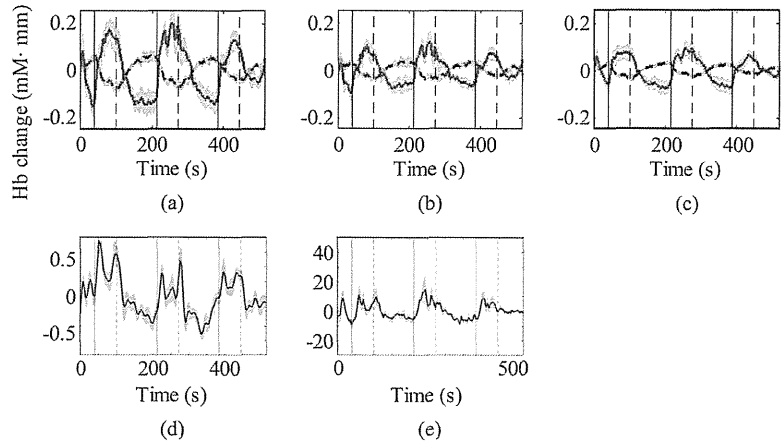

Fig. 3 Grand average of continuous data of fNIRS (original, deep, and shallow) and gray matter blood-oxygenation-level dependent (GM-BOLD) signals for activation channel, and laser-Doppler-flowmetry (LDF) signal changes (channel 2) obtained during verbal fluency task (VFT). Translucent patches indicate the standard error at each time point. Vertical solid and dashed lines indicate task onset and end timings, respectively. (a) Original signal; (b) deep signal; (c) shallow signal [oxy-Hb (solid line), deoxy-Hb (dashed line)]; (d) GM-BOLD signal change (%); and (e) LDF signal change (arb. unit).

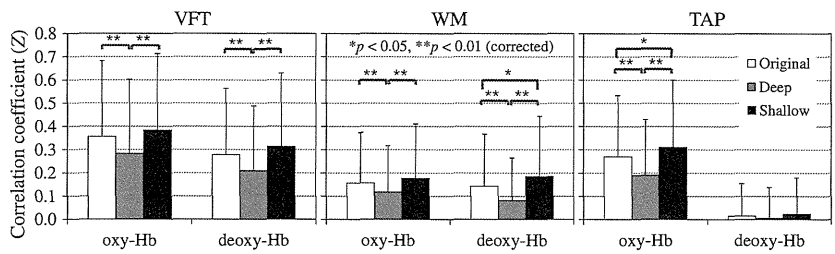

Fig. 4 Correlation coefficients (Fisher's Z) between oxy- and deoxy-Hb signals (original, deep, and shallow) and LDF signal (including both channels 1 and 2) during a verbal fluency task (VFT), a working memory task (WM), and a finger tapping task (TAP). Error bars indicate the standard deviations. Single (*) and double (**) asterisks denote the statistical significance at $p < 0.05$ and 0.01 (corrected for multiple comparisons), respectively.

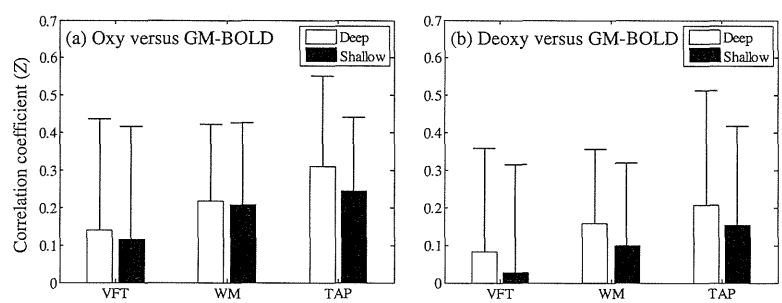

Fig. 5 Correlation coefficients (Fisher's Z) for (a) fNIRS oxy-Hb (deep and shallow) and spatially weighted GM-BOLD signals in target areas, and for (b) fNIRS deoxy-Hb (deep and shallow) and spatially weighted GM-BOLD signals in target areas during a VFT, a WM, and a TAP. Error bars indicate the standard deviations. The sign of deoxy-Hb signal is inverted. For any individual task, no significant differences were found between the correlations of deep and shallow fNIRS and GM-BOLD signals.

Table 4 Means and standard deviations of deep-layer pooled contribution ratio (%) at activation channels.

	oxy-Hb	deoxy-Hb
VFT	50.0 ± 17.1	55.1 ± 17.0
WM	56.2 ± 10.7	64.3 ± 9.7
TAP	60.9 ± 11.6	74.8 ± 5.3

lower than those reported in a previous study,[38] over half of the contribution of Hb signal to the fNIRS signals is originated from the deep layer, especially for oxy-Hb.

4 Discussion

4.1 *Correlations between Functional Near-Infrared Spectroscopy and Laser-Doppler-Flowmetry Signals*

The correlations between the fNIRS and LDF signals (shown in Fig. 4) are very similar to those obtained in a previous study[38] that showed the correlation coefficients for deep fNIRS and LDF signals are significantly lower than those for shallow fNIRS and LDF signals. The correlation coefficient between deoxy-Hb and LDF signals under the TAP condition was extremely low. That was possibly because the deep-layer pooled contribution ratio of deoxy-Hb under the TAP condition was relatively high (Table 4) in the present study. The low correlation between deoxy-Hb and skin blood flow can be caused by the low contribution of the shallow signal. It should be noted that we did not temporally integrate the LDF signal, but the integrated LDF signal can be more correlated with fNIRS signal when a proper integration time is selected.[29]

From the aspect of correlation between fNIRS and LDF signals, it was shown that the fNIRS signals were reasonably divided into signals with either higher or lower correlations with the LDF signal. It should be noted that during the finger-tapping task, the LDF signal had a higher correlation with the shallow fNIRS signal than that with the deep fNIRS signal, even if the target channels were located mainly in somatosensory or motor areas (BAs 1, 2, 3, 4, and 40) far from the LDF probes (attached on the forehead or temple). The result suggests that the LDF signal correlates with the shallow fNIRS signal in the broad area during the finger-tapping task.

4.2 *Correlation between Deep Functional Near-Infrared Spectroscopy Signals and Gray Matter Blood Oxygenation Level–Dependent Signals*

The mean of the correlation coefficients of the deep signal was significantly higher than that of the shallow signal. This is partly because deep (brain) and shallow (scalp) tissue layers are anatomically governed by different blood vessel systems (internal or external carotid artery). Different correlation coefficients for the fNIRS and the GM-BOLD signals would, therefore, be expected, i.e., the deep fNIRS signal should have stronger correlation with the GM-BOLD signal than that between the shallow fNIRS and the GM-BOLD signals.

From this point of view, the results of the correlation between deep fNIRS and GM-BOLD signals (Fig. 5) showed that the MD-ICA method successfully separates fNIRS signals into deep and shallow signals that have higher and lower correlations, respectively, with spatially weighted GM-BOLD signals.

Deep and shallow signals can be similar as a result of MD-ICA method. The high correlation between deep and shallow signals was also obtained in previous studies.[38,54] This can happen because the same independent components are commonly used for reconstructing deep and shallow signals, and the systemic signals did not be removed in order to quantify the contributions of deep and shallow signals. If the contributions of components that included both deep and shallow signals are almost the same, then the correlations of the shallow and deep signals with the GM-BOLD would be almost equivalent. In the present case, however, different correlations were obtained. This means that deep and shallow signals were different enough from each other to evaluate the separation performance. Although mean deep and shallow signals seem very similar, as shown in Fig. 3 for example, individual deep and shallow signals are different and have different correlations with LDF or GM-BOLD signals.

4.3 *Deep-Layer Contribution Ratio Obtained by Multidistance Independent Component Analysis*

It should be noted that the MD-ICA method quantifies the contribution ratios of both deep and shallow layers, but the ratios include the effect of systemic interference because the MD-ICA method discriminates fNIRS signals on the basis of signal depth only. Even if the deep-layer pooled contribution ratio is high, for example, it is possible that the systemic contribution in the deep layer is dominant.

It has been reported that there is interindividual variability in the correlation between the fNIRS signal and the scalp blood flow or mean blood pressure[66] and that the systemic changes that also affect extracranial signals can lead to false positives in fNIRS signals.[33] It should be noted that the effect of posture (sitting or supine) on the contribution of deep-layer tissue to fNIRS signal and its dependency on kind of task have not been investigated. Such an effect might cause the difference between the contribution ratios obtained in the current study and in a previous study.

4.4 *Limitations*

In regard to the proposed deep-shallow separation method (MD-ICA method), the structural parameter Xi_{gr} was fixed for all participants and for all measurement channels. In this study, it was confirmed that the fixed parameter is effective, even in the case where the structural differences depending on individuals and positions within individuals are not considered and neither MRI structural data nor x-ray CT data are available. It should be noted that the deep- or shallow-tissue condition may be changed by changing the posture. The deep/shallow contribution ratio calculated in this study (i.e., supine posture) is not necessarily the same as that calculated for a sitting posture.

The measurement area was limited to only prefrontal, somatosensory, and motor cortices on the left side of the head. Other areas should be covered by the proposed method, so occipital and temporal areas should be further investigated. All participants in this study were male; it would, therefore, be more helpful to validate the proposed method by using female participants.

5 Conclusion

Though very few studies have validated a multidistance scalp-effect-removal method with concurrent fNIRS-fMRI measurement, this study shows that the previously proposed deep/shallow separation method (MD-ICA method) successfully separates fNIRS signals into "spatially" deep and shallow signals by comparing these signals with spatially weighed GM-BOLD and LDF signals. The result shows that the accuracy and reliability of the fNIRS signal will be greatly improved with the MD-ICA method. The correlation coefficients for shallow fNIRS and LDF signals were larger than those for deep fNIRS and LDF signals. This result is consistent with the results obtained in a previous study.[38] This method needs only small numbers of probes [at least two middle-distance (>10.5 mm) channels], so it will easily contribute to broad area (e.g., whole head) brain-imaging studies using cost-effective equipment.

Acknowledgments

The authors thank Mr. Tsuyoshi Miyashita, Dr. Hirokazu Tanaka, and Dr. Eisuke Sakakibara for their assistance with the experiments, Dr. Daisuke Suzuki, Mr. Michiyuki Fujiwara, and Mr. Shingo Kawasaki for providing technical assistance, Dr. Akiko Obata and Dr. Ryuta Aoki for their helpful comments on the experimental design, and Dr. Shizu Takeda and Dr. Atsushi Maki for their general support. This study was supported by Grants-in-Aid for Scientific Research on Innovative Areas [Nos. 23118001 and 23118004 (Adolescent Mind & Self-Regulation) to KK, No. 32118003 to MF, and Comprehensive Brain Science Network to KK], a Grant-in-Aid for Young Scientists (B) (Nos. 23791309 and 26860914) to RT, a Grant-in-Aid for Scientific Research (B) (No. 23390286) to MF, and a Grant-in-Aid for Challenging Exploratory Research (No. 22659209) to MF from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT). A part of this study was also the result of the "Development of Biomarker Candidates for Social Behavior" interdisciplinary project carried out under the Strategic Research Program for Brain Sciences by MEXT. This study was also supported in part by Health and Labor Sciences Research Grants (H23-seishin-ippan-002 to RT, YN, and MF; H25-seishin-jitsuyoka-ippan-002 to KK; and H25-seishin-ippan-002 to MF) and Health and Labour Science Research Grant on the Practical Application of Medical Technology for Intractable Diseases and Cancer: New development of medical technology for the diagnosis and treatment of psychiatric diseases and cancer by construction of virtual mega-hospital for clinical trials to MF from the Ministry of Health, Labour and Welfare, and Intramural Research Grant for Neurological and Psychiatric Disorders (No. 24-1 and 23-10 to MF and 26-3 to MF and KK) from the National Center for Neurology and Psychiatry. *Conflict of interest.* Hitachi Medical Corporation provided a material support [temporary rental of an fNIRS (Optical Topography) ETG-4000 system] for this study.

References

1. B. Chance et al., "Cognition-activated low-frequency modulation of light absorption in human brain," *Proc. Natl. Acad. Sci. U. S. A.* **90**(8), 3770–3774 (1993).
2. Y. Hoshi and M. Tamura, "Detection of dynamic changes in cerebral oxygenation coupled to neuronal function during mental work in man," *Neurosci. Lett.* **150**(1), 5–8 (1993).
3. F. F. Jöbsis, "Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters," *Science* **198**(4323), 1264–1267 (1977).
4. T. Kato et al., "Human visual cortical function during photic stimulation monitoring by means of near-infrared spectroscopy," *J. Cereb. Blood Flow Metab.* **13**(3), 516–520 (1993).
5. A. Villringer et al., "Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults," *Neurosci. Lett.* **154**(1–2), 101–104 (1993).
6. H. Koizumi et al., "Higher-order brain function analysis by trans-cranial dynamic near-infrared spectroscopy imaging," *J. Biomed. Opt.* **4**(4), 403–413 (1999).
7. A. Maki et al., "Spatial and temporal analysis of human motor activity using noninvasive NIR topography," *Med. Phys.* **22**(12), 1997–2005 (1995).
8. D. A. Boas et al., "Twenty years of functional near-infrared spectroscopy: introduction for the special issue," *NeuroImage* **85**(1), 1–5 (2014).
9. M. Ferrari and V. Quaresima, "A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application," *NeuroImage* **63**(2), 921–935 (2012).
10. F. Scholkmann et al., "A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology," *NeuroImage* **85**(1), 6–7 (2014).
11. T. Suto et al., "Multichannel near-infrared spectroscopy in depression and schizophrenia: cognitive brain activation study," *Biol. Psychiatry* **55**(5), 501–511 (2004).
12. R. Takizawa et al., "Neuroimaging-aided differential diagnosis of the depressive state," *NeuroImage* **85**(1), 498–507 (2014).
13. E. Watanabe et al., "Non-invasive assessment of language dominance with near-infrared spectroscopic mapping," *Neurosci. Lett.* **256**(1), 49–52 (1998).
14. S. Pu et al., "Association between cognitive insight and prefrontal function during a cognitive task in schizophrenia: a multichannel near-infrared spectroscopy study," *Schizophr. Res.* **150**(1), 81–87 (2013).
15. T. Grossmann and M. H. Johnson, "Selective prefrontal cortex responses to joint attention in early infancy," *Biol. Lett.* **6**(4), 540–543 (2010).
16. F. Homae et al., "Development of global cortical networks in early infancy," *J. Neurosci.* **30**(14), 4877–4882 (2010).
17. F. Homae, "A brain of two halves: insights into interhemispheric organization provided by near-infrared spectroscopy," *NeuroImage* **85**(1), 354–362 (2014).
18. Y. Minagawa-Kawai et al., "Prefrontal activation associated with social attachment: facial-emotion recognition in mothers and infants," *Cereb. Cortex* **19**(2), 284–292 (2009).
19. G. Taga et al., "Spontaneous oscillation of oxy- and deoxy-hemoglobin changes with a phase difference throughout the occipital cortex of newborn infants observed using non-invasive optical topography," *Neurosci. Lett.* **282**(1–2), 101–104 (2000).
20. G. Taga et al., "Brain imaging in awake infants by near-infrared optical topography," *Proc. Natl. Acad. Sci. U. S. A.* **100**(19), 10722–10727 (2003).
21. H. Atsumori et al., "Development of wearable optical topography system for mapping the prefrontal cortex activation," *Rev. Sci. Instrum.* **80**(4), 043704 (2009).
22. T. Funane et al., "Synchronous activity of two people's prefrontal cortices during a cooperative task measured by simultaneous near-infrared spectroscopy," *J. Biomed. Opt.* **16**(7), 077011 (2011).
23. Y. Ito et al., "Assessment of heating effects in skin during continuous wave near infrared spectroscopy," *J. Biomed. Opt.* **5**(4), 383–390 (2000).
24. M. Kiguchi et al., "Comparison of light intensity on the brain surface due to laser exposure during optical topography and solar irradiation," *J. Biomed. Opt.* **12**(6), 062108 (2007).
25. T. J. Germon et al., "Sensitivity of near infrared spectroscopy to cerebral and extra-cerebral oxygenation changes is determined by emitter-detector separation," *J. Clin. Monit. Comput.* **14**(5), 353–360 (1998).
26. S. Kohri et al., "Quantitative evaluation of the relative contribution ratio of cerebral tissue to near-infrared signals in the adult human head: a preliminary study," *Physiol. Meas.* **23**(2), 301–312 (2002).

