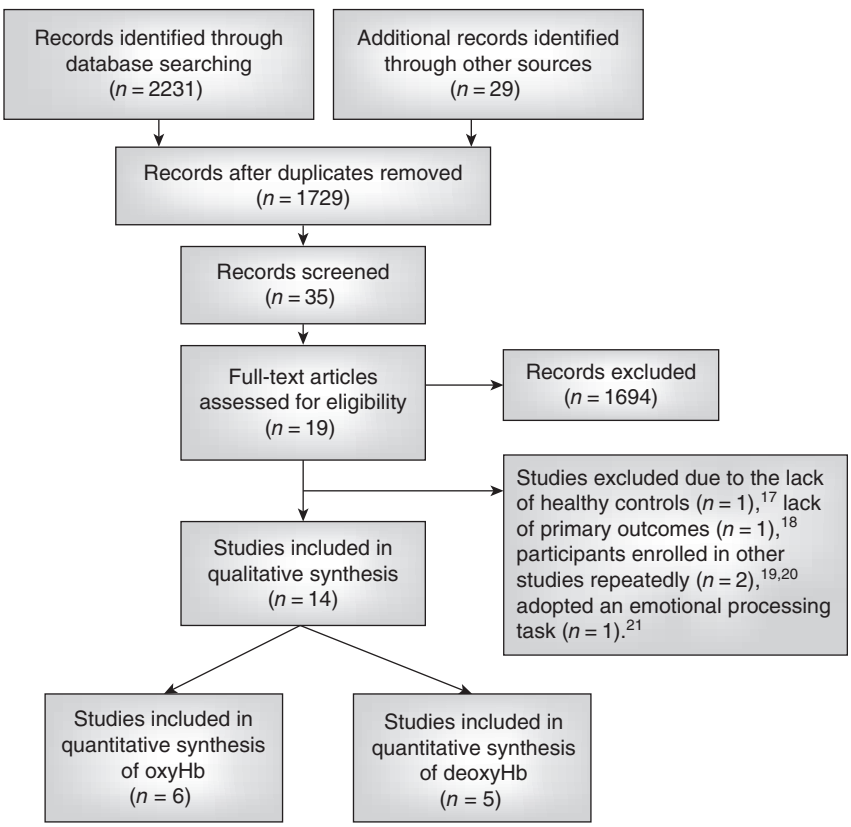

Figure 1. Flow diagram illustrating literature search. deoxyHb, deoxygenated hemoglobin; oxyHb, oxygenated hemoglobin.

Table 1. Important characteristics of included studies

Source	Setting	Sample size (male/female)	Age (Mean ± standard deviation)	Diagnostic criteria (Diagnostic instrument)	Psychopathology measures	Medication (number of patients)	Exclusion criteria	NIRS device	Cognitive paradigm	Results
Matsuo 2002[22]	Japan	MDD: 14 (4/10) HC: 21 (3/18)	MDD: 56.1 ± 17.3 HC: 50.3 ± 12.6	DSM-IV (MINI)	MDD: 5.3 ± 4.0 (HAMD)	MDD: antidepressants (14), antipsychotics (8), lithium (2) HC: free	Substance dependence disorder, endocrine disease, head trauma, neurological diseases, or any chronic general medical disease requiring lifelong treatment	Single-channel NIRS (HEO-200)	VFT	The increase of oxyHb during a verbal fluency task was significantly lower in the MDD group than in the controls.
Herrmann 2004[23]	Germany	MDD: 9 (5/4) HC: 9 (5/4)	MDD: 37.3 ± 13.8 HC: 35.1 ± 5.5	ICD-10 (Unspecified)	MDD: 16.56 ± 8.14 (BDI) HC: 1.56 ± 0.44 (BDI)	MDD: tricyclic antidepressants (8), lithium (1) HC: free	MDD: acute or prior comorbidities with other psychiatric illnesses. HC: acute or former neurological or psychiatric disorder	Two-channel NIRS (NIRO-300)	VFT	Patients had significantly lower activation bilaterally during the cognitive task.
Matsuo 2005[24]	Japan	MDD: 10 (5/5) HC: 10 (6/4)	MDD: 62.2 ± 4.8 HC: 58.7 ± 5.8	DSM-IV (MINI)	MDD: 4.0 ± 2.2 (HAMD), 28.1 ± 2.3 (MMSE) HC: 28.3 ± 2.2 (MMSE)	MDD: SSRI (4), drug-free for more than 2 weeks (6) HC: free	Substance dependence disorder, endocrine disease, head trauma, neurological diseases, or any chronic general medical disease requiring lifelong treatment	24-channel NIRS (Hitachi ETG-100)	VFT	Activation of the prefrontal cortex during the cognitive task was significantly less in patients compared to controls.
Kameyama 2006[25]	Japan	MDD: 11 (9/2) HC: 17 (13/4)	MDD: 44.8 ± 13.1 HC: 42.8 ± 4.5	DSM-IV (Unspecified)	MDD: 10.4 ± 9.5 (HAMD)	MDD: Antidepressants (11), mood stabilizers (2), antipsychotics (1) HC: free	Neurological disorders, substance abuse, head injuries or major physical illnesses	Two 24-channel NIRS (Hitachi ETG-100)	VFT	The oxyHb increases in the major depression group were significantly lower than those in the healthy control group in eight frontal channels and four left and four right lower anterior temporal channels primarily during the early task period.
Zhao 2007[26]	China	MDD: 12 (7/5) HC: 12 (5/7)	MDD: 68.42 ± 6.33 HC: 70.33 ± 7.05	DSM-IV (Unspecified)	MDD: 10.25 ± 1.60 (HAMD), 28.50 ± 1.51 (MMSE) HC: 28.3 3 ± 1.37 (MMSE)	MDD: SSRI (12) HC: free	Neurological disorders, cerebral organic lesions, substance abuse, head injuries or major physical illnesses	28-channel CW NIRS	VFT	Activation of the left prefrontal cortex during the VFT was significantly less in patients compared to controls.
Ohta 2008[27]	Japan	MDD: 17 (5/12) HC: 24 (12/12)	MDD: 42.8 ± 18.2 HC: 36.2 ± 16.5	DSM-IV (MINI)	MDD: 16.6 ± 4.5 (HAMD)	MDD: Antidepressants (17) HC: free	Unspecified	52-channel NIRS (ETG-4000)	VFT	Patients with MDD showed attenuated increases in oxyHb during the VFT in the bilateral frontal regions compared to healthy controls.

