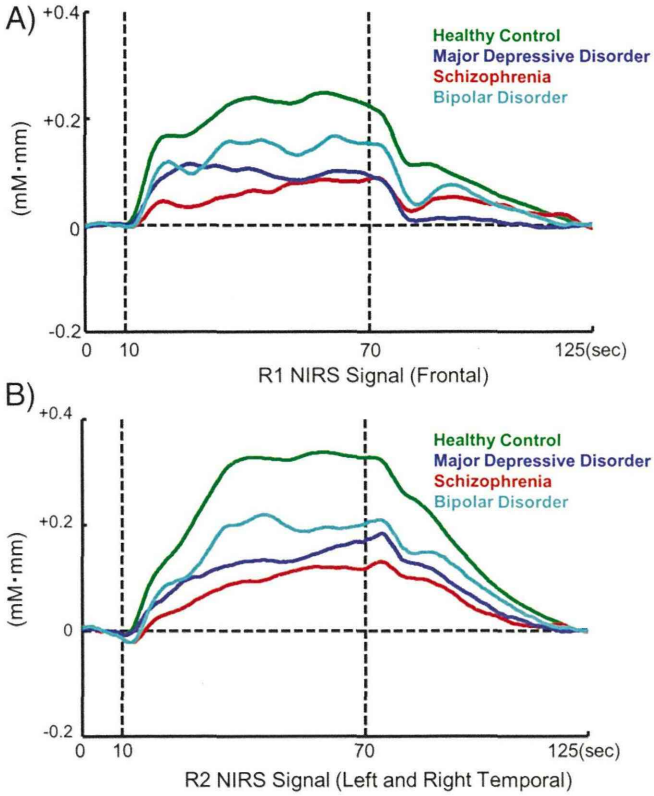


Fig. 3. Time courses of the haemodynamic responses in Region 1 (R1) and Region 2 (R2) in the 4 diagnostic groups. Panels A and B show the time courses of the haemodynamic responses in R1 and R2, respectively.

(= 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

purpose for which the variables were used (Table 2). For example, if the optimal thresholds of the integral values of R1 and R2 derived from the initial site (73 and 104) were applied to the independent test data from the remaining 6 sites, the sensitivities were 0.73 (proportion of patients/measurement: 96/131) and 0.79 (104/131) and the specificities were 0.63 (proportion of HCs/measurement: 326/514) and 0.63 (324/514) for R1 (positive predictive value (PPV) = 0.37, negative predictive value (NPV) = 0.90) and R2 (PPV = 0.40, NPV = 0.92), respectively.

Test for differentiation of patients with unipolar MDD from those with BP and SZ

Using the preliminary data from the initial site, one-way ANOVA performed between the patients with MDD and those with one of the other 2 disorders of interest (BP or SZ) revealed a significant difference in the R1 centroid values [$F(1,53) = 9.54, p < 0.01$; $d = 0.96$, 95% CI, (0.25 to 1.62)], but not in the R1 [$F(1,53) = 0.14, p = 0.71$] or the R2 [$F(1,53) = 0.05, p = 0.83$] integral values.

As the significant R1 centroid value proved to be the most useful variable, we applied it to ROC analysis for the differentiation of patients with unipolar MDD from those with non-MDD disorders. The resulting area under the ROC curve (Az) value was 0.74 [95% CI, (0.61 to 0.87)] and the optimal threshold was 54 [s] from the extreme top left point of the ROC curve (eFig. S4).