27. L. Minati et al., "Intra- and extra-cranial effects of transient blood pressure changes on brain near-infrared spectroscopy (NIRS) measurements," *J. Neurosci. Methods* **197**(2), 283–288 (2011).
28. P. Smielewski et al., "Clinical evaluation of near-infrared spectroscopy for testing cerebrovascular reactivity in patients with carotid artery disease," *Stroke* **28**(2), 331–338 (1997).
29. T. Takahashi et al., "Influence of skin blood flow on near-infrared spectroscopy signals measured on the forehead during a verbal fluency task," *NeuroImage* **57**(3), 991–1002 (2011).
30. S. N. Davie and H. P. Grocott, "Impact of extracranial contamination on regional cerebral oxygen saturation: a comparison of three cerebral oximetry technologies," *Anesthesiology* **116**(4), 834–840 (2012).
31. H. Sørensen et al., "Cutaneous vasoconstriction affects near-infrared spectroscopy determined cerebral oxygen saturation during administration of norepinephrine," *Anesthesiology* **117**(2), 263–270 (2012).
32. S. B. Erdoğan, M. A. Yücel, and A. Akin, "Analysis of task-evoked systemic interference in fNIRS measurements: insights from fMRI," *NeuroImage* **87**, 490–504 (2014).
33. I. Tachtsidis et al., "False positives in functional near-infrared topography," *Adv. Exp. Med. Biol.* **645**, 307–314 (2009).
34. E. Kirilina et al., "The physiological origin of task-evoked systemic artefacts in functional near infrared spectroscopy," *NeuroImage* **61**(1), 70–81 (2012).
35. L. Gagnon et al., "Improved recovery of the hemodynamic response in diffuse optical imaging using short optode separations and state-space modeling," *NeuroImage* **56**(3), 1362–1371 (2011).
36. R. B. Saager, N. L. Telleri, and A. J. Berger, "Two-detector corrected near infrared spectroscopy (C-NIRS) detects hemodynamic activation responses more robustly than single-detector NIRS," *NeuroImage* **55**(4), 1679–1685 (2011).
37. Q. Zhang, E. N. Brown, and G. E. Strangman, "Adaptive filtering to reduce global interference in evoked brain activity detection: a human subject case study," *J. Biomed. Opt.* **12**(6), 064009 (2007).
38. T. Funane et al., "Quantitative evaluation of deep and shallow tissue layers' contribution to fNIRS signal using multi-distance optodes and independent component analysis," *NeuroImage* **85**(1), 150–165 (2014).
39. S. Kohno et al., "Removal of the skin blood flow artifact in functional near-infrared spectroscopic imaging data through independent component analysis," *J. Biomed. Opt.* **12**(6), 062111 (2007).
40. A. T. Eggebrecht et al., "Mapping distributed brain function and networks with diffuse optical tomography," *Nat. Photonics* **8**(6), 448–454 (2014).
41. D. K. Joseph et al., "Diffuse optical tomography system to image brain activation with improved spatial resolution and validation with functional magnetic resonance imaging," *Appl. Opt.* **45**(31), 8142–8151 (2006).
42. D. A. Boas, A. M. Dale, and M. A. Franceschini, "Diffuse optical imaging of brain activation: approaches to optimizing image sensitivity, resolution, and accuracy," *NeuroImage* **23**(Suppl 1), S275–S288 (2004).
43. X. Cui et al., "A quantitative comparison of NIRS and fMRI across multiple cognitive tasks," *NeuroImage* **54**(4), 2808–2821 (2011).
44. F. B. Haeussinger et al., "Reconstructing functional near-infrared spectroscopy (fNIRS) signals impaired by extra-cranial confounds: an easy-to-use filter method," *NeuroImage* **95**, 69–79 (2014).
45. S. Heinzel et al., "Variability of (functional) hemodynamics as measured with simultaneous fNIRS and fMRI during intertemporal choice," *NeuroImage* **71**(1), 125–134 (2013).
46. R. Aoki et al., "Relationship of negative mood with prefrontal cortex activity during working memory tasks: an optical topography study," *Neurosci. Res.* **70**(2), 189–196 (2011).
47. R. Aoki et al., "Correlation between prefrontal cortex activity during working memory tasks and natural mood independent of personality effects: an optical topography study," *Psychiatry Res.: Neuroimaging* **212**(1), 79–87 (2013).
48. H. Sato et al., "Correlation of within-individual fluctuation of depressed mood with prefrontal cortex activity during verbal working memory task: optical topography study," *J. Biomed. Opt.* **16**(12), 126007 (2011).
49. H. Sato et al., "Intersubject variability of near-infrared spectroscopy signals during sensorimotor cortex activation," *J. Biomed. Opt.* **10**(4), 044001 (2005).
50. S. Abrahams et al., "Functional magnetic resonance imaging of verbal fluency and confrontation naming using compressed image acquisition to permit overt responses," *Hum. Brain Mapp.* **20**(1), 29–40 (2003).
51. W. B. Edmister et al., "Improved auditory cortex imaging using clustered volume acquisitions," *Hum. Brain Mapp.* **7**(2), 89–97 (1999).
52. H. Sato et al., "A NIRS-fMRI investigation of prefrontal cortex activity during a working memory task," *NeuroImage* **83**, 158–173 (2013).
53. D. T. Delpy et al., "Estimation of optical pathlength through tissue from direct time of flight measurement," *Phys. Med. Biol.* **33**(12), 1433–1442 (1988).
54. T. Funane et al., "Greater contribution of cerebral than extracerebral hemodynamics to near-infrared spectroscopy signals for functional activation and resting-state connectivity in infants," *Neurophotonics* **1**(2), 025003 (2014).
55. J. F. Cardoso, "Blind signal separation: statistical principles," *Proc. IEEE* **86**(10), 2009–2025 (1998).
56. T. Katura et al., "Extracting task-related activation components from optical topography measurement using independent components analysis," *J. Biomed. Opt.* **13**(5), 054008 (2008).
57. Y. Fukui, Y. Ajichi, and E. Okada, "Monte Carlo prediction of near-infrared light propagation in realistic adult and neonatal head models," *Appl. Opt.* **42**(16), 2881–2887 (2003).
58. A. Hirasawa et al., "Influence of skin blood flow and source-detector distance on near-infrared spectroscopy-determined cerebral oxygenation in humans," *Clin. Physiol. Funct. Imaging* (2014).
59. G. E. Strangman, Q. Zhang, and Z. Li, "Scalp and skull influence on near infrared photon propagation in the Colin27 brain template," *NeuroImage* **85**(1), 136–149 (2014).
60. T. Yamada, S. Umeyama, and K. Matsuda, "Separation of fNIRS signals into functional and systemic components based on differences in hemodynamic modalities," *PLoS One* **7**(11), e50271 (2012).
61. T. Funane et al., "Dynamic phantom with two stage-driven absorbers for mimicking hemoglobin changes in superficial and deep tissues," *J. Biomed. Opt.* **17**(4), 047001 (2012).
62. C. Rorden and M. Brett, "Stereotaxic display of brain lesions," *Behav. Neurol.* **12**(4), 191–200 (2000).
63. A. Sassaroli et al., "Spatially weighted BOLD signal for comparison of functional magnetic resonance imaging and near-infrared imaging of the brain," *NeuroImage* **33**(2), 505–514 (2006).
64. F. Aletti et al., "Deep and surface hemodynamic signal from functional time resolved transcranial near infrared spectroscopy compared to skin flowmotion," *Comput. Biol. Med.* **42**(3), 282–289 (2012).
65. A. Seiyama et al., "Circulatory basis of fMRI signals: relationship between changes in the hemodynamic parameters and BOLD signal intensity," *NeuroImage* **21**(4), 1204–1214 (2004).
66. I. Tachtsidis et al., "Investigation of frontal cortex, motor cortex and systemic haemodynamic changes during anagram solving," *Adv. Exp. Med. Biol.* **614**, 21–28 (2008).

Tsukasa Funane is a researcher at Central Research Laboratory, Hitachi, Ltd., Japan. Since 2006, he has been a member of the research group working on optical topography, optical brain function monitoring technology based on near-infrared spectroscopy (NIRS). His main responsibilities include basic research on new measurement methods as well as hardware developments and their application to human brain study such as social cognitions.

Hiroki Sato received his BA degree from International Christian University in 1998, his MA degree from the University of Tokyo in 2000, and his PhD degree from Keio University in 2006. He is currently a senior researcher of Central Research Laboratory, Hitachi, Ltd., and promotes the basic research to develop new applications of near-infrared spectroscopy (NIRS) in a wide range of fields.

Noriaki Yahata received his BA degree from the University of Tokyo, Japan, in 1993 and his PhD degree from the State University of New York in 2001. He was an assistant professor at the Department of Pharmacology, Nippon Medical School (2005 to 2008) and the Department of Neuropsychiatry, the University of Tokyo (2009 to 2012). Currently, he is an assistant professor at the Department of Youth Mental Health, the University of Tokyo. His research interests include development of a neuroimaging-based biomarker for neuropsychiatric disorders.

Ryu Takizawa received his BA degree in psychology in 1999 and his MD degree in 2003. After clinical training in the Department of Neuropsychiatry, the University of Tokyo, he received his PhD degree in medicine in 2010. Currently, he is an assistant professor at the Department of Neuropsychiatry, the University of Tokyo, and a Newton International fellow, Institute of Psychiatry, King's College London. His interests include studies on clinical biomarkers and gene-environmental interplays in mental health from a life-course developmental perspective.

Yukika Nishimura received her BA degree in experimental psychology from Keio University in 2001, her MSc degree in 2003, and her PhD degree in 2007 in medical science from Mie University, Japan. She was a research resident of the Japan Foundation for Neuroscience and Mental Health (2008 to 2011), and she is currently a project research associate in the Department of Neuropsychiatry at the University of Tokyo, Japan. Her research interest is the cognitive neuroscience of psychiatric disorders.

Akihide Kinoshita received his MD degree from Tokyo Medical University in 2006 and his PhD degree from the University of Tokyo, Japan, in 2014. After clinical training in psychiatry at the Tokyo Metropolitan Bokutoh Hospital, he received clinical research training in neuroimaging in the Department of Neuropsychiatry, Graduate School of Medicine, University of Tokyo, Japan. He engages in research in the Department of Neuropsychiatry, the University of Tokyo, and his major research interest is neuroimaging in schizophrenia.

Takusige Katura is a researcher at Central Research Laboratory, Hitachi, Ltd., Japan. Since 2001, he has been a member of the research group working on optical topography, optical brain function monitoring technology based on near-infrared spectroscopy (NIRS). His main responsibilities include basic research on new measurement methods as well as signal analysis and its application to human brain study such as social cognitions.

Hirokazu Atsumori is a researcher in the Central Research Laboratory, Hitachi, Ltd., Japan. He has been working on the research and development of optical topography, a functional neuroimaging technique based on near-infrared spectroscopy, since 2002. He is now engaged in the development of a wearable and compact optical topography system for monitoring prefrontal cortex activities and its application to new research fields.

Masato Fukuda received his MD degree in 1983 and his PhD degree in 1997 from the University of Tokyo, Japan. His professional achievements include being an assistant professor in the Department of Neuropsychiatry, the University of Tokyo, and an associate professor in the Department of Psychiatry and Neuroscience, Gunma University. He is currently the professor and chair in the Department of Psychiatry and Neuroscience, Gunma University, and his major research interest is clinical neurophysiology and neuroimaging in psychiatry.

Kiyoto Kasai received his MD degree in 1995 and his PhD degree in 2004 from the University of Tokyo, Japan. After clinical training in psychiatry at the University of Tokyo Hospital and National Center of Neurology and Psychiatry, he received clinical research training in neuroimaging in psychiatry at Harvard Medical School. He is now the professor and chair in the Department of Neuropsychiatry, the University of Tokyo, and his major research interest is clinical neurophysiology and neuroimaging in schizophrenia.

Hideaki Koizumi joined Hitachi, Ltd. in 1971 after receiving his BA degree from the University of Tokyo [PhD (physics), 1976]. He is a fellow and corporate officer of Hitachi, Ltd., a vice president of the Engineering Academy of Japan, a member of the Science Council of Japan, and a foreign member of the Chinese Academy of Engineering. He has proposed many new concepts in human security and well-being and methodologies, especially in the field of spectroscopy.

Masashi Kiguchi has studied various optical measurements: nonlinear spectroscopy, time-resolved spectroscopy, near-field spectroscopy, and near-infrared spectroscopy, and his background is physics and laser spectroscopy. He has studied the problems related to the principle of near-infrared spectroscopy (NIRS) measurement and has been taking the lead in the development of new techniques for observing brain activities to open new research fields and in basic studies for putting them to practical use.