Table 1. (*Continued*)

Source	Setting	Sample size (male/female)	Age (Mean ± standard deviation)	Diagnostic criteria (Diagnostic instrument)	Psychopathology measures	Medication (number of patients)	Exclusion criteria	NIRS device	Cognitive paradigm	Results
Pu 2008[28]	Japan	MDD: 24 (6/18) HC: 30 (14/16)	MDD: 72.3 ± 5.5 HC: 72.0 ± 4.7	DSM-IV (MINI)	MDD: 18.4 ± 3.8 (HAMD), 20.1 ± 9.5 (BDI), 27.3 ± 1.7 (MMSE) HC: 5.2 ± 3.7 (BDI), 27.9 ± 2.1 (MMSE)	MDD: free HC: free	Central nervous system disorders, head trauma, stroke, substance abuse, psychotic symptoms, electroconvulsive therapy	52-channel NIRS (ETG-4000)	VFT	LOD patients had less activation in a broad area covering both prefrontal and superior temporal cortices than healthy controls.
Yamagata 2008[29]	Japan	MDD: 23 (5/18) HC: 13 (8/5)	EOD: 68.4 ± 5.6 LOD: 70.2 ± 1.9 HC: 70.3 ± 4.4	DSM-IV (Unspecified)	EOD: 15.4 ± 6.3 (HAMD) LOD: 19.0 ± 6.8 (HAMD)	MDD: paroxetine (6), milnacipran (5), tricyclic antidepressants (12) HC: free	Dementia, seizures, alcohol or other substance abuse, neurological disorders, traumatic brain injury	52-channel NIRS (ETG-4000)	VFT	Increases in oxyHb were mildly attenuated in EOD and severely attenuated in LOD in most channels.
Pu 2011[30]	Japan	MDD: 24 (12/12) HC: 26 (8/18)	MDD: 47.9 ± 13.9 HC: 42.4 ± 9.3	DSM-IV (MINI)	MDD: 22.1 ± 12.7 (BDI), 20.3 ± 9.2 (HAMD) HC: 8.0 ± 8.0 (BDI)	MDD: SSRI (13), SNRI (8), TCA (3) HC: free	Central nervous system disorders, head trauma, stroke, substance abuse, electroconvulsive therapy	52-channel NIRS (ETG-4000)	2-back task	MDD patients showed a smaller increase in lateral prefrontal and superior temporal cortex activation during the 2-back task than healthy controls.
Schecklmann 2011[31]	Germany	UNI: 16 (9/7) HC: 15 (7/8)	UNI: 43.4 ± 9.8 HC: 40.9 ± 8.0	ICD-10 (SCID-I)	UNI: 18.2 ± 9.2 (BDI) HC: 4.1 ± 3.0 (BDI)	UNI: Antidepressants (15), antipsychotics (9), benzodiazepine (9), mood stabilizers (1) HC: free	Neurological illness or severe somatic illness	52-channel NIRS (ETG-4000)	Object and spatial visual WM	Patient group showed diminished brain activity in all working memory conditions.
Pu 2012[32]	Japan	MDD: 26 (11/15) HC: 30 (12/18)	MDD: 47.9 ± 19.2 HC: 50.5 ± 19.7	DSM-IV (MINI)	MDD: 22.9 ± 10.4 (BDI), 17.6 ± 7.0 (HAMD) HC: 3.6 ± 3.1 (BDI)	MDD: SSRI (13), SNRI (9), TCA (4) HC: free 13	Central nervous system disorders, head trauma, stroke, substance abuse, electroconvulsive therapy	52-channel NIRS (ETG-4000)	VFT	Regional hemodynamic changes were significantly smaller in the MDD group than in the control group in prefrontal and temporal regions.
Pu 2012[33]	Japan	MDD: 36 (9/27) HC: 35 (11/24)	MDD: 71.8 ± 5.1 HC: 70.9 ± 4.3	DSM-IV (MINI)	MDD: 19.6 ± 3.7 (HAMD), 20.1 ± 8.8 (BDI), 27.2 ± 1.9 (MMSE) HC: 5.4 ± 3.9 (BDI), 27.8 ± 1.9 (MMSE)	MDD: free HC: free	Central nervous system disorders, head trauma, stroke, substance abuse, psychotic symptom, electroconvulsive therapy	52-channel NIRS (ETG-4000)	2-back task	LOD patients were associated with reduced increase in prefrontal and temporal regions activation compared to healthy controls.