To validate the optimal threshold calculated, we applied it to the independent test data of the remaining 6 sites, to differentiate the patients with MDD from those with SZ and BP [Az = 0.81, 95% CI, (0.74 to 0.89); $d = 1.17$, 95% CI, (0.79 to 1.54); optimal threshold = 54 [s], PPV = 0.79, NPV = 0.82; Fig. 4]. Using this threshold (54 [s]), 74.6% of the patients with MDD (proportion of patients/measurement: 41/55) and 85.5% of those with SZ or BP (65/76) were classified correctly [76.9% of BP patients (20/26) and 90.0% of SZ patients (45/50)] (Fig. 5). The ROC curves of MDD v. BP [Az = 0.74, 95% CI, (0.62 to 0.85); $d = 0.81$, 95% CI, (0.32 to 1.29); optimal threshold = 54 [s], PPV = 0.87, NPV = 0.59] and MDD v. SZ [Az = 0.86, 95% CI, (0.78 to 0.93); $d = 1.40$, 95% CI, (0.96 to 1.82); optimal threshold = 54 [s], PPV = 0.89, NPV = 0.78] are shown separately in eFig. S5.

For reference, the test performed for the differentiation between patients with BP and those SZ is shown in Supplementary Material (VII).

Correlational analysis of demographic and clinical confounding factors

Correlational analysis showed no significant correlations between any of the significant dependent variables (among the R1 and R2 integral values and the R1 centroid value of NIRS signals) and any of the demographic confounding factors [performance (number of correct words), education years and pre-morbid IQ; $p > 0.05$] for all patients with psychiatric disorders (MDD, BP and SZ).

Regarding clinical confounding factors, a stepwise regression analysis of each significant dependent variable for each disorder revealed that there was no entry clinical variable in the linear regression models, with the exception of the global assessment of functioning (GAF) score (beta = 0.50, $p < 0.01$) for the R2 integral value ($F = 10.73$, $p < 0.01$; $R = 0.50$, $R^2 = 0.25$, adjusted $R^2 = 0.23$) in patients with MDD, and the GAF score (beta = 0.58, $p = 0.01$) for the R2 integral value ($F = 8.43$, $p = 0.01$; $R = 0.59$, $R^2 = 0.35$, adjusted $R^2 = 0.30$) in patients with BP who exhibited depressive symptoms. Thus, only one clinical variable (i.e., GAF score) among all of the medication and clinical variables examined had a significant impact on the R2 integral values for patients with MDD or BP who exhibited depressive symptoms.

Table 2
Sensitivities and specificities of the integral values of Region 1 (R1) and Region 2 (R2) signals between healthy controls and all patients with psychiatric disorders, based on the independent data collected from the 6 additional sites.

Integral value	R1		R2	
	Sensitivity	Specificity	Sensitivity	Specificity
160	0.95	0.27	0.90	0.39
150	0.93	0.30	0.89	0.43
140	0.92	0.34	0.88	0.47
130	0.90	0.37	0.87	0.52
120	0.88	0.42	0.85	0.57
110	0.85	0.46	0.82	0.61
100	0.82	0.50	0.78	0.64
90	0.78	0.54	0.76	0.68
80	0.74	0.61	0.73	0.73
70	0.72	0.65	0.64	0.76
60	0.66	0.71	0.57	0.78
50	0.57	0.75	0.52	0.81
40	0.47	0.80	0.43	0.86
30	0.38	0.84	0.33	0.89

Discussion

The present multi-site study is the first large-scale, case–control study that demonstrates the utility of NIRS for the differential diagnosis of major psychiatric disorders. The main strengths of this study include the application of a neuroimaging biomarker in clinical practice that allows the clinically useful differential diagnosis of depressive states. The frontal centroid value, which represents the timing of frontal NIRS signal patterns, was a significant variable for differential diagnosis and the optimal threshold derived from the ROC analysis correctly discriminated patients with unipolar MDD (74.6%) from those with non-MDD disorders (85.5%; BP, 76.9% and SZ, 90.0%).