Journal of Psychiatric Research 57 (2014) 74–83

Contents lists available at ScienceDirect

Journal of Psychiatric Research

journal homepage: www.elsevier.com/locate/psychires

Near-infrared spectroscopic study of frontopolar activation during face-to-face conversation in major depressive disorder and bipolar disorder

Yuichi Takei*, Masashi Suda, Yoshiyuki Aoyama, Noriko Sakurai, Minami Tagawa, Tomokazu Motegi, Miho Yamaguchi, Kosuke Narita, Masato Fukuda

Department of Psychiatry and Neuroscience, Gunma University of Graduate School of Medicine, 3-39-22 Showa, Maebashi, Gunma 371-8511, Japan

ARTICLE INFO

Article history:
Received 31 January 2014
Received in revised form
11 June 2014
Accepted 18 June 2014

Keywords:
Major depression
Bipolar disorder
Mood disorder
Talk
Near-infrared spectroscopy
Social cognition

ABSTRACT

Major depressive disorder (MDD) and bipolar disorder (BD) patients show speech characteristics that vary greatly according to mood state. In a previous study, we found impaired temporal and right inferior frontal gyrus (IFG) activation in schizophrenia during face-to-face conversation; no study had, however, previously investigated mood disorders during face-to-face conversation. Here, we investigated frontal and temporal lobe activation during conversation in patients with MDD and BD. Frontal and temporal lobe activation was measured using near-infrared spectroscopy (NIRS) in 29 patients with MDD, 31 patients with BD, and 31 normal controls (NC). We compared continuous activation and rapid change of activation with talk/listen phase changes during the conversation and analyzed the correlation between these indices and clinical variables. Both the MDD and BD groups showed decreased continuous activation in the left dorsolateral prefrontal (DLPFC) and left frontopolar cortices (FPCs); they also showed decreased rapid change in bilateral FPC activation. In the MDD group, the rapid change of activation was positively correlated with Global Assessment of Functioning (GAF) scores. In the BD group, continuous activation was negatively correlated with age of onset. These results indicate that frontal activation during conversation decreases in both MDD and BD. However, both continuous activation and rapid change may reflect the pathophysiological character of MDD and BD; in particular, the reduced amount of rapid change in the right FPC may be related to impaired adaptive ability in MDD.

© 2014 Elsevier Ltd. All rights reserved.

1. Objectives of the study and background

Psychiatrists investigating mental illness can diagnose specific diseases using various characteristics exhibited by patients during the course of a conversation. Patients suffering from major depressive disorder (MDD) and bipolar disorder (BD) show particular conversational characteristics, which can be divided into those that do or do not change according to disease state (American Psychiatric Association, 1994; Bouhuys and Sam, 2000).

In the depressive state, patients use few words, show psychomotor retardation, and exhibit poor choices of conversation topics. Their emotional reactivity becomes poor, and smiles are not exhibited even when discussing pleasant topics. In contrast, in a manic or hypomanic state, patients with BD are more talkative than usual and speak one-sidedly; it may be difficult to understand the content of their speech due to the manifestation of flights of ideas. In the euthymic state, these symptoms disappear. Although these state-dependent characteristics of conversation are not specific to MDD and BD, it is essential to observe them because they are key to diagnosing these disorders using the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM–IV; American Psychiatric Association, 1994). These state-dependent characteristics are crucial in clinical assessments (Hamilton Rating Scale for Depression; Hamilton, 1960; Young Mania Rating Scale; Young et al., 1978).

There are several studies that investigated links between personality and social communication, as well as the direct communication between patients and interviewers. Patients with BD and MDD in the euthymic state have been found to differ from normal control (NC) participants in that they show high scores for harm avoidance and/or low scores for self-directedness and cooperativeness, as measured by the Temperament and Character Inventory (TCI; Celikel et al., 2009; Hansenne et al., 1999). With

* Corresponding author. Tel./fax: +81 27 220 8183.
E-mail address: tyuichi@gunma-u.ac.jp (Y. Takei).

http://dx.doi.org/10.1016/j.jpsychires.2014.06.009
0022-3956/© 2014 Elsevier Ltd. All rights reserved.

regards to direct communication, Coyne et al. stated that deficits in human communication play an important role in theories of development, persistence, and recurrence of depression (Coyne and Downey, 1991). Bouhuys et al. indicated that the lack of coordination during MDD interviews may be a risk factor for the condition's recurrence (Bouhuys and Sam, 2000). These characteristics may be associated with vulnerability in patients with mood disorders.

Cognitive and emotional dysfunctions underlying these characteristic conversational differences between healthy adults and mood disorder patients have been previously investigated by functional magnetic resonance imaging (fMRI). These studies suggest an altered activation of the amygdala, as well as the frontal, cingulate, and temporal cortices in patients with MDD and BD during various cognitive tasks (Savitz and Drevets, 2009). Although such data may assist in diagnosis, and may be directly related to the Global Assessment of Functioning (GAF), brain activation during conversation has not yet been investigated in patients with MDD and BD due to methodological difficulties.

Near-infrared spectroscopy (NIRS) is a recently developed noninvasive functional neuroimaging technique (Koizumi et al., 1999; Strangman et al., 2002). NIRS can detect regional cerebral blood volume (rCBV) changes through the fluctuating concentrations of oxyhemoglobin ([oxy-Hb]) and deoxyhemoglobin ([deoxy-Hb]). NIRS has some advantages over other functional neuroimaging methodologies due to (i) its complete noninvasiveness, enabling repeated measurements; (ii) the portability and compactness of the NIRS apparatus, enabling measurements under natural conditions with participants comfortably seated; and (iii) little sensitivity to motion artifacts, allowing the NIRS to be used in the study of conversation. Considering these advantages, NIRS allows for brain activation to be evaluated in a naturalistic environment. Indeed, several recent studies reported the use of NIRS during face-to-face interaction (Costantini et al., 2013; Cui et al., 2012; Konvalinka and Roepstorff, 2012).

Our group previously investigated frontal and temporal lobe activation during face-to-face conversations with normal control participants (Suda et al., 2010, 2011). Our results showed activity in the frontal and temporal lobes, as well as substantial cyclical activity in the frontal lobe corresponding to the time course of the conversation task — where participants were required to talk to a person facing them — especially around the frontopolar region. Both of these studies assessed only the sustained activities as grand average [oxy-Hb] data (GAOD) of the conversation and control tasks. In the present study, we sought to assess this substantial cyclical activity, measured by the averaged amount of [oxy-Hb] change (AAOC) over time. We measured this fluctuation in activity during the speech and listening phases superimposed on the base large [oxy-Hb] change during the task in order to evaluate the effects of switching turns during the conversation and control tasks. Although the meaning of the AAOC has not yet been clarified, we presumed that there are different physiological meanings assigned to the GAOD and AAOC. We posited that the GAOD represents the global functioning of an individual required to face to another person (i.e., a conversation situation-related function), while the AAOC mainly reflects speech itself (i.e., a speech-related function).

Because the characteristics of BD and MDD are well reflected in conversation, we hypothesized that (i) the GAOD and AAOC of the frontal cortex in MDD and BD would be altered compared to NC participants; and (ii) the GAOD or AAOC of the frontal cortex in MDD and BD would correlate with both GAF and depressive symptoms.

Table 1
Participant characteristics.

Sex (male)	MDD ($n = 29$)			BD ($n = 31$)			NC ($n = 31$)	
	M	F		M	F		M	F
	14	15		14	17		11	20
	Mean	SD		Mean	SD		Mean	SD
Age	34.5	9.0		34.9	6.6		33.6	10.0
Age range	19–51			20–45			23–58	
Age of onset	30.3	8.9		26.0	6.4		–	–
Illness duration	4.1	3.6		9.7	7.0		–	–
HRSD	9.8	4.4		6.4	5.5		–	–
YMRS	–	–		1.9	3.6		–	–
GAF	56.7	8.2		54.1	12.5		–	–
Subtype	–	–		BP I	1/31		–	–
	–	–		BP II	30/31			
	Mean	SD	n	Mean	SD	n		
Antidepressant (imipramine equivalent dose mg/day)	72.0	50.6	22/29	122.7	85.2	17/31	–	–
Antipsychotic (chlorpromazine equivalent dose mg/day)	112.5	63.5	6/29	213.5	122.8	11/31	–	–
Anxiolytic (diazepam equivalent dose mg/day)	7.1	6.4	14/29	12.6	17.2	15/31	–	–
Hypnotic (flunitrazepam equivalent dose mg/day)	2.0	1.5	16/29	2.5	2.0	16/31	–	–
Lithium mg/day	300.0	–	1/29	618.8	242.8	16/31	–	–
Carbamazepine mg/day	0	–	0/29	500.0	258.2	4/31	–	–
Valproate mg/day	300.0	200.0	2/29	552.9	194.0	17/31	–	–

M, male; F, female; MDD, major depressive disorder; BD, bipolar disorder; NC, normal controls; BPI, bipolar I disorder; BPII, bipolar II disorder; HRSD, 17-item Hamilton Rating Scale for Depression; YMRS, Young Mania Rating Scale; GAF, The Global Assessment of Functioning.

2. Materials and methods

2.1. Sample

Twenty-nine patients with MDD, 31 patients with BD, and 31 healthy volunteers (NC) were recruited from the Department of Psychiatry and Neuroscience of Gunma University Hospital in Japan to participate in this study. All participants were right-handed native Japanese speakers. Participants had been previously diagnosed with MDD or BD according to the DSM-IV criteria (American Psychiatric Association, 1994). Patients over 60 years old were not included in this study in order to eliminate the effects of additional pathophysiological factors such as aging and possible cerebrovascular changes. Depressive symptoms among patients with MDD and BD were evaluated using the 17-item Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960). Manic symptoms in BD patients were evaluated using the Young Mania Rating Scale (YMRS; Young et al., 1978). Nearly all participants were on medication such as mood stabilizers, antipsychotics, antidepressants, anxiolytics, and/or hypnotics. Equivalent dosages were calculated for each class of medication, as follows: chlorpromazine-equivalent dose of antipsychotics; imipramine-equivalent dose of antidepressants; diazepam-equivalent dose of anxiolytics; and flunitrazepam-equivalent dose of hypnotics (Inagaki and Inada, 2006).

Exclusion criteria for MDD and BD groups were determined using clinical interviews. Criteria included any observable brain abnormality in magnetic resonance imaging (MRI) results, past or present neurological illness, traumatic brain injury with any of the known cognitive consequences or loss of consciousness for more than 5 min, substance use or addiction, and present hearing or vision impairment. Exclusion criteria for the NC participants were also determined using clinical interviews. These included a history of any major psychiatric disorder or major physical illness and a current prescription for major psychiatric medication.

This study was performed in accordance with the Helsinki Declaration, as revised in 1989, and was approved by the Institutional Review Board of Gunma University Hospital. Written informed consent was obtained from all participants prior to study initiation. For patients under 20 years of age or those who had been forcibly committed to hospital, written informed consent was obtained from the patient's legal representative. Because we could not obtain the behavioral data from conversations with participants who had not provided consent for videotape recordings, we describe the clinical characteristics of all participants using the behavioral data listed in Table 1.

2.2. Activation tasks

Two types of activation tasks—a conversation and a control task—were used to assess brain activation during conversation (Fig. 1). The order of the two tasks was counterbalanced between participants.

2.3. Conversation task

The conversation task, which comprised the pre-task, task, and post-task segments, was designed to simulate typical, daily life conversation in an experimental setting. The interviewers who led the conversation task had been selected from hospital staff that had not been previously acquainted with the participants. Each session began after NIRS probes had been placed on the participants' frontal and temporal regions while they were seated in a comfortable chair, face-to-face (1 m apart) with an interviewer. To eliminate the possible influence of facial cues before and after engaging in conversation, a partition was placed between the participant and interviewer during the pre-task and post-task segments, and was removed during the task.

To prevent the emergence of qualitative and quantitative differences in the nature of conversations from one participant to the next, all participants were instructed to engage in face-to-face conversation with the interviewer during the task segment according to 2 criteria. First, they were to follow a set time course of conversation, according to which the participant and the interviewer alternated speaking every 15 s; this plan was reinforced through spoken cues about the elapsed time from the experimenter every 5 s. The task thus consisted of 6 cycles of 30-s speech segments, with the entire conversation lasting as long as 180 s. Second, participants were to limit the subjects of the conversation to food-related topics. During the pre-task and post-task segments, participants were instructed to repeat the syllables /a/, /i/, /u/, /e/, and /o/, that is, the Japanese counterparts of the English sounds "A," "B," and "C," to exclude the effects of phonation and stabilize baseline conditions. The conversations of 23 patients with MDD, 17 patients with BD, and 28 NCs who had given consent for recording were videotaped for later analysis of visual and audio data.

Conversation task performance was evaluated both quantitatively and qualitatively. First, the amount of conversation contributed by participants was quantitatively evaluated using speaking time (ST)—the duration of the participants' speech as measured by videotaped data analysis. Second, the content of their speech was qualitatively evaluated by 4 expert psychiatrists in terms of the receiving aspect (RS), or the appropriateness of speech content in the context of a conversation; and the sending aspect (SS), or the extent of production of new topics. To measure the RS, the participants' replies to the preceding interviewer's speech were scored as inappropriate, partially inappropriate, partially appropriate, or

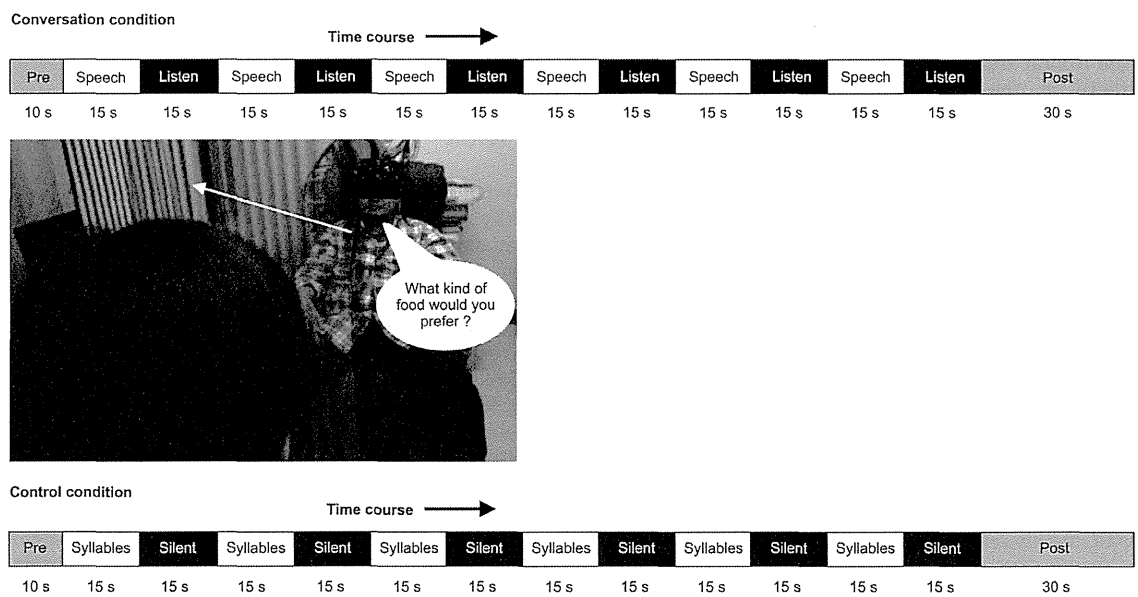

Fig. 1. Task procedures and picture of the measurement setting. Two types of activation task conditions, a conversation condition and a control condition, were employed in this study. Both tasks consisted of three segments: pre-task, task, and post-task. Under the conversation condition, participants were required to speak with the interviewer facing them during the task segment. The task period consisted of six cycles of such 30-s talks; the total conversation lasted up to 180 s. Before starting (pre-task) and after finishing the experiment (post-task), the participant and interviewer were separated by a partition so that they could not see each other. Under the control condition, participants were instructed to repeat meaningless syllables during their turns to speak during the task period. As shown in the picture, the participant is wearing the near-infrared spectroscopy probe on his forehead, and the participant and interviewer are seated facing each other.

appropriate. To measure the SS, the participants' questions to the interviewer were analyzed to determine whether they had introduced no new topic(s), nearly the same topic(s), partially new topic(s), or completely new topic(s).