Table 1. (*Continued*)

Source	Setting	Sample size (male/female)	Age (Mean ± standard deviation)	Diagnostic criteria (Diagnostic instrument)	Psychopathology measures	Medication (number of patients)	Exclusion criteria	NIRS device	Cognitive paradigm	Results
Noda 2012[34]	Japan	MDD: 30 (14/16) HC: 30 (14/16)	MDD: 36.7 ± 11.6 HC: 35.1 ± 9.4	ICD-10 (SCID-I)	MDD: 16.7 ± 4.8 (HAMD)	MDD: antidepressant (27), anxiolytics (20), hypnotics (16), mood stabilizers (7), antipsychotics (9) HC: free	Head trauma, neurological illness, history of electroconvulsive therapy, alcohol/substance abuse or addiction	52-channel NIRS (ETG-4000)	VFT	The oxyHb increase during the VFT was significantly smaller in patients than in controls.
Takizawa 2014[35]	Japan	MDD: 55 HC: 514	MDD: 43.8 ± 12.7 HC: 43.9 ± 15.7	DSM-IV (SCID)	MDD: 14.1 ± 6.7 (HAMD)	MDD: antidepressants HC: free	Neurological illness, traumatic brain injury with any known cognitive consequences and alcohol/substance abuse or addition	52-channel NIRS (ETG-4000)	VFT	Unspecific

BDI, Beck Depression Inventory; EOD, early onset major depression; HAMD, Hamilton Depression Scale; HC, healthy controls; LOD, late-onset major depression; MDD, major depressive disorder; MINI, Mini-International Neuropsychiatric Interview; MMSE, Mini-mental state examination; NIRS, near-infrared spectroscopy; oxyHb, oxygenated hemoglobin; SCID-I, Structured Clinical Interview for DSM-IV Axis I Disorders; SNRI, serotonin-norepinephrine reuptake inhibitors; SSRI, specific serotonin reuptake inhibitors; TCA, tricyclic antidepressants; UNI, unipolar depression; VFT, verbal fluency task; WM, working memory.

single-channel NIRS (HEO-200, Omron, Tokyo, Japan),[22] one study used two-channel NIRS (NIRO-300, Hamamatsu Photonics K.K., Hamamatsu City, Japan),[23] two studies used 24-channel NIRS (Hitachi ETG-100, Hitachi Medical Corporation, Tokyo, Japan),[24,25] one study used 28-channel CW5 NIRS (TechEn, Milford, MA, USA),[25] and nine studies used 52-channel NIRS (Hitachi ETG-4000).[27–35]

According to our review of the authors' conclusions, all of the studies reached the unanimous conclusion that patients with MDD had significantly lower activation of the prefrontal cortex during cognitive tasks than controls. However, for quantitative meta-analysis, the original data needed were incomplete or missing in most of the included studies. Six studies were ultimately included in the meta-analysis of oxyHb (the values needed for analysis were directly extracted from the tables in two studies,[23,26] or were converted from line charts using GetaData Graph Digitizer software,[22,24] or were provided by the authors after email contacts[31,35]). Five studies were included in the meta-analysis of deoxyHb (the values needed for analysis were directly extracted from the tables in one study,[23] or were converted from line charts using GetaData Graph Digitizer software,[22,24] or were provided by the authors after email contacts[31,35]).

For six studies included in the meta-analysis of oxyHb, one study used single-channel NIRS (with a pair of probes placed on the left prefrontal region),[22] whereas other studies used multi-channel systems (two-channel system located in the bilateral prefrontal cortex,[23] 24-channel system placed on the bilateral prefrontal areas,[24] 28-channel system covering the bilateral prefrontal areas,[25] and 52-channel system measuring the bilateral prefrontal and superior temporal cortical regions[31,35]). The pooled effects of all of the channels were calculated, with one exception: one multisite study divided the 52 channels into two regions (frontal and temporal) for separate analysis.[35] We thus selectively analyzed the signals of Region 1 (the frontopolar and dorsolateral prefrontal cortical regions) and abandoned the signals of the temporal regions.

Five of the above six studies were also included in the meta-analysis of deoxyHb.[22–24,31,35] The same analysis strategy was adopted.

Table 2. Risks of bias within studies

	Selection				Comparability	Exposure		
Study sources	Is the case definition adequate?	Representative of cases	Selection of controls	Definition of controls		Determination of exposure	Same method for determining cases and controls	Non-response rate
Matsuo 2002[22]	★	0	★	★	★	★	★	★
Herrmann 2004[23]	★	0	★	★	★	★	★	★
Matsuo 2005[24]	★	0	★	★	★	★	★	★
Kameyama 2006[25]	★	0	★	★	★	★	★	★
Zhao 2007[26]	★	0	★	★	★	★	★	★
Ohta 2008[27]	★	0	★	★	★	★	★	★
Pu 2008[28]	★	0	★	★	★★	★	★	★
Yamagata 2008[29]	★	0	★	★	★	★	★	★
Pu 2011[30]	★	0	★	★	★	★	★	★
Schecklmann 2011[31]	★	0	★	★	★	★	★	
Pu 2012[32]	★	0	★	★	★	★	★	★
Pu 2012[33]	★	0	★	★	★★	★	★	★
Noda 2012[34]	★	0	★	★	★	★	★	★
Takizawa 2014[35]	★	0	★	★	★	★	★	★

Risk of bias in the included studies

The NOS was used for assessing the quality of the included studies. The results for case–control studies are shown in Table 2. Two studies[28,33] are given eight stars, 11 studies are given seven stars and only one study[31] is given six stars. All 14 studies met the NOS criteria for case definition; however, none was clear about the representativeness of cases. Regarding the comparability of the study groups, only two studies[28,33] adjusted for medication or were matched according to medication. The other studies were adjusted for or were matched on other important potential confounders, such as age, sex, intelligence or graduation and handedness. To determine exposure, NIRS was applied to examine brain activation for both cases and controls in all of the studies. Only one study[31] did not have comparable non-response rates; five patients were excluded from group analyses because of difficulties in understanding the task.

Outcome measurement

OxyHb analysis

A total of six studies[22–24,26,31,35] were included in the analysis of the mean alterations of oxyHb comprising 111 patients with MDD and 581 healthy controls (Fig. 2). There was no evidence of heterogeneity ($I^2 = 0\%$, $P = 0.48$). The pooled SMD of oxyHb change was −0.74 (95% confidence interval [CI], −0.97 to −0.52, $P < 0.00001$), which indicated that patients with MDD were associated with reduced increase in oxyHb during the cognitive activation period in the prefrontal regions compared to the healthy controls. However, the results from the present subgroup analysis of mood states failed to suggest a significant difference. The pooled SMD for the non-remitting MDD subgroup was −0.74 (95%CI, −0.99 to −0.49). The pooled SMD for the remitting MDD subgroup was −0.75 (95%CI, −1.31 to −0.19). Also, the results of the subgroup analysis in VFT and WM failed to suggest a significant difference. The pooled SMD for VFT was −0.76 (95%CI, −0.99 to −0.52). The pooled SMD for WM was −0.57 (95%CI, −1.37 to 0.23).

DeoxyHb analysis

Five studies[22–24,31,35] were included in terms of the differences of mean changes in deoxyHb during the cognitive activation period between patients with MDD ($n = 99$) and healthy controls ($n = 569$) (Fig. 3). The pooled SMD was 0.18 (95%CI, −0.20 to 0.56, $P = 0.36$), which indicated no significant differences between the groups. Additionally, there was evidence of low heterogeneity between studies ($I^2 = 41\%$, $P = 0.15$). There were no significant subgroup differences between the remitting and non-