2.4. Control task

To examine brain activation and artifact contamination induced by phonation alone, the participants were instructed to perform a control task consisting of repeating meaningless syllables (e.g., "a," "ka," "sa," "ta," and "na") during their turn to speak over the course of the task segment of the conversation task. All participants were physically able to repeat the syllables without interruption.

2.5. NIRS measurement

Changes in [oxy-Hb] were measured as an index of changes in cerebral blood volume and in [deoxy-Hb] using a 52-channel NIRS machine (Hitachi ETG-4000; Hitachi Medical Systems, Tokyo, Japan). The correspondence between NIRS channels and measurement points on the cerebral cortex was confirmed by comparison with the results of a multi-participant study of anatomical cranio-cerebral correlation (Okamoto et al., 2004), and was displayed on the basis of the results obtained using the virtual registration method (Fig. 2; Tsuzuki et al., 2007). For details about NIRS measurement, refer to Supplement No. 1.

2.6. Data analyses

We calculated Ebel's intraclass correlation coefficient (Ebel's ICC) for SS and RS of both the groups to investigate inter-rater reliability. The behavioral data (ST, RS, and SS) were compared between the three groups using a one-way analysis of variance (ANOVA). The waveforms of [Hb] changes in all 52 channels during the conversation and control conditions were calculated for all participants. NIRS data from channels 1–21 that clearly contained motion artifacts, as determined by close observation of the participants, were excluded from further analyses. The most common causes of artifacts were large movement and poor contact with the probes due to abundant hair. A probe placed on an area with abundant hair could not be securely fastened to the head and was easily displaced.

The GAOD was calculated for both the conversation and control tasks by averaging by each channel and task, excluding the pre- and post-task segments. The GAOD of the conversation and control tasks were analyzed using a mixed-design repeated measures ANOVA with diagnosis (MDD, BD, or NC) as the between-subjects variable and task (conversation or control task) as the repeated measure variable. Results were corrected for the number of channels using a false discovery rate (FDR) correction to avoid Type I errors. When an interaction was found, Scheffé's post hoc *t*-test with diagnosis was conducted for both conditions ($p < 0.05$).

The AAOC was calculated to evaluate the effect of switching turns during the conversation and control tasks (Fig. 2). The first two time segments were excluded from the analysis because they served as baseline measurements. The obtained amounts of [oxy-Hb] change value were averaged by task. The AAOC of the conversation and control tasks were analyzed using a mixed-design repeated measures ANOVA, with diagnosis (MDD, BD, or NC) as the between-subjects variable and task (conversation or control task) as the repeated measure variable. Results were corrected for the number of channels using the FDR correction ($p < 0.05$). When an interaction was indicated, Scheffé's post hoc *t*-test with diagnosis was conducted for both conditions ($p < 0.05$).

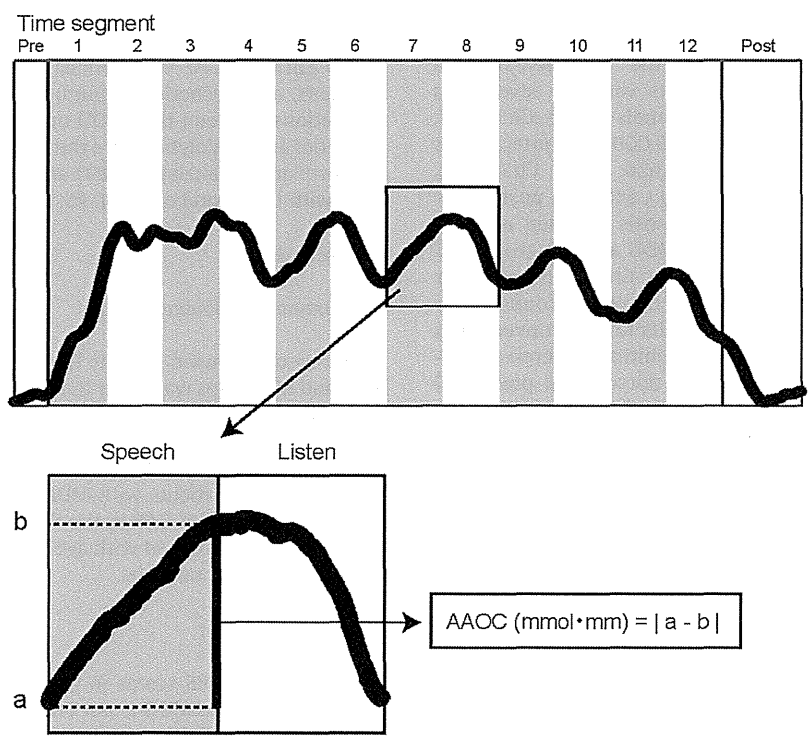

Fig. 2. Formula for calculating the averaged amount of [oxy-Hb] change (AAOC). In this formula, "a" is the [oxy-Hb] data at segment start, and "b" is the [oxy-Hb] data at segment end. The upper panel shows the waveform of an individual during the conversation task. The segment represented by the square in the top image is enlarged in the lower image. The gray panels indicate the speech phase while the white panels indicate the listening phase.

Pearson's r was calculated for both correlations between GAOD/AAOC and behavioral data, which included ST, RS, SS, current age, age of onset, illness duration, GAF score, HRSD score, YMRS score, and drug dosage. Results were corrected for the number of channels by way of FDR correction ($p < 0.05$).

3. Results

3.1. Sample characteristics

Table 1 summarizes the demographic characteristics of the sample. Age and sex ratios were not significantly different between the three groups ($F[2, 90] = 0.2$, $p = 0.83$; chi square (2) = 1.1, $p = 0.58$). At the time of the study, the Patients with MDD were euthymic or depressive, as indicated by their scores on the HRSD. Patients with BD were euthymic, depressive, or hypomanic, as indicated by their scores on the HRSD and the YMRS.

3.2. Behavioral data analysis (Table 2, Fig. 3)

Ebel's ICC for the RS of the NC ($r = 0.441$, $p = 0.023$) and MDD ($r = 0.718$, $p = 0.000$) groups indicated significant correlations, while that of the BD group ($r = 0.392$, $p = 0.084$) indicated a trend–level correlation. Ebel's ICC for SS of the NC ($r = 0.712$, $p = 0.000$), MDD ($r = 0.524$, $p = 0.010$), and BD ($r = 0.607$ $p = 0.005$) groups indicated significant correlations.

The results of the one-way ANOVA for RS revealed a significant main effect of diagnosis, and Scheffé's post hoc t-test of the RS revealed a lower score for MDD patients than controls (Table 2). The ST and SS were not significantly different between the three groups.

3.3. Analysis of GAOD during conversation and control tasks

The results of the mixed-design repeated measures ANOVA of the GAOD, with diagnosis as the between-subjects variable and task as the within-subjects variable, revealed a significant main effect of task in 31 channels (Ch22–Ch52; $F[1, 85] = 16.22$–143.01; FDR-corrected $p = 0.000$) and interactions of diagnosis and task in 3 channels (Ch28, Ch39, Ch50; $F[2, 85] = 7.16$–8.16; FDR-corrected $p = 0.001$). Scheffé's post hoc t-test of the conversation task with diagnosis revealed a larger activation in the NC group than in the MDD and BD groups for Ch28 and Ch39, and in the BD group for Ch50. On the other hand, there was no significant difference between groups in the control task. Scheffé's post hoc t-test results for the conversation task indicated that the brain areas exhibiting differences between the groups tended to be the left dorsolateral prefrontal cortex (DLPFC), frontopolar cortex (FPC), and inferior frontalgyrus (IFG), according to the virtual registration method (Fig. 4).

3.4. AAOC during conversation and control tasks (Fig. 5)

The results of the mixed-design repeated measures ANOVA of the AAOC, with diagnosis as the between-subjects variable and task as the within-subjects variable, revealed a significant main effect of task in 31 channels (Ch22–Ch52; $F[1, 85] = 8.63$–62.63; FDR-corrected p: 0.000–0.004) and interactions of diagnosis and task in 7 channels (Ch25–Ch29, Ch36, Ch37; $F[2, 85] = 4.74$–7.58; FDR-corrected p: 0.001–0.011). Scheffé's post hoc t-test of the conversation task with diagnosis revealed a larger activation in the NC group than in the MDD and BD groups for Ch25–Ch28 and Ch37, and in MDD for Ch29 and Ch36. No significant difference was found between groups in the control task.

Table 2
Behavioral data results.

	MDD		BD		NC		F	p
	Mean	SD	Mean	SD	Mean	SD		
ST (s)	73.2	12.0	74.1	10.2	77.7	4.9	1.736	0.184
RS	3.6	0.4	3.7	0.3	3.9	0.2	4.863	0.011
SS	3.2	0.6	3.2	0.5	3.3	0.7	0.164	0.849

MDD, major depressive disorder; BD, bipolar disorder; NC, normal controls; ST, speaking time; RS, scoring of qualitative evaluation of receiving aspects; SS, scoring of qualitative evaluation of sending aspects.

3.5. Correlation analysis of brain activation, clinical symptoms, medications, and behavioral data (Fig. 6)

NC group: The GAOD during the conversation task was not correlated with either behavioral data (ST, RS, and SS) or current age. The AAOC during the conversation task was likewise not correlated with behavioral data (ST, RS and SS) or current age.

MDD group: The AAOC during the conversation task was positively correlated with GAF in 3 PFC channels: Ch36, Ch38, and Ch46 ($r = 0.53$–0.61; FDR-corrected p: 0.000–0.003, Fig. 6). However, the GAOD during the conversation task was not correlated with GAF, and HRSD was correlated with neither GAOD nor AAOC. The GAOD and AAOC during the conversation task were not correlated with behavioral data (ST, RS, and SS), current age, age of onset, illness duration, or drug dosage (antidepressant and anxiolytics). HRSD score was not correlated with behavioral data (ST, RS, and SS).

BD group: We did not find significant correlations between the GAOD or AAOC and GAF and HRSD. However, we found a significant negative correlation between the GAOD during the conversation task and current age in 2 channels (Ch37 and Ch39; $r = -0.58$ to -0.55; FDR-corrected $p = 0.001$), and also age of onset in 3 channels (Ch35, Ch42, Ch43; $r = -0.63$ to -0.48; FDR-corrected p: 0.000–0.008, Fig. 6). Neither GAOD nor AAOC during the conversation task were correlated with behavioral data (ST, RS, and SS), illness duration, GAF score, HRSD score, or drug dosage (antidepressants, antipsychotics, lithium, and valproate). Any significant correlations between the GAOD or AAOC and YMRS or dosage of hypnotics and anxiolytics were the result of the effect of outliers on the correlational analysis. HRSD and YMRS score were not correlated with behavioral data (ST, RS, and SS).

4. Discussion

4.1. Summary of obtained results

Our results showed that the GAOD in the left DLPFC and FPC of MDD and BD groups was lower than that of the NC group during the conversation task, and that the AAOC in the FPC of both the MDD and BD groups decreased during the conversation task. Both the GAOD and AAOC during the control task did not differ between the 3 groups. In the patients with MDD, the AAOC tended to be positively correlated with GAF in the FPC. In the BD group, the GAOD was negatively correlated with age of onset in the right DLPFC and both middle temporal lobes.

4.2. Behavioral data

We found low RS scores in patients with MDD. This result is reasonable, because MDD is not only associated with various neuropsychological deficits (to memory, learning, attention, alertness, and executive functions) during a depressive episode, but also during a euthymic state (Austin et al., 1992; Tham et al., 1997; Veiel, 1997; Zakzanis et al., 1998). However there were no significant

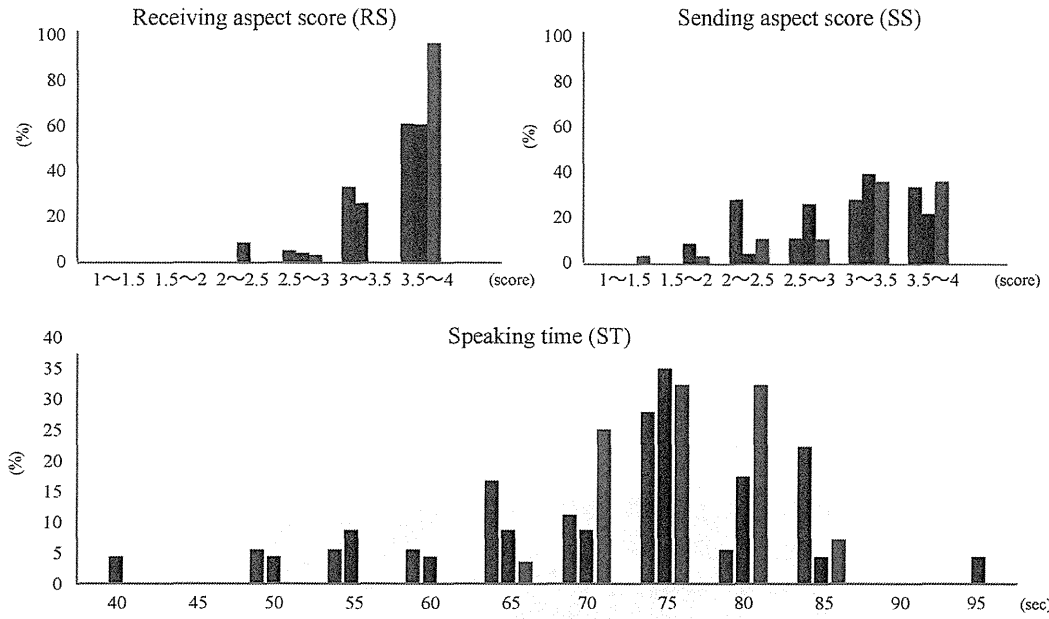

Fig. 3. Percent histogram of behavioral data results: averaged receiving aspect score (RS, left) and sending aspect score (SS, right), as evaluated by two expert psychiatrists, and speaking time (ST, bottom). The x-axes of the left and right figures indicate RS and SS, and that of the bottom figure indicates ST. The y-axes of the three figures indicate the percentage of subjects for each score. The green bar represents normal controls; the blue bar, patients with major depression; and the red bar, patients with bipolar disorder.

differences in ST and SS among three groups, potentially due to a ceiling effect: almost all participating patients were in the euthymic state, and performed well. The lack of correlation between behavioral data and HRSD or YMRS in both MDD and BD patients may also be due to patient state.

4.3. Difference in physiological meaning between the GAOD and the AAOC

We found that the frontal and temporal regions were continuously activated during face-to-face conversation. In addition to the blockwise activations of the frontal and temporal lobes during the 180-s conversation task, rhythmic activations tended to be superimposed on the FPC area, as shown in Fig. 2. In our previous studies, we obtained similar waveform data (Suda et al., 2010, 2011).

Although we cannot clearly explain the cognitive meanings of the GAOD and AAOC, we posit that their physiological meanings differ. In terms of task construction, the GAOD is obtained by subtracting the response to repeated syllables from responses to all the cognitive elements necessary for speech, such as context processing, monitoring, referencing, emotional processing, and mentalizing. In terms of waveform patterns, this basic activation is stable from beginning to end during the conversation. Thus, the GAOD may represent the global function required to interact with another person and may be related to mentalizing as well as processing various types of information received from another person (i.e., conversation situation-related function).