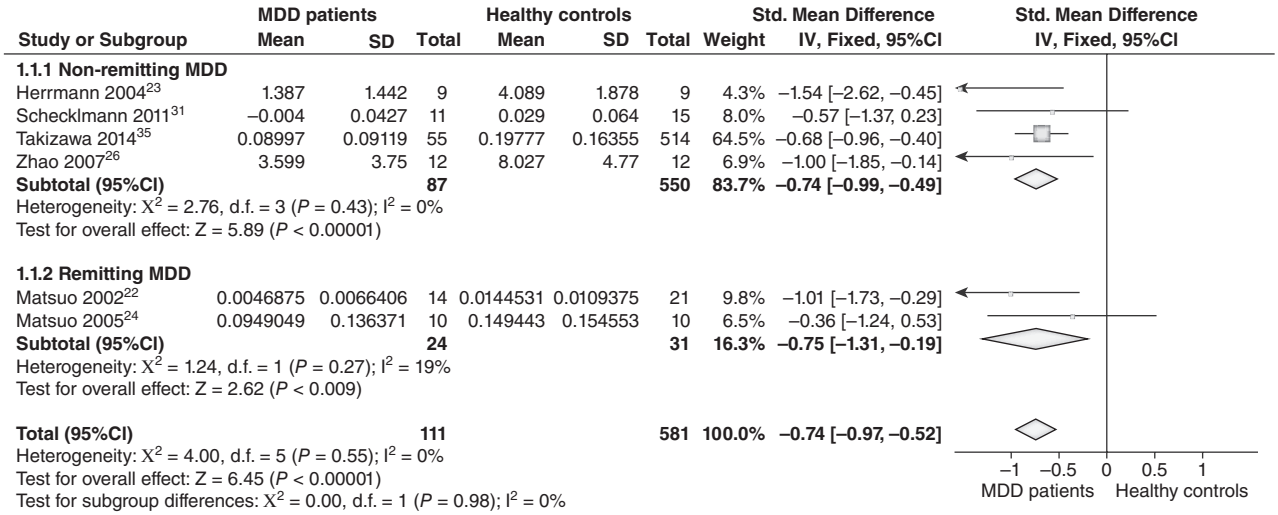

Figure 2. Comparison of mean alteration of oxygenated hemoglobin during cognitive activation period measured using near-infrared spectroscopy in patients with major depressive disorder (MDD) versus healthy controls. Std. Mean Difference, standardized mean difference; CI, confidence interval.

remitting MDD subgroups. The pooled SMD for the non-remitting MDD subgroup was 0.23 (95%CI, −0.16 to 0.62). The pooled SMD for the remitting MDD subgroup was 0.08 (95%CI, −1.03 to 1.19).

DISCUSSION

Consequently, the meta-analyses of the studies included in the present article provided evidence for a distinct characteristic pattern of blood oxygenation changes of the prefrontal cortex in response to different cognitive tasks between patients with MDD and healthy controls using NIRS. The present review showed oxyHb activation during VFT and WF to be significantly smaller in patients with MDD than in healthy controls. We found a large effect size of oxyHb changes (SMD, −0.74; 95%CI, −0.97 to −0.52). However, in contrast to significantly smaller

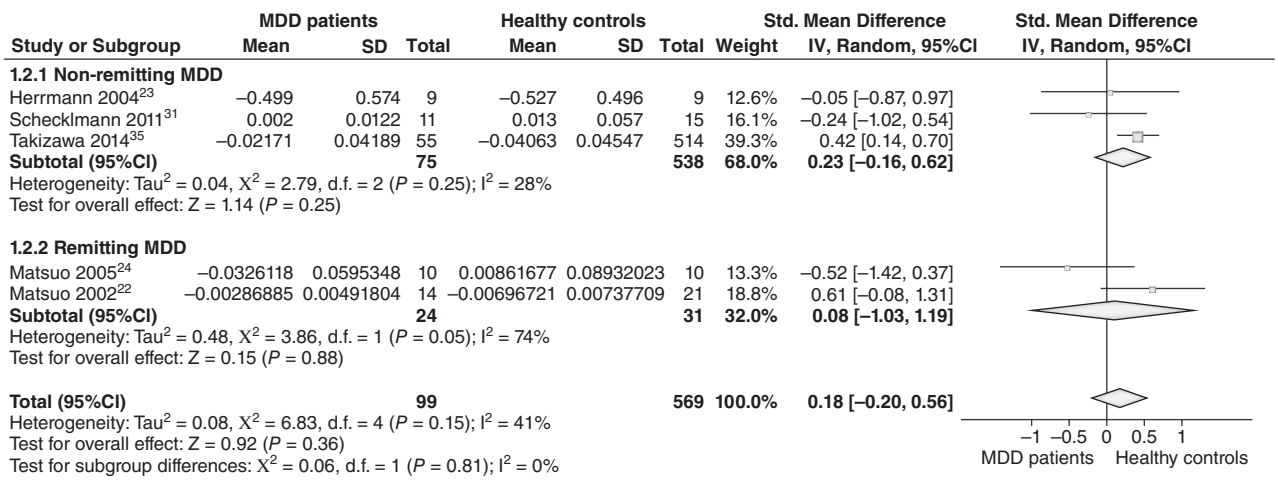

Figure 3. Comparison of mean alteration of deoxygenated hemoglobin during cognitive activation period measured using near-infrared spectroscopy in patients with major depressive disorder (MDD) versus healthy controls. Std. Mean Difference, standardized mean difference; CI, confidence interval.

oxyHb changes, the meta-analysis of deoxyHb failed to show a significant difference between the groups (SMD, 0.18; 95%CI, −0.20 to 0.56). This finding is consistent with findings in previous literature reports,[36,37] which indicates that although both the oxyHb increase and the deoxyHb decrease detected by NIRS are assumed to reflect cortical activation, oxyHb was positively correlated with CBF and was demonstrated to be the more sensitive marker of CBF changes than deoxyHb. From the above meta-analysis of oxyHb changes, patients with MDD had significantly lower prefrontal activation during cognitive tasks using NIRS than healthy subjects. Many neuropsychological studies demonstrated impaired executive function in patients with MDD.[38,39] Current neuroimaging research using positron emission tomography (PET), single-photon emission computed tomography (SPECT), and functional magnetic resonance imaging (fMRI) suggested dysfunctional prefrontal–subcortical circuitry underlying the cognitive deficits in MDD.[40] However, the findings were not consistent or were insufficiently reproducible. Several studies demonstrated 'hypofrontality' (i.e. decreased activity in glucose metabolism and regional cerebral blood flow in the frontal lobes) under a resting state or during task-related activation in patients with MDD,[41–43] whereas the contradictory findings of 'hyperfrontality' and normal prefrontal activation have also been reported.[44–46]

To be diagnostically useful, ideal biomarkers should be reproducible, sensitive, specific and easily accessible. The findings of our systematic review confirmed a consistent characteristic pattern of attenuated prefrontal activation in patients with MDD when performing cognitive tasks that engage the prefrontal cortex by using the NIRS technique. NIRS is inexpensive, non-invasive and easy to use. The role of NIRS prompted its usage as an objective tool for clinically diagnosing MDD.

A further subgroup analysis of mood states demonstrated no significant differences between the remitting and non-remitting MDD subgroups, suggesting that the attenuated frontal activation detected by NIRS during cognitive tasks may be a trait-dependent property of MDD. However, as noted in a previous study, the oxyHb changes appeared to be a state-dependent marker of MDD.[34] Follow-up studies are needed to determine whether the frontal hemodynamic responses measured by NIRS are a trait- or state-dependent marker of MDD, because the existing evidence is inadequate and many confounding factors cannot be ruled out.

Our review has several limitations. First, we included only observational studies that had a small sample size and that were at great risk, compared to RCT, for selective and confounding bias, thereby reducing the strength of evidence. Second, we could not include all 14 eligible studies in each quantitative analysis because of insufficiency of original data. We narratively reviewed the primary results. Nearly without exception in all of the studies, patients with MDD were noted to generate reduced prefrontal activation during cognitive tasks. Third, we did not perform a meta-analysis on the total hemoglobin concentration, due to a lack of data, although some researchers assert that total hemoglobin concentrations are more closely related to regional cerebral blood volume than to oxyHb or deoxyHb.[47] Fourth, because most patients were taking antidepressants during the NIRS examination, we could not rule out that antidepressants interfered with NIRS results. Fifth, we performed a meta-analysis on the pooled effects of all channels of hemoglobin concentrations. Although such a procedure enables a comprehensive assessment of global prefrontal activation, it also could miss the larger effect size in more restricted prefrontal regions. Sixth, the publication bias might account for several of the observed effects. The small number of studies included in the oxyHb or deoxyHb analysis did not allow us to assess the possibility for publication bias.

The shortcomings of the NIRS methodology are as follows: NIRS enables measurement of hemoglobin concentration changes only as relative values, not as absolute values; NIRS has relatively low spatial resolution and low cerebral penetration depth; and the contributions from extracerebral tissue, such as the skin and skull, may contaminate the NIRS signal. Due to a lack of standard quantification, the acquired hemoglobin data from a variety of NIRS instruments in this meta-analysis are provided as relative values and are measured in different units (i.e. mmol · mm, mmol/L, or arbitrary unit). To ensure that the data were combined properly, we used the SMD rather than the mean difference as a summary statistic; this allowed us to standardize the hemoglobin concentration of the studies to a uniform scale, as described in the statistical analysis section.

In conclusion, there is preliminary evidence supporting a characteristic pattern of decreased prefrontal cortical activation in patients with MDD during cog-

nitive tasks measured using the NIRS technique. This finding may imply the probability of widespread clinical use of NIRS as an objective diagnostic tool for depression. However, there were several influential methodological problems and publication bias. Additional research in large patient groups is warranted.

ACKNOWLEDGMENTS

The authors have no conflicts of interest to declare. We are grateful to Professor Martin Schecklmann, Dr Masato Fukuda, Dr Kiyoto Kasai, Dr Ryu Takizawa, Dr Masaru Mimura, Dr Shingo Kawasaki, Dr Shenghong Pu, Dr Takamasa Noda, Dr Kazuyuki Nakagome, Dr Shin-ichi Niwa, Dr Yuji Okazaki, and Dr Yoshihiro Satomura for their assistance in providing the original data required for the meta-analyses.