The AAOC is obtained by subtracting the response to the ending time point data of the listening phase from the ending time point data of the talk phase. However, it is difficult to compare these results with those of previous studies since prior research did not utilize the same analytical techniques; therefore, we interpret the physiological meaning of AAOC solely within the context of the task design. In the present study, participants needed to speak, monitor their speech, and change speech content in response to the partner's statements during the conversation; thus, we considered the AAOC to be a speech-related function.

4.4. Decreased GAOD and clinical variables in MDD and BD

The GAOD decreased in the left DLPFC and FPC in both MDD and BD groups, and in the IFG in the BD group, when compared to the NC group. In the BD group, the GAOD was negatively correlated with age of onset in the right DLPFC and both middle temporal lobes, and with current age in the FPC.

Various neuroimaging studies have found decreased frontal activation in patients with MDD (Drevets et al., 2002; Kameyama et al., 2006; Kennedy et al., 2001; Ohta et al., 2008; Suto et al., 2004) and BD (Frangou et al., 2008; Kronhaus et al., 2006; Roth et al., 2006; Strakowski et al., 2005). These findings support the decreased GAOD and AAOC of the frontal regions in MDD and BD. The GAOD in the left DLPFC and FPC of MDD and BD groups tended to be lower. In our previous study, when compared to NC, patients with schizophrenia exhibited decreased activation in the bilateral temporal lobes and the right IFG during conversation (Takei et al., 2013). These results indicate a hemodynamic difference that reflects conversation situation-related function in MDD, BD, and schizophrenia patients. This different regional pattern of mood disorders (MDD and BD) and schizophrenia is supported by meta-analyses of volumetric MRI studies (Arnone et al., 2009; Koutsouleris et al., 2008; Sacher et al., 2012). These previous studies reported frontal volume reduction of MDD, BPD, and schizophrenia; temporal volume reduction was also emphasized in schizophrenia. The GAOD reflects multiple functions related to conversation situations; therefore, it may reflect the impaired brain region central to the pathophysiology of the disorder. To be clear, the lack of difference in the GAOD and AAOC during the control task indicates that the obtained difference was not due to brain volume reduction or phonation.

In the BD group, the GAOD was negatively correlated with age of onset in the right DLPFC and both middle temporal lobes, and with

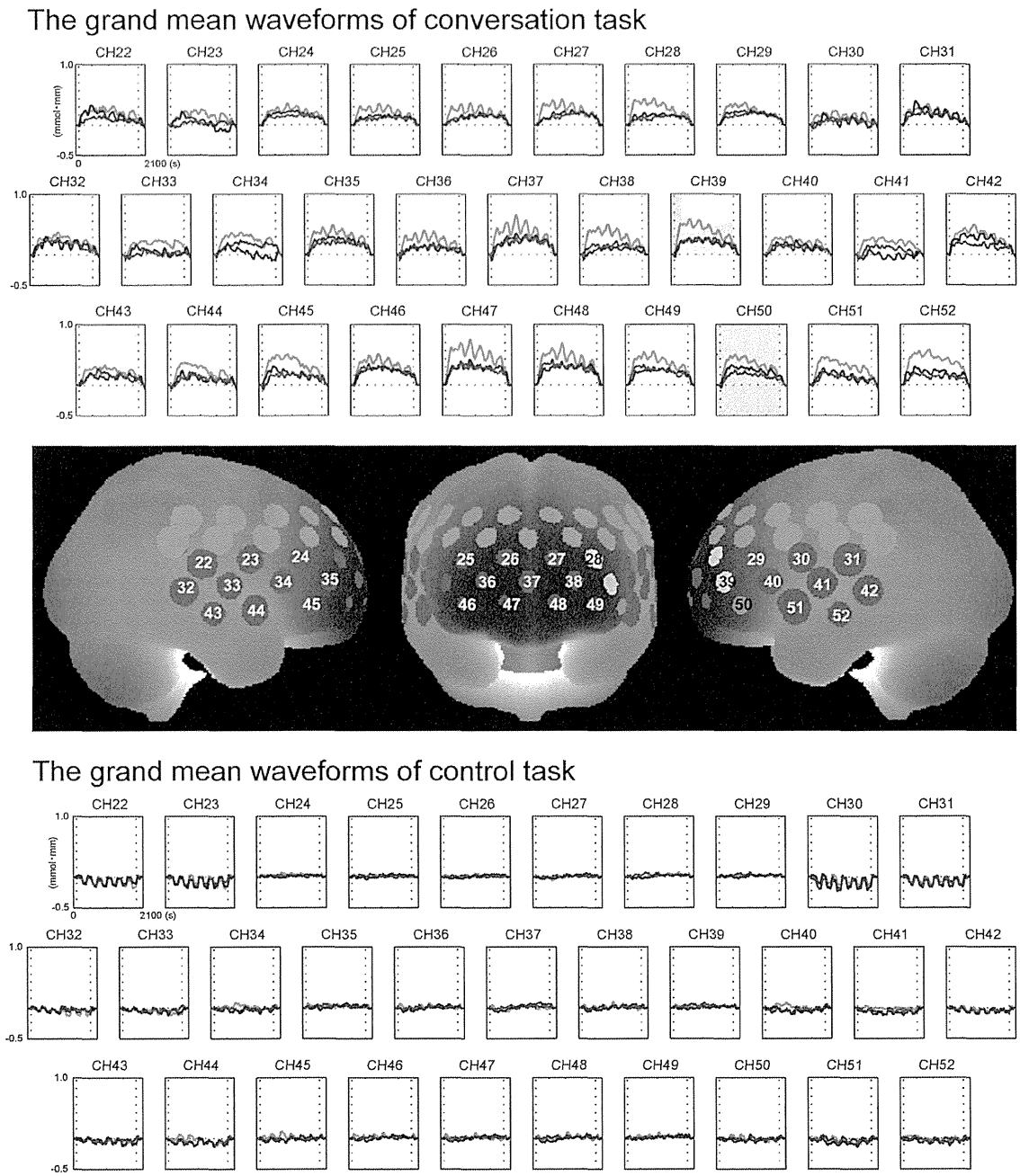

Fig. 4. The upper 31 figures (Ch22–Ch52) indicate the grand mean waveforms of the conversation task: green line, control participant; blue line, major depression; red line, bipolar disorder. The yellow channels of the upper figures indicate a larger grand averaged [oxy-Hb] data (GAOD) in the NC group than in the MDD and BD groups, and the pink channel indicates a larger GAOD in the NC group than in the BD by mixed-design repeated measures ANOVA, followed by Scheffé's post hoc *t*-test. The middle 3 figures below show the probabilistic estimation and anatomical labeling of the locations of NIRS channels in the standard brain space in accordance with Tsuzuki et al. (2007), and the yellow and pink areas indicate the corresponding brain areas that differed between the groups according to the results of the post hoc *t*-test. Light gray channels without a number are channels that were excluded because of detection of clear motion artifacts. The lower 31 figures (Ch22–Ch52) indicate the grand mean waveforms of the control task: green line, control participant; blue line, major depression; red line, bipolar disorder.

current age in the FPC, but not correlated with illness duration. Past epidemiological research suggests that early- and late-onset BD differ in clinical expression and familial risk, and may therefore be considered different subforms of manic-depressive illness (Schurhoff et al., 2000). However, the correlation between GAOD and age of onset in this study may reflect genetic differences between early and late onset BD; indeed, there was only one case of BD with an age of onset over 40. To clarify the influence of age of onset on GAOD, more cases of BD with onset under 20 and over 40 years should be investigated in the future.

4.5. Decreased AAOC and clinical variables in MDD and BD

The AAOC tended to be decreased in the bilateral FPC and the left DLPFC in both MDD and BD groups, compared to the NC group.

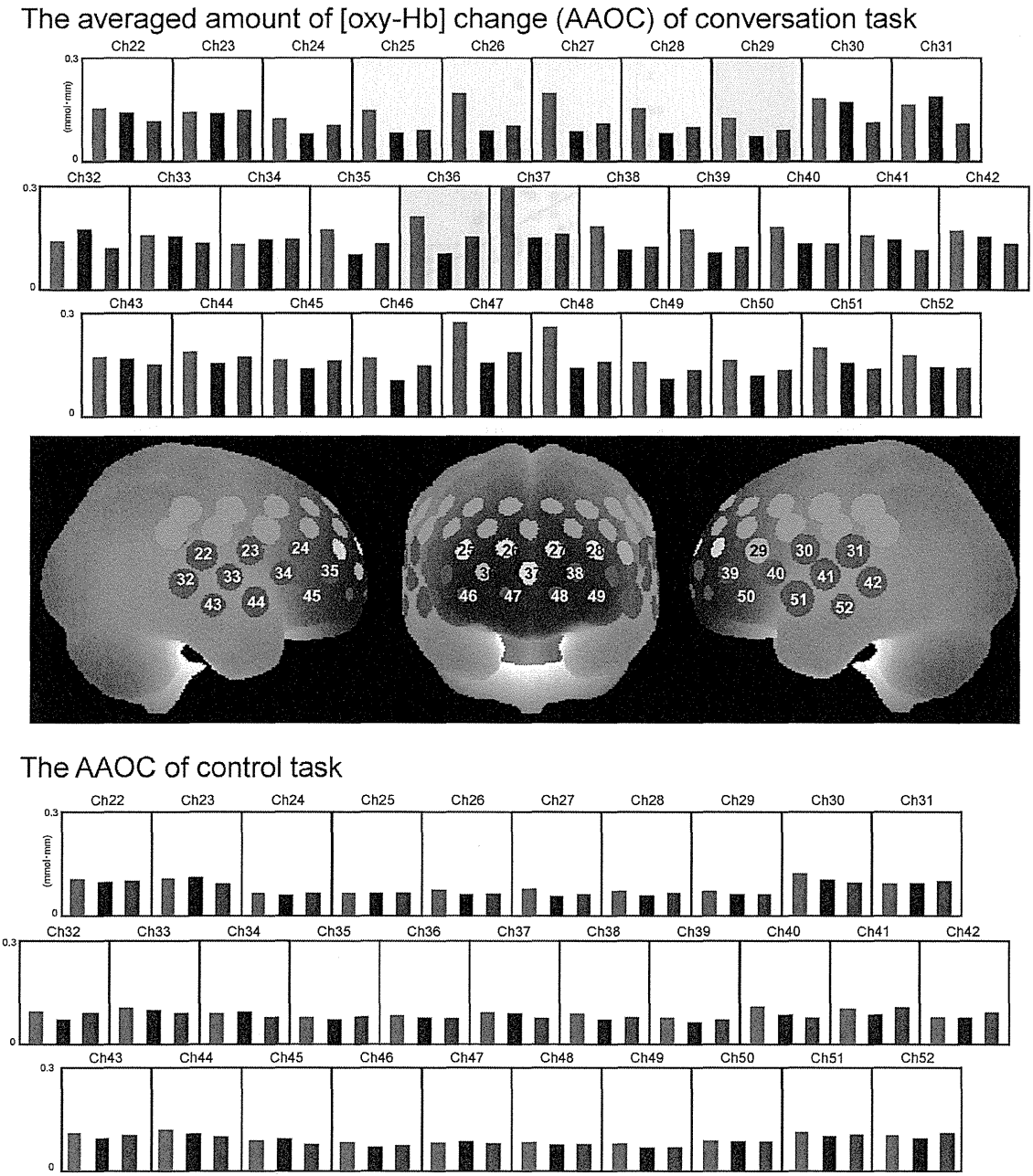

Fig. 5. The upper 31 figures (Ch22–Ch52) indicate the averaged amount of [oxy-Hb] change (AAOC) of the conversation task: green bar, control participant; blue bar, major depression; red bar, bipolar disorder. The yellow channels of the upper figures indicate a larger AAOC in the NC group than in the MDD and BD groups, and the light blue channels indicate a larger AAOC in the NC group than in the MDD group by mixed-design repeated measures ANOVA, followed by Scheffé's post hoc *t*-test. The middle 3 figures below show the probabilistic estimation and anatomical labeling of the locations of NIRS channels in the standard brain space in accordance with Tsuzuki et al. (2007), and the yellow and light blue areas indicate the corresponding brain areas that differed between the groups according to the results of the post hoc *t*-test. Light gray channels without a number are channels that were excluded following the detection of clear motion artifacts. The lower 31 figures (Ch22–Ch52) indicate the AAOC of the control task: green bar, control participant; blue bar, major depression; red bar, bipolar disorder.

However, the decreased area of AAOC for MDD was wider than for BD, contrary to our hypothesis.

The main difference between MDD and BD was the decreased AAOC in the right FPC found in only the MDD group. The right frontal lobe is involved in generating unusual or distant verbal associations while the left frontal lobe is involved in generating "usual" associations (Kiefer et al., 1998; Seger et al., 2000). Clinical research suggests that individuals with MDD are cognitively inflexible, exhibiting ruminative, rigid, and automatic thoughts within a negative schema (Deveney and Deldin, 2006; Remijnse et al., 2013). Researchers reported that individuals with MDD were less flexible than to healthy controls. These previous studies may indicate that decreased AAOC in the right FPC reflected the cognitive inflexibility seen in MDD.

Among the patients with MDD, the AAOC tended to be positively correlated with the GAF in the right FPC. Pu et al. (2012) reported that reduced activation in the prefrontal and temporal regions during a working memory task was significantly related to lower

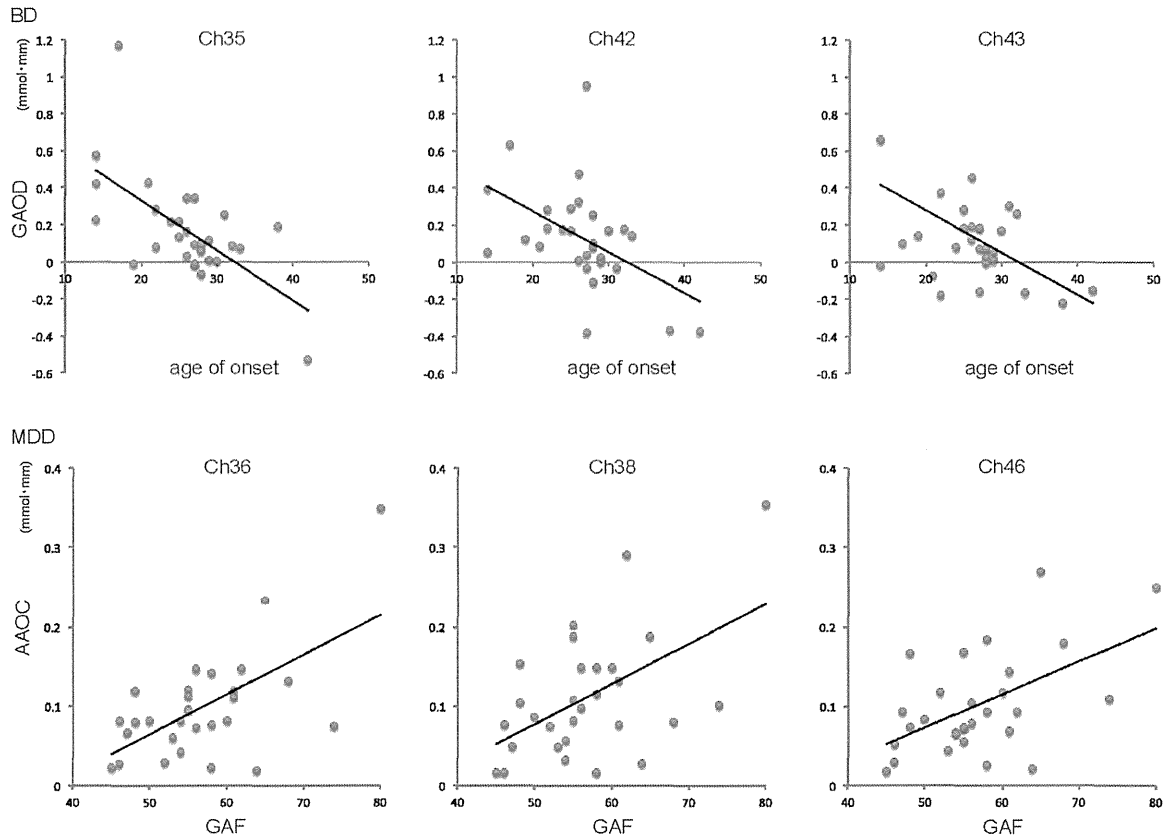

Fig. 6. Upper 3 figures (Ch35, Ch42, and Ch43): Significant negative correlation between the grand averaged [oxy-Hb] data (GAOD) during the conversation task of BD and age of onset. Lower 3 figures (Ch36, Ch38, and Ch46): Significant positive correlation between the AAOC data during the conversation task for MDD and GAF.

scores on the Social Adaptation Self-Evaluation Scale (SASS) in late-onset depression. Our results indicate that decreased AAOC in the right FPC may be reflected in the cognitive inflexibility of MDD, especially during face-to-face conversations. This may also be related to social adaptation through communication with others.

4.6. Limitations and future directions

The limitations of this study are as follows: (i) the results were mainly drawn from patients in the euthymic state, since few patients in depressive or hypomanic states were included; (ii) we were unable to investigate the correlation between NIRS measurements and psychotropic medication because we did not include drug-free patients, or patients taking a single drug; (iii) we did not evaluate neurocognitive function and intelligence quotient by other neuropsychological methods; therefore, we could not evaluate relationships between conversation performance and cognitive domains such as executive function or working memory; (iv) we could not identify an observable index of conversation that is more sensitive to the GAOD and AAOC. Future studies addressing these limitations are planned.

4.7. Conclusion

We used NIRS to investigate frontal lobe activation in MDD, BD, and NC participants during face-to-face conversations in situ, and demonstrated decreased activation and different temporal characteristics among the three groups. GAOD, which may reflect conversation situation-related function, was decreased in the left DLPFC and FPC in both MDD and BD groups. AAOC, which may reflect speech-related function, was decreased in the FPC in both the MDD and BD groups; this was shown more strongly in the right FPC for the MDD group. GAOD was negatively correlated with age of onset in BD, while AAOC was positively correlated with GAF in MDD. However, both continuous activation and rapid change may reflect the pathophysiological character of MDD and BD; in particular, the decrease of AAOC in the right FPC may be related to the impaired adaptive ability exhibited in MDD.

Conflict of interest

Authors have no conflicts of interest to declare.

Contributors

Masashi Suda and Yuichi Takei designed the tasks; Masashi Suda, Yuichi Takei, Yoshiyuki Aoyama, Kosuke Narita, Miho Yamaguchi, Minami Tagawa, Tomokazu Motegi, and Noriko Sakurai conducted the experiments and analyzed the data; and Yuichi Takei, Masashi Suda, and Masato Fukuda, wrote the final version of the manuscript.

Role of the funding source

This work was supported in part by grants awarded to TY from the Ministry of Education, Culture, Sports, Science, and Technology (Grant-in-Aid for Scientific Research on Innovative Areas [4301]; the Japan Society for the Promotion of Science (Grants-in-Aid for

Scientific Research [B] [No. 12014473]) and MF from the Ministry of Education, Culture, Sports, Science and Technology (Grant-in-Aid for Scientific Research on Innovative Areas (23118001 & 23118003; Adolescent Mind & Self-Regulation); the Japan Society for the Promotion of Science (Grants-in-Aid for Scientific Research [B] [No. 23390286] and for Challenging Exploratory Research [No. 22659209]; the Ministry of Health, Labour and Welfare (Health and Labour Sciences Research Grants, Comprehensive Research on Disability, Health and Welfare, No. H25-Seishin-Ippan-002); the National Center for Neurology and Psychiatry (Intramural Research Grant for Neurological and Psychiatric Disorders, No. 21-1 and 23-10); Japanese Ministry of Health, Labour and Welfare, Health and Labour Science Research Grant on the Practical Application of Medical Technology for Intractable Diseases and Cancer: New development of medical technology for the diagnosis and treatment of psychiatric diseases and cancer by construction of virtual mega-hospital for clinical trials (H25-Jitsuyouka (kokusai) Shitei-002); and the SENSHIN Medical Research Foundation.

Acknowledgments

The authors thank the trainee doctors (Dr. Kameyama, Dr Narita, Dr. Majima, and Dr. Yonemura, Dr Mikuni) of Gunma University Hospital and the students of the Gunma University Faculty of Medicine for participating in the study. Gunma University (Dr. Fukuda) and the Hitachi Group (Advanced Research Laboratory, Hitachi Ltd., and the Research and Developmental Center, Hitachi Medical Corporation) have maintained an official contract with Gunma University Hospital for a collaborative study of clinical applications of NIRS in psychiatric disorders since 2002.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jpsychires.2014.06.009.

References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Washington, DC: American Psychiatric Publishing; 1994.
Arnone D, Cavanagh J, Gerber D, Lawrie SM, Ebmeier KP, McIntosh AM. Magnetic resonance imaging studies in bipolar disorder and schizophrenia: meta-analysis. Brit J Psychiatry 2009;195:194–201.
Austin MP, Ross M, Murray C, O'Carroll RE, Ebmeier KP, Goodwin GM. Cognitive function in major depression. J Affect Disord 1992;25:21–9.
Bouhuys AL, Sam MM. Lack of coordination of nonverbal behaviour between patients and interviewers as a potential risk factor to depression recurrence: vulnerability accumulation in depression. J Affect Disord 2000;57:189–200.
Celikel FC, Kose S, Cumurcu BE, Erkorkmaz U, Sayar K, Borckardt JJ, et al. Cloninger's temperament and character dimensions of personality in patients with major depressive disorder. Compr Psychiatry 2009;50:556–61.
Costantini M, Di Vacri A, Chiarelli AM, Ferri F, Romani GL, Merla A. Studying social cognition using near-infrared spectroscopy: the case of social Simon effect. J Biomed Opt 2013;18:250–5.
Coyne JC, Downey G. Social factors and psychopathology: stress, social support, and coping processes. Annu Rev Psychol 1991;42:401–25.
Cui X, Bryant DM, Reiss AL. NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. Neuroimage 2012;59:2430–7.
Deveney CM, Deldin PJ. A preliminary investigation of cognitive flexibility for emotional information in major depressive disorder and non-psychiatric controls. Emotion 2006;6:429–37.
Drevets WC, Bogers W, Raichle ME. Functional anatomical correlates of antidepressant drug treatment assessed using PET measures of regional glucose metabolism. Eur Neuropsychopharm 2002;12:527–44.
Frangou S, Kington J, Raymont V, Shergill SS. Examining ventral and dorsal prefrontal function in bipolar disorder: a functional magnetic resonance imaging study. Eur Psychiatry 2008;23:300–8.
Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62.
Hansenne M, Reggers J, Pinto E, Kjiri K, Ajamier A, Ansseau M. Temperament and character inventory (TCI) and depression. J Psychiatr Res 1999;33:31–6.
Inagaki A, Inada T. Dose equivalence of psychotropic drugs. Jpn J Psychopharm 2006;9:1443–7.
Kameyama M, Fukuda M, Yamagishi Y, Sato T, Uehara T, Ito M, et al. Frontal lobe function in bipolar disorder: a multichannel near-infrared spectroscopy study. Neuroimage 2006;29:172–84.
Kennedy SH, Evans KR, Kruger S, Mayberg HS, Meyer JH, McCann S, et al. Changes in regional brain glucose metabolism measured with positron emission tomography after paroxetine treatment of major depression. Am J Psychiatry 2001;158:899–905.
Kiefer M, Weisbrod M, Kern I, Maier S, Spitzer M. Right hemisphere activation during indirect semantic priming: evidence from event-related potentials. Brain Lang 1998;64:377–408.
Koizumi H, Yamashita Y, Maki A, Yamamoto T, Ito Y, Itagaki H, et al. Higher-order brain function analysis by trans-cranial dynamic near-infrared spectroscopy imaging. J Biomed Opt 1999;4:403–13.
Konvalinka I, Roepstorff A. The two-brain approach: how can mutually interacting brains teach us something about social interaction? Front Hum Neurosci 2012;6:215.
Koutsouleris N, Gaser C, Jager M, Bottlender R, Frodl T, Holzinger S, et al. Structural correlates of psychopathological symptom dimensions in schizophrenia: a voxel-based morphometric study. Neuroimage 2008;39:1600–12.
Kronhaus DM, Lawrence NS, Williams AM, Frangou S, Brammer MJ, Williams SC, et al. Stroop performance in bipolar disorder: further evidence for abnormalities in the ventral prefrontal cortex. Bipolar Disord 2006;8:28–39.
Ohta H, Yamagata B, Tomioka H, Takahashi T, Yano M, Nakagome K, et al. Hypofrontality in panic disorder and major depressive disorder assessed by multichannel near-infrared spectroscopy. Depress Anxiety 2008;25:1053–9.
Okamoto M, Dan H, Sakamoto K, Takeo K, Shimizu K, Kohno S, et al. Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping. Neuroimage 2004;21:99–111.
Pu S, Yamada T, Yokoyama K, Matsumura H, Mitani H, Adachi A, 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;203:222–8.
Remijnse PL, van den Heuvel OA, Nielen MMA, Vriend C, Hendriks GJ, Hoogendijk WJG, et al. Cognitive inflexibility in obsessive-compulsive disorder and major depression is associated with distinct neural correlates. PLOS ONE 2013;8.
Roth RM, Koven NS, Randolph JJ, Flashman LA, Pixley HS, Ricketts SM, et al. Functional magnetic resonance imaging of executive control in bipolar disorder. Neuroreport 2006;17:1085–9.
Sacher J, Neumann J, Funfstuck T, Soliman A, Villringer A, Schroeter ML. Mapping the depressed brain: a meta-analysis of structural and functional alterations in major depressive disorder. J Affect Disord 2012;140:142–8.
Savitz J, Drevets WC. Bipolar and major depressive disorder: neuroimaging the developmental-degenerative divide. Neurosci Biobehav Rev 2009;33:699–771.
Schurhoff F, Bellivier F, Jouvent R, Mouren-Simeoni MC, Bouvard M, Allilaire JF, et al. Early and late onset bipolar disorders: two different forms of manic-depressive illness? J Affect Disord 2000;58:215–21.
Seger CA, Desmond JE, Glover GH, Gabrieli JD. Functional magnetic resonance imaging evidence for right-hemisphere involvement in processing unusual semantic relationships. Neuropsychology 2000;14:361–9.
Strakowski SM, Adler CM, Holland SK, Mills NP, DelBello MP, Eliassen JC. Abnormal FMRI brain activation in euthymic bipolar disorder patients during a counting Stroop interference task. Am J Psychiatry 2005;162:1697–705.
Strangman G, Boas DA, Sutton JP. Non-invasive neuroimaging using near-infrared light. Biol Psychiatry 2002;52:679–93.
Suda M, Takei Y, Aoyama Y, Narita K, Sato T, Fukuda M, et al. Frontopolar activation during face-to-face conversation: an in situ study using near-infrared spectroscopy. Neuropsychologia 2010;48:441–7.
Suda M, Takei Y, Aoyama Y, Narita K, Sakurai N, Fukuda M, et al. Autistic traits and brain activation during face-to-face conversations in typically developed adults. PLOS ONE 2011;6.
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;55:501–11.
Takei Y, Suda M, Aoyama Y, Yamaguchi M, Sakurai N, Narita K, et al. Temporal lobe and inferior frontal gyrus dysfunction in patients with schizophrenia during face-to-face conversation: a near-infrared spectroscopy study. J Psychiatr Res 2013;47:1581–9.
Tham A, Engelbrektson K, Mathe AA, Johnson L, Olsson E, Aberg-Wistedt A. Impaired neuropsychological performance in euthymic patients with recurring mood disorders. J Clin Psychiatry 1997;58:26–9.
Tsuzuki D, Jurcak V, Singh AK, Okamoto M, Watanabe E, Dan I. Virtual spatial registration of stand-alone fNIRS data to MNI space. Neuroimage 2007;34:1506–18.
Veiel HO. A preliminary profile of neuropsychological deficits associated with major depression. J Clin Exp Neuropsychol 1997;19:587–603.
Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. Brit J Psychiatry 1978;133:429–35.
Zakzanis KK, Leach L, Kaplan E. On the nature and pattern of neurocognitive function in major depressive disorder. Neuropsychiatry Neuropsychol Behav Neurol 1998;11:111–9.