REFERENCES

1. Kessler RC, Bromet EJ. The epidemiology of depression across cultures. *Annu. Rev. Public Health* 2013; **34**: 119–138.
2. Murray CJL, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet* 1997; **349**: 1498–1504.
3. Sheehan DV. Depression: Underdiagnosed, undertreated, underappreciated. *Manag. Care* 2004; **13** (6 Suppl. Depression): 6–8.
4. Schneider B, Prvulovic D. Novel biomarkers in major depression. *Curr. Opin. Psychiatry* 2013; **26**: 47–53.
5. Schneider B, Prvulovic D, Oertel-Knöchel V *et al.* Biomarkers for major depression and its delineation from neurodegenerative disorders. *Prog. Neurobiol.* 2011; **95**: 703–717.
6. Cyranoski D. Neuroscience: Thought experiment. *Nature* 2011; **469**: 148–149.
7. Jöbsis FF. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science* 1977; **198**: 1264–1267.
8. Strangman G, Boas DA, Sutton JP. Non-invasive neuroimaging using near-infrared light. *Biol. Psychiatry* 2002; **52**: 679–693.
9. Taber KH, Hillman EMC, Hurley RA. Optical imaging: A new window to the adult brain. *J. Neuropsychiatry Clin. Neurosci.* 2010; **22**: iv, 357–360.
10. Schytz HW, Guo S, Jensen LT *et al.* A new technology for detecting cerebral blood flow: A comparative study of ultrasound tagged NIRS and 133Xe-SPECT. *Neurocrit. Care* 2012; **17**: 139–145.
11. Hock C, Villringer K, Müller-Spahn F *et al.* Decrease in parietal cerebral hemoglobin oxygenation during performance of a verbal fluency task in patients with Alzheimer's disease monitored by means of near-infrared spectroscopy (NIRS) – correlation with simultaneous rCBF-PET measurements. *Brain Res.* 1997; **755**: 293–303.
12. Ehlis AC, Schneider S, Dresler T, Fallgatter AJ. Application of functional near-infrared spectroscopy in psychiatry. *Neuroimage* 2014; **85**: 478–488.
13. Fukuda M, Mikuni M. Clinical application of near-infrared spectroscopy (NIRS) in psychiatry: The advanced medical technology for differential diagnosis of depressive state. *Seishin Shinkeigaku Zasshi* 2012; **114**: 801–806 (in Japanese).
14. Wells GA, Shea B, O'Connell D *et al.* First do no harm. Simple tools to diagnose mental illness should not be offered without sound supporting evidence. *Nature* 2011; **469**: 132.
15. Wells GA, Shea B, O'Connell D *et al.* The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. 2011. [Cited 25 November 2013.] Available from URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
16. Ferrari M, Quaresima V. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage* 2012; **63**: 921–935.
17. Onishi Y, Kikuchi S, Watanabe E, Kato S. Alterations in prefrontal cortical activity in the course of treatment for late-life depression as assessed on near-infrared spectroscopy. *Psychiatry Clin. Neurosci.* 2008; **62**: 177–184.
18. Okada F, Takahashi N, Tokumitsu Y. Dominance of the 'nondominant' hemisphere in depression. *J. Affect. Disord.* 1996; **37**: 13–21.
19. Matsuo K, Kato T, Fukuda M, Kato N. Alteration of hemoglobin oxygenation in the frontal region in elderly depressed patients as measured by near-infrared spectroscopy. *J. Neuropsychiatry Clin. Neurosci.* 2000; **12**: 465–471.
20. Suto T, Fukuda M, Ito M, Uehara T, Mikuni M. Multichannel near-infrared spectroscopy in depression and schizophrenia: Cognitive brain activation study. *Biol. Psychiatry* 2004; **55**: 501–511.
21. Matsubara T, Matsuo K, Nakashima M *et al.* Prefrontal activation in response to emotional words in patients with bipolar disorder and major depressive disorder. *Neuroimage* 2014; **85**: 489–497.
22. Matsuo K, Kato N, Kato T. Decreased cerebral haemodynamic response to cognitive and physiological tasks in mood disorders as shown by near-infrared spectroscopy. *Psychol. Med.* 2002; **32**: 1029–1037.
23. Herrmann MJ, Ehlis AC, Fallgatter AJ. Bilaterally reduced frontal activation during a verbal fluency task in depressed patients as measured by near-infrared spectroscopy. *J. Neuropsychiatry Clin. Neurosci.* 2004; **16**: 170–175.
24. Matsuo K, Onodera Y, Hamamoto T, Muraki K, Kato N, Kato T. Hypofrontality and microvascular dysregulation in remitted late-onset depression assessed by functional near-infrared spectroscopy. *Neuroimage* 2005; **26**: 234–242.