NeuroImage 85 (2014) 498–507

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Neuroimaging-aided differential diagnosis of the depressive state

Ryu Takizawa [a,1], Masato Fukuda [b,*,1], Shingo Kawasaki [c], Kiyoto Kasai [a], Masaru Mimura [d,e], Shenghong Pu [f], Takamasa Noda [g], Shin-ichi Niwa [h], Yuji Okazaki [i,j]
on behalf of the Joint Project for Psychiatric Application of Near-Infrared Spectroscopy (JPSY-NIRS) Group

[a] Department of Neuropsychiatry, The University of Tokyo Graduate School of Medicine, Tokyo 113-8655, Japan
[b] Department of Psychiatry and Neuroscience, Gunma University Graduate School of Medicine, Gunma 371-8511, Japan
[c] Hitachi Medical Corporation, Tokyo 101-0021, Japan
[d] Department of Neuropsychiatry, Showa University School of Medicine, Tokyo 157-8577, Japan
[e] Department of Psychiatry, Keio University School of Medicine, Tokyo 160-8582, Japan
[f] Division of Neuropsychiatry, Department of Multidisciplinary Internal Medicine, Tottori University Faculty of Medicine, Tottori 683-8504, Japan
[g] National Centre of Neurology and Psychiatry Hospital, Tokyo 187-8551, Japan
[h] Department of Neuropsychiatry, Fukushima Medical University School of Medicine, Fukushima 960-1295, Japan
[i] Department of Psychiatry, Mie University Graduate School of Medicine, Mie 514-8507, Japan
[j] Tokyo Metropolitan Matsuzawa Hospital, Tokyo 156-0057, Japan

ARTICLE INFO

Article history:
Accepted 31 May 2013
Available online 10 June 2013

Keywords:
Neuroimaging
Near-infrared spectroscopy (NIRS)
Differential diagnosis
Depressive state
Psychiatric disorder

ABSTRACT

A serious problem in psychiatric practice is the lack of specific, objective biomarker-based assessments to guide diagnosis and treatment. The use of such biomarkers could assist clinicians in establishing differential diagnosis, which may improve specific individualised treatment. This multi-site study sought to develop a clinically suitable neuroimaging-guided diagnostic support system for differential diagnosis at the single-subject level among multiple psychiatric disorders with depressive symptoms using near-infrared spectroscopy, which is a compact and portable neuroimaging method. We conducted a multi-site, case–control replication study using two cohorts, which included seven hospitals in Japan. The study included 673 patients (women/men: 315/358) with psychiatric disorders (major depressive disorder, bipolar disorder, or schizophrenia) who manifested depressive symptoms, and 1007 healthy volunteers (530/477). We measured the accuracy of the single-subject classification in differential diagnosis among major psychiatric disorders, based on spatiotemporal characteristics of fronto-temporal cortical haemodynamic response patterns induced by a brief (<3 min) verbal fluency task. Data from the initial site were used to determine an optimal threshold, based on receiver-operator characteristics analysis, and to generate the simplest and most significant algorithm, which was validated using data from the remaining six sites. The frontal haemodynamic patterns detected by the near-infrared spectroscopy method accurately distinguished between patients with major depressive disorder (74.6%) and those with the two other disorders (85.5%; bipolar disorder or schizophrenia) that presented with depressive symptoms. These results suggest that neuroimaging-guided differential diagnosis of major psychiatric disorders developed using the near-infrared spectroscopy method can be a promising biomarker that should aid in personalised care in real clinical settings. Potential confounding effects of clinical (e.g., age, sex) and systemic (e.g., autonomic nervous system indices) variables on brain signals will need to be clarified to improve classification accuracy.
© 2013 Elsevier Inc. All rights reserved.

Introduction

Among non-communicable diseases, neuropsychiatric conditions, including depression, contribute most significantly to overall disability-adjusted life years (DALYs), surpassing both cardiovascular disease and cancer (Mathers and Loncar, 2006; Prince et al., 2007). Therefore, early and accurate diagnosis and treatment are critical in psychiatric disorders, for which the development of specific biomarkers is of special importance. Currently, however, the diagnostic process in psychiatry is mainly based on patients' reports of symptoms, observed behaviours and disease course. Overcoming the limitations of relying on clinical interviews alone for the diagnosis of psychiatric disorders has been a great challenge.

To complicate this issue further, the manifestation of only a major depressive episode hampers the reliable differentiation of major depressive disorder (MDD) from bipolar disorder (BP) or

* Corresponding author at: Department of Psychiatry and Neuroscience, Gunma University Graduate School of Medicine, 3-39-22 Showa-machi, Maebashi, Gunma 371-8511, Japan. Fax: +81 27 220 8187.
E-mail address: fkdpsy@med.gunma-u.ac.jp (M. Fukuda).
[1] These authors contributed equally to this paper.

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.05.126

schizophrenia (SZ) based on the Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria alone (Zimmermann et al., 2009). Although many clinical symptoms are common to various psychiatric disorders, depressive symptoms are particularly ubiquitous in the disease process or clinical staging of various psychiatric disorders (Hafner et al., 2005). For instance, differentiation between BP presenting with depressive symptoms and unipolar MDD is a topical issue (Akiskal et al., 1995). Indeed, most patients with BP with depressive symptoms are initially diagnosed with and treated for MDD (Akiskal et al., 1995; Goldberg et al., 2001). Therefore, biomarkers that can facilitate early and accurate differentiation of BP with depressive symptoms from MDD are necessary.

In addition, depressive symptoms that fulfil the operational diagnostic criteria for a depressive episode/major depression can also occur at any stage of SZ and can contribute substantially to its associated morbidity and even mortality (an der Heiden et al., 2005). The differentiation of SZ from MDD, especially in the early stages, is also important because patients with SZ also exhibit non-psychotic and non-specific prodromal symptoms (e.g., depressive or negative symptoms and cognitive deficits) for several years before the onset of full-blown psychosis (McGorry et al., 2008). Therefore, the availability of clinically useful and cost-effective biomarkers for the differential diagnosis of major psychiatric disorders would likely enhance patient management, improve treatment/therapeutic response and lead to targeted therapies tailored to the individual (Holsboer, 2008). Despite their potential, to date, no such biomarkers have been established.

Functional imaging studies are one source of potential biomarkers (Gur et al., 2007; Phillips and Vieta, 2007); these studies have previously elucidated subtle brain abnormalities in patients with major psychiatric disorders relative to healthy control (HC) individuals and have been applied to the differential diagnosis of psychiatric disorders (e.g., to differentiate MDD from SZ, Barch et al., 2003 or BP, Almeida et al., 2009). However, some functional neuroimaging techniques are limited by the fact that, during the procedure, the individuals need to be placed in an uncomfortable or unnatural setting (e.g., lying in a supine position in a narrow gantry with the head fixed during the entire examination), for accurate measurement.

In contrast, multi-channel near-infrared spectroscopy (NIRS) using near-infrared light provides a completely non-invasive measurement of the spatiotemporal characteristics of brain function in ordinary clinical settings and allows patients to be comfortably seated in a well-lit room; therefore, it is considered a method for 'real-world neuroimaging'. Additionally, NIRS has relatively low maintenance costs and does not involve ionising radiation or objectionable noise; thus, it can be repeated as needed even for patients with psychiatric disorders. The utility and limitations of NIRS have been discussed extensively in previous reports (Ferrari and Quaresima, 2012; Obrig and Villringer, 2003; Strangman et al., 2002a). NIRS allows the measurement of haemoglobin concentration changes (1) only in the cortical surface area located immediately beneath the probes, but not in deeper brain structures, and (2) with limited spatial resolution, although it has a high temporal resolution. In NIRS, typical cortical activation represents not only decreased concentration of deoxy-haemoglobin ([deoxy-Hb]), which is considered the main source of blood oxygenation level-dependent (BOLD) contrast increase in functional magnetic resonance imaging (fMRI), but also a relatively larger increase in oxy-haemoglobin concentration ([oxy-Hb]) (Fig. 1).

The verbal fluency task (VFT) is a cognitive task that is used as a neuropsychological test or a neuroimaging task. The VFT elicits different abnormalities relevant to each diagnostic group of major psychiatric disorders (Curtis et al., 2001; Zanelli et al., 2010). We previously developed a very brief (<3 min) VFT and used it to investigate the differential fronto-temporal haemodynamic pattern between MDD and SZ (Suto et al., 2004) or MDD and BP (Kameyama et al., 2006), as well as the relationship between NIRS signals and functional impairment in SZ (Takizawa et al., 2008). We also found functional NIRS abnormalities in individuals at ultra-high risk for SZ and patients with first-episode psychosis (Koike et al., 2011). However, the clinical applicability of NIRS to the differential diagnosis of individuals remains uncertain. In this study, we extended our translational approach to replicate our previous findings (Kameyama et al., 2006; Suto et al., 2004) in a seven-site collaborative study using a large, fully independent sample set, and to evaluate the application of NIRS to psychiatric differential diagnosis in natural clinical settings.

Specifically, we used NIRS with wide coverage of the prefrontal and temporal cortices to investigate whether the frontal and temporal brain haemodynamic responses induced by cognitive activation could serve as biomarkers of underlying major psychiatric disorders with depression. To validate the reproducibility and generalisability of the results, we applied an algorithm developed using the data generated at the initial site to the test data derived from the remaining 6 sites. We hypothesised that the spatiotemporal characteristics of the haemodynamic responses detected by NIRS would not only differentiate patients with psychiatric disorders from HCs with acceptable sensitivity and specificity, but would also differentiate correctly and with a high concordance rate patients with MDD from patients with bipolar disorder and schizophrenia who present with depressive symptoms.

Material and methods

Participants

This multi-site study was performed in 7 hospitals: 6 were affiliated with universities (Fukushima, Gunma, Mie, Tokyo, Showa, and Tottori) and one was affiliated with the National Centre of Neurology and Psychiatry of Japan. The sites were situated in the Tokyo metropolitan area and in moderate-scale prefectural capital cities (Fukushima, Maebashi, Tsu and Yonago). The participants were recruited from June 2004 to June 2009, with the exception of recruitment at the initial site (Gunma University Hospital in Maebashi City), which was conducted over 6 years (March 2003 to March 2009). The ethics committees of the participating hospitals approved this collaborative study. In accordance with the Declaration of Helsinki, all participants gave written informed consent after receiving a complete explanation of the study.

Six hundred and seventy-three in-patients and out-patients with psychiatric disorders (MDD, BP and SZ), in addition to 1007 HC volunteers (Flow diagram (1)), were initially enrolled. Of note, these individuals were not the same as those included in our previous studies (Kameyama et al., 2006; Suto et al., 2004). The patients were diagnosed by experienced psychiatrists based on the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID) (First et al., 1997). The HC volunteers were hospital staff members, university students and members of the general population who responded to website or newspaper advertisements in each city. The SCID non-patient edition was used to screen HC individuals. The exclusion criteria of the initial enrolment were neurological illness, traumatic brain injury with any known cognitive consequences and alcohol/substance abuse or addiction. All participants were native Japanese speakers who were capable of performing a Japanese version of the VFT easily.

On the day of NIRS measurement, the depressive symptoms of participants were evaluated using the 17-item Hamilton Rating Scale for Depression (HAMD) (Hamilton, 1960) and their psychotic and manic symptoms were evaluated using both the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1991) and the Young Mania Rating Scale (YMRS) (Young et al., 1978), respectively, by well-trained psychiatrists with no knowledge of the NIRS data. During the study, all patients with psychiatric disorders were medicated with one or more agents (anti-psychotics, anti-depressants, anxiolytics and/or anti-parkinsonian agents), with the exception of 10 drug-free patients with MDD and 5 drug-free patients with SZ.

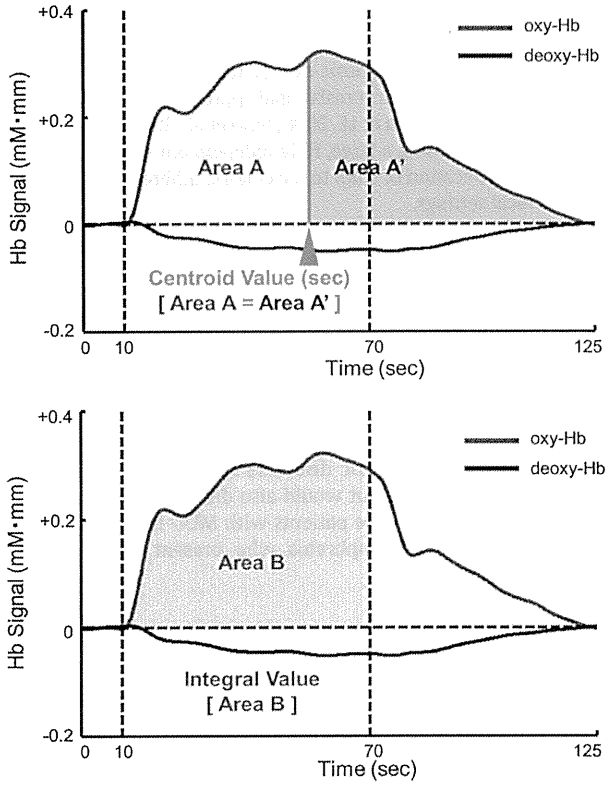

Fig. 1. Typical time-course pattern of near-infrared spectroscopy (NIRS) signals coupled with the verbal fluency task. The 'centroid value' is indicated by the time [s], which is indicated by a perpendicular line from the centroid of an NIRS signal-change area (calculated with positive change) throughout all the task periods. Oxy-Hb: oxygenated haemoglobin signal; deoxy-Hb: deoxygenated haemoglobin signal.

To minimise the influence of confounding factors, we performed group matching for age and gender among the 4 diagnostic groups using one-way analysis of variance (ANOVA) and a chi-squared test, which excluded randomly selected individuals and brought the mean age of the HC individuals and patients with MDD or SZ in closer alignment with that of the patients with BP (44.0 ± 14.9 years old (y.o.)), which was the group with the fewest individuals (Table 1 and Flow diagram (2)). For confirmation, we also analysed demographically non-matched samples that are identified in the Supplementary Material (I). The overall results were the same as those described in the main manuscript and the reduction in the total number of study participants after demographical matching did not appear to have an influence on the development of the algorithm (see Supplementary Material (I)).

Because our clinically valuable target were help-seeking unremitted patients, subsequently we excluded study participants with extremely mild symptoms (HAMD score ≤ 5, PANSS depression item score ≤ 1, PANSS negative symptom score ≤ 11, PANSS general psychopathology score ≤ 21, or PANSS positive symptom score − negative symptom score ≤ 11; the latter 3 criteria were based on the criteria from the PANSS manual for the 5th percentile of patients with mild SZ, Kay et al., 1991). We also excluded patients in a manic phase (YMRS score > 10) from the NIRS measurement; rather, we focused on patients with BP who were in the depressive phase because the different phases may produce different brain dysfunctions in patients with BP (Phillips and Vieta, 2007), and manic patients with BP were diagnosed without apparent difficulty (Flow diagram (3)).

Activation task

The activation task used in this study was similar to that used in our previous studies (Kameyama et al., 2006; Suto et al., 2004; Takizawa et al., 2008). Briefly, a VFT (letter version) was administered and NIRS signal ([oxy-Hb] and [deoxy-Hb]) changes were measured during a 10 s pre-task baseline period, a 60 s activation period and a 55 s post-task baseline period. During the activation period, the participants were instructed to utter as many Japanese words beginning with a designated syllable as possible. For the pre- and post-task baseline periods, the individuals were instructed to simply repeat Japanese vowels out loud. The total number of correct words generated during the 60 s activation period was used as the measure of task performance (Table 1).

Among the many neuropsychological tasks used for detecting neurocognitive deficits in patients with major psychiatric disorders (Barrett et al., 2009; Zanelli et al., 2010), we selected the VFT because it is an executive task that exhibits distinct differences in performance and neuroimaging data among each diagnostic group of major psychiatric disorders (Costafreda et al., 2006; Curtis et al., 2001; Zanelli et al., 2010). In addition, the VFT is easy to understand and execute; in fact, all participants generated more than one word during the VFT. Therefore, this task is suitable for translational research aimed at identifying practical biomarkers.

NIRS measurement

The NIRS apparatus and measurement procedure were described in full previously (Takizawa et al., 2008). Briefly, we used a 52-channel NIRS system (ETG-4000; Hitachi Medical Co., Tokyo, Japan). The preparation of the apparatus, including the audiovisual on-screen instructions, usually took less than 7 min and our brief version of the VFT took less than 3 min, which is less demanding for participants (10–15 min is necessary for the entire procedure).