25. Kameyama M, Fukuda M, Yamagishi Y *et al*. Frontal lobe function in bipolar disorder: A multichannel near-infrared spectroscopy study. *Neuroimage* 2006; **29**: 172–184.
26. Zhao LS, Ma X, Bao F *et al*. Prefrontal lobe function in remitted geriatric depression: A study by using 28-channel near-infrared spectroscopy. *J. Clin. Rehabil. Tissue Eng. Res.* 2007; **11**: 8974–8978 (in Chinese).
27. Ohta H, Yamagata B, Tomioka H *et al*. Hypofrontality in panic disorder and major depressive disorder assessed by multi-channel near-infrared spectroscopy. *Depress. Anxiety* 2008; **25**: 1053–1059.
28. Pu S, Matsumura H, Yamada T *et al*. Reduced frontopolar activation during verbal fluency task associated with poor social functioning in late-onset major depression: Multichannel near-infrared spectroscopy study. *Psychiatry Clin. Neurosci.* 2008; **62**: 728–737.
29. Yamagata B, Tomioka H, Takahashi T, Isomura AJ, Kobayashi H, Mimura M. Differentiating early and late-onset depression with multichannel near-infrared spectroscopy. *Psychogeriatrics* 2008; **8**: 79–87.
30. Pu S, Yamada T, Yokoyama K *et al*. A multi-channel near-infrared spectroscopy study of prefrontal cortex activation during working memory task in major depressive disorder. *Neurosci. Res.* 2011; **70**: 91–97.
31. Schecklmann M, Dresler T, Beck S *et al*. Reduced prefrontal oxygenation during object and spatial visual working memory in unipolar and bipolar depression. *Psychiatry Res.* 2011; **194**: 378–384.
32. Pu S, Nakagome K, Yamada T *et al*. The relationship between the prefrontal activation during a verbal fluency task and stress-coping style in major depressive disorder: A near-infrared spectroscopy study. *J. Psychiatr. Res.* 2012; **46**: 1427–1434.
33. Pu S, Yamada T, Yokoyama K *et al*. Reduced prefrontal cortex activation during the working memory task associated with poor social functioning in late-onset depression: Multi-channel near-infrared spectroscopy study. *Psychiatry Res.* 2012; **203**: 222–228.
34. Noda T, Yoshida S, Matsuda T *et al*. Frontal and right temporal activations correlate negatively with depression severity during verbal fluency task: A multi-channel near-infrared spectroscopy study. *J. Psychiatr. Res.* 2012; **46**: 905–912.
35. Takizawa R, Fukuda M, Kawasaki S *et al*. Neuroimaging-aided differential diagnosis of the depressive state. *Neuroimage* 2014; **1**: 498–507.
36. Strangman G, Culver JP, Thompson JH, Boas DA. A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. *Neuroimage* 2002; **17**: 719–731.
37. Hoshi Y, Kobayashi N, Tamura M. Interpretation of near-infrared spectroscopy signals: A study with a newly developed perfused rat brain model. *J. Appl. Physiol.* 2001; **90**: 1657–1662.
38. Fossati P, Ergis AM, Allilaire JF. Executive functioning in unipolar depression: A review. *Encephale* 2002; **28**: 97 (abstract) (in French).
39. Boeker H, Schulze J, Richter A, Nikisch G, Schuepbach D, Grimm S. Sustained cognitive impairments after clinical recovery of severe depression. *J. Nerv. Ment. Dis.* 2012; **200**: 773–776.
40. Murrough JW, Iacoviello B, Neumeister A, Charney DS, Iosifescu DV. Cognitive dysfunction in depression: Neurocircuitry and new therapeutic strategies. *Neurobiol. Learn. Mem.* 2011; **96**: 553–563.
41. Okada G, Okamoto Y, Morinobu S, Yamawaki S, Yokota N. Attenuated left prefrontal activation during a verbal fluency task in patients with depression. *Neuropsychobiology* 2003; **47**: 21–26.
42. Audenaert K, Goethals I, Van Laere K *et al*. SPECT neuropsychological activation procedure with the Verbal Fluency Test in attempted suicide patients. *Nucl. Med. Commun.* 2002; **23**: 907–916.
43. Siegle GJ, Thompson W, Carter CS, Steinhauer SR, Thase ME. Increased amygdala and decreased dorsolateral prefrontal BOLD responses in unipolar depression: Related and independent features. *Biol. Psychiatry* 2007; **61**: 198–209.
44. Diener C, Kuehner C, Brusniak W, Ubl B, Wessa M, Flor H. A meta-analysis of neurofunctional imaging studies of emotion and cognition in major depression. *Neuroimage* 2012; **61**: 677–685.
45. Fitzgerald PB, Srithiran A, Benitez J *et al*. An fMRI study of prefrontal brain activation during multiple tasks in patients with major depressive disorder. *Hum. Brain Mapp.* 2008; **29**: 490–501.
46. Videbech P, Ravnkilde B, Kristensen S *et al*. The Danish PET/depression project: Poor verbal fluency performance despite normal prefrontal activation in patients with major depression. *Psychiatry Res.* 2003; **123**: 49–63.
47. Huppert TJ, Hoge RD, Diamond SG *et al*. A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *Neuroimage* 2006; **29**: 368–382.