NIRS is based on the principle that near-infrared light is preferentially absorbed by haemoglobin and less so by other tissues. Near-infrared light emitted from the skin travels into the body, is reflected and absorbed by the internal tissues and reappears on the skin. Thus, the absorption of near-infrared light reflects haemoglobin concentration ([Hb]) in the tissue placed beneath emission and detection probe pairs. Measurements taken using 2 or more wavelengths of near-infrared light enable the determination of [oxy-Hb] and [deoxy-Hb] changes because their absorptions are different at different wavelengths. The ETG-4000 system measures relative changes in [oxy-Hb] and [deoxy-Hb] using 2 wavelengths (695 and 830 nm) of infrared light, based on the modified Beer–Lambert law.(Yamashita et al., 1996) In this continuous-wave NIRS system, these [Hb] values include a differential pathlength factor (DPF); therefore, the unit of this form of NIRS measurement is mM·mm. The distance between pairs of source-detector probes was set to 3.0 cm and each measurement area located between pairs of source-detector probes was defined as one 'channel'. It is assumed that a machine in which the source-detector spacing is 3.0 cm measures points at a depth of 2–3 cm from the scalp (i.e., the surface of the cerebral cortex) (Okada and Delpy, 2003). The temporal resolution of NIRS was set to 0.1 s.

The arrangement of the probes measured relative [oxy-Hb] and [deoxy-Hb] signal changes in the bilateral prefrontal cortical area (i.e., dorso-lateral [Brodmann areas (BAs) 9 and 46], ventro-lateral [BAs 44, 45, and 47] and fronto-polar [BA 10] regions) and in the superior and middle temporal cortical surface regions, which was corroborated by a multi-individual study of anatomical cranio-cerebral correction via the international 10–20 system (Fig. 2 and Table S1) (Tsuzuki et al., 2007). However, in the 10–20 system, the anterior parts of the probes (e.g., Fpz) can be positioned precisely, whereas the position errors of more lateral probes might be increased due to inter-individual head size variability. In addition, although we initially aimed to analyse single-individual and single-channel levels in this study, studies of repeated NIRS measures have demonstrated acceptable reliability of the NIRS signal at the group and cluster levels, whereas retest reliability was unsatisfactory at the single-individual and single-channel levels (Schecklmann et al., 2008).

Table 1
Demographic and clinical characteristics of the 4 age- and gender-matched diagnostic groups at all 7 study sites.

	Major depressive disorder		Schizophrenia		Bipolar disorder		Healthy control		Group difference
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	*p*-Value
n	153		136		134		590		
Age years	43.8	12.7	43.7	12.1	44.0	14.9	43.9	15.7	0.99
Gender, women/men	77/76		67/69		69/65		314/276		0.81[a]
Education, years	14.0	1.9	15.2	2.0	15.6	2.0	16.1	2.4	<0.01
Estimated premorbid IQ	106.0	10.1	103.7	11.2	106.9	8.6	107.2	10.1	0.23
Task performance	13.0	3.8	13.6	4.4	12.0	3.6	15.3	4.8	<0.01
Age at onset, years	39.2	11.3	23.4	7.4	32.9	12.4	–		
PANSS									
Positive	–		16.3	5.0	–		–		
Negative	–		21.0	6.0	–		–		
General psychopathology	–		37.0	7.6	–		–		
HAM-D	14.1	6.7	–		8.4	7.0	–		
YMRS	–		–		4.7	5.9	–		
GAF	53.9	9.7	47.3	11.4	55.5	13.3	–		

Abbreviations: IQ, intelligence quotient; PANSS, Positive and Negative Syndrome Scale; GAF, global assessment of functioning.
[a] Chi-square test was used for testing group difference. Otherwise, one-way ANOVA was used.

Therefore, instead of undertaking a full analysis at the single-individual and single-channel levels, here we performed an analysis of NIRS signals at the single-individual and cluster levels. A principal component analysis (PCA) of NIRS [oxy-Hb] signal changes in targeted fronto-temporal channels was performed at the initial study site as a preliminary analysis to capture a channel cluster of the analogous time-course pattern in HC individuals. Subsequently, the weight maps of the first and second principal component graphs were used to identify 2 cluster components.

These analyses suggested that 2 cluster components were identified and that the 2 clusters included the frontal region (11 channels) and the bilateral temporal region (10 channels each) (see Supplementary Material (II) and eFig. S1). The channels in these 2 respective regions of interest were averaged and transformed into representative 'Region 1 (R1)' and 'Region 2 (R2)' NIRS signals for each individual (Fig. 3). According to registration into the LONI Probabilistic Brain Atlas 40 (LBPA40) (Fig. 2) (see Supplementary Table S1 for LBPA40 anatomical labels) (Shattuck et al., 2008), the R1 NIRS signal consisted of signals from channels located approximately in the fronto-polar and dorsolateral prefrontal cortical regions (i.e., superior and middle frontal gyri), whereas the R2 NIRS signal consisted of signals from channels located approximately in the ventro-lateral prefrontal cortex and the superior and middle temporal cortical regions (i.e., inferior frontal gyrus and superior and middle temporal gyri).

An automatic artefact-rejection procedure (see Supplementary Material (III)) was followed and individual data were excluded when there were fewer than 6 remaining channels from each of the 2 cluster regions (Flow diagram (4)).

Statistical analyses

Taking into consideration the potential application of the technique in ordinary clinical settings and personalised care, a conservative receiver operating characteristic (ROC) analysis was performed and used to generate simple indices of NIRS signal patterns, to aid individual diagnoses.

The spatiotemporal characteristics of the frontal and temporal haemodynamic responses induced by VFT were assessed and subsequently applied to an algorithm using the simplest and fewest variables for differential diagnosis. Because previous studies have shown that the best way to differentiate patients with MDD from those with BP or SZ is to describe the time-course of changes in the NIRS signal associated with the VFT (Kameyama et al., 2006; Suto et al., 2004), we chose to create 2 simple visual indices, referred to here as 'integral value' and 'centroid value', to capture these time-course changes.

The integral value describes the size of the haemodynamic response during the 60 s activation task period, whereas the centroid value serves as an index of time-course changes throughout the task, with periods representing the timing of the haemodynamic response. The centroid value is indicated by a time shown with a perpendicular line from the centroid of the NIRS signal change area (calculated as a positive change) throughout the task periods (from 0 [s] to 125 [s]

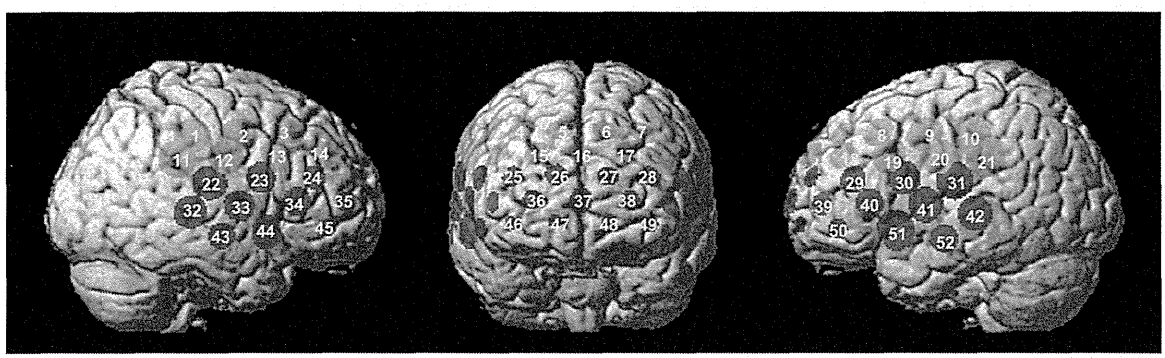

Fig. 2. Regions of interest (Regions 1 and 2) of the near-infrared spectroscopy (NIRS) signals. The locations of near-infrared spectroscopy (NIRS) measurements were probabilistically estimated and anatomically labelled in the standard brain space (LBPA40) according to Tsuzuki et al. (2007). Region 1: (ch 25–28, ch 36–38 and ch 46–49); Region 2, Right: (ch 22–24, ch 32–35 and ch 43–45); Left: (ch 29–31, ch 39–42 and ch 50–52).

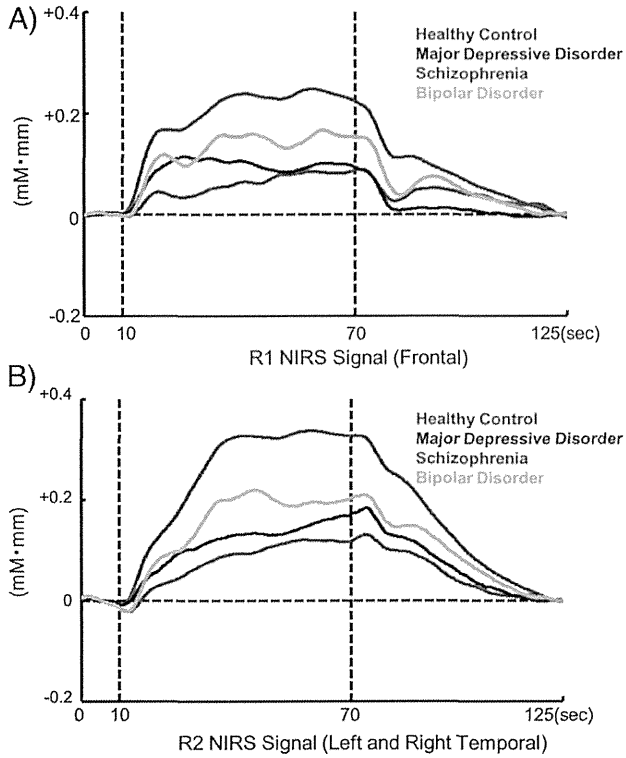

Fig. 3. Time courses of the haemodynamic responses in Region 1 (R1) and Region 2 (R2) in the 4 diagnostic groups. Panels A and B show the time courses of the haemodynamic responses in R1 and R2, respectively.

(= 10 [s] + 60 [s] + 55 [s])) (see Fig. 1). To confirm the reproducibility of each single index between the 2 measurements, a test–retest analysis (single-measure intra-class correlation (ICC) analysis using a one-way random effect model) revealed the presence of significant intra-class correlation coefficients for both the R1 and R2 integral values [$r = 0.47$, $p = 0.01$; $r = 0.59$, $p < 0.01$, respectively] and the R1 centroid value [$r = 0.65$, $p < 0.01$], but not for the R2 centroid value [$r = 0.20$, $p = 0.19$] (see Supplementary material (IV)). PCA and ICC analyses revealed that the 2 indices of NIRS analysis during VFT were acceptable at the single-individual and cluster levels. Thus, the 3 significant variables were used for further analysis.

The 2 representative R1 and R2 NIRS signals obtained from each individual were averaged separately for each type of [Hb] and the integral and centroid values were calculated using parametric statistical tests. Further analyses focused on the increases in [oxy-Hb], because these appear to reflect task-related cortical activation more directly than do decreases in [deoxy-Hb], as evidenced by the stronger correlation between the former and the blood oxygenation level-dependent signal measured by fMRI (Strangman et al., 2002b) and by the results of animal studies (Hoshi et al., 2001). As the typical [oxy-Hb] activation pattern had a positive direction (Fig. 1), data with positive [oxy-Hb] changes (i.e., data with an integral value ≥ 0) in R1 and R2 were used to create an algorithm (Flow diagram (4)). Data exhibiting negative [oxy-Hb] changes were added to the analysis and were described in the results as being appropriate. The analysis of [deoxy-Hb] changes was reported in Supplementary Material (V); however, no significant variable was found regarding [deoxy-Hb] changes.

First, as a preliminary analysis to identify the variable that differentiates patients with psychiatric diseases from HCs most robustly, the 3 variables, including both integral and centroid values of the R1 NIRS signal and the integral value of the R2 NIRS signal, were compared among all of the patients and the age- and gender-matched controls at the initial study site using ANOVA. The resulting significant variables were applied to ROC analyses at the remaining 6 sites.

Because mental health professionals in real clinical settings must differentiate patients with MDD from those with BP or SZ presenting with depression as accurately as possible, the second main analysis performed here aimed to determine the most informative variable and the optimal threshold to discriminate patients with MDD from those with non-MDD disorders. In the present study, the 3 variables, including both integral and centroid indices of R1 and the integral R2 index of the NIRS signal, were compared among patients with MDD and those with either of the other 2 disorders using ANOVA; the variables that were deemed to be significant were applied to the ROC analysis. The preliminary data from the initial site were used to determine an optimal threshold, which was then validated using the test data from the remaining 6 sites.

Third, Pearson's correlation analysis was performed between the significant variables and demographic confounding factors. Data were tested for a normal distribution using the Kolmogorov–Smirnov test. Data that were not normally distributed were analysed using Spearman's correlation analysis.

In particular, regarding clinical confounding factors, such as symptoms (HAMD, YMRS and PANSS scores) and medication doses (antidepressants: imipramine (IMP) equivalent dose; antipsychotics: chlorpromazine (CPZ) equivalent dose; anxiolytics: diazepam equivalent dose; and anti-parkinsonian drugs: biperiden equivalent dose, lithium dose, sodium valproate dose and carbamazepine dose), a stepwise multiple linear regression analysis was performed with a probability of F for conservative entry and removal criteria of 0.01 and 0.05, respectively, to elucidate the complicated relationships among these clinical confounding factors in each diagnostic group.

All data are expressed as mean and standard deviation (SD). The significance level was set to alpha = 0.05. When a difference was considered significant, we presented both the effect size (Cohen's *d*) and the 95% confidence interval (CI). Statistical analyses were performed using the SPSS 16.0.1J software (SPSS Inc., Tokyo, Japan).

Results

Demographic characteristics

Table 1 shows the demographic and clinical characteristics of the 4 age- and gender-matched diagnostic groups used in this study. One-way ANOVA revealed an absence of significant age differences among the groups ($p = 0.99$) and a chi-squared test showed an absence of gender differences among the groups ($p = 0.81$). In addition, the age and gender distributions among the 4 diagnostic groups were not significantly different at the initial site (Gunma University, MDD: 39.9 (11.7) y.o., 12/15; BP: 41.1 (13.2) y.o., 22/15; SZ: 40.1 (14.9) y.o., 11/20; and HC: 40.0 (4.2) y.o., 7/10) (age, $p = 0.98$; gender, $p = 0.24$) and at the other 6 sites (MDD: 44.6 (12.7) y.o., 65/61; BP: 45.1 (15.4) y.o., 47/50; SZ: 44.8 ± 11.0 y.o., 56/49; and HC: 44.0 ± 15.9 y.o., 307/266) (age, $p = 0.89$; gender, $p = 0.81$).

Preliminary test of the difference between HCs and patients

Although it was not the main theme of this study, to compare our results with those of studies of biomarkers performed only to detect functional abnormalities in patients against a control group, we also analysed the differences between HC individuals and patients to confirm the significance of the 3 variables chosen for analysis. Full analyses are described in Supplementary Material (VI).

From the analyses performed using data from the initial site, we adopted both R1 and R2 integral values as statistically significant variables for the algorithm. Thresholds were dependent on the