Single-individual diagnostic classification analyses among various psychiatric disorders

The present study was not only a case–control study of group comparisons, but also a study specifically designed for examining the practical utility of single-individual diagnostic classification in various psychiatric disorders. Several studies have reported the single-individual diagnostic classification of one psychiatric disorder compared with HCs by applying multivariate statistical methods (e.g.,

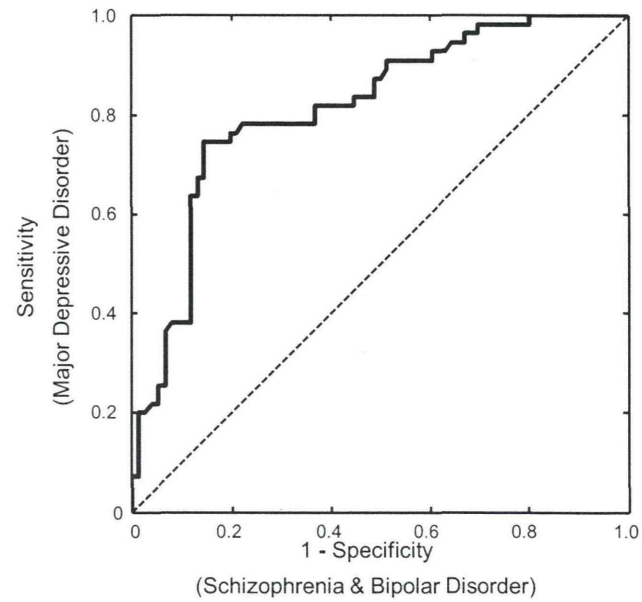


Fig. 4. Receiver operating characteristic analysis of the centroid value of Region 1 (R1) near-infrared spectroscopy signal between patients with major depressive disorder and those with either of the other 2 disorders of interest (bipolar disorder and schizophrenia) based on the independent data collected from the 6 additional sites.

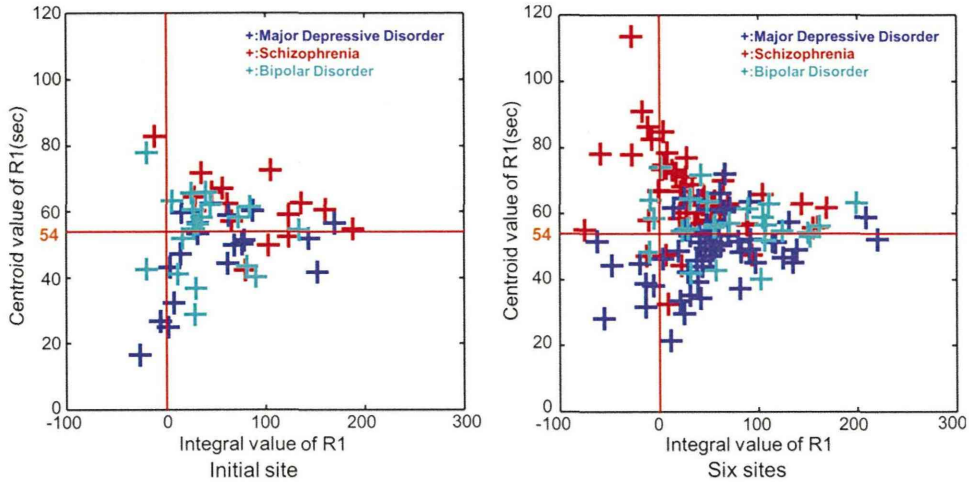


Fig. 5. Scatter plots of the centroid and integral values of Region 1 (R1) signal in the patients, both at the initial site (Gunma University) and at the 6 additional sites.

neuroanatomical pattern classification) to structural MRI data (Davatzikos et al., 2005) and NIRS data (Hahn et al., 2013) from SZ and high-risk psychosis samples (Koutsouleris et al., 2009), as well as to functional MRI data from patients with depression (Hahn et al., 2011). These studies were technically sophisticated; however, more research must be performed to test their reproducibility and generalisability in the advanced stage of clinical application, because (1) they were designed for the analysis of one diagnostic classification based on comparison to HCs, and not for differential diagnosis among multiple psychiatric disorders; and (2) they were performed using one relatively small cohort; thus, they must be replicated in another cohort including larger sample groups.

Furthermore, we will discuss briefly our results in comparison with those of other single-individual diagnostic classification studies (Davatzikos et al., 2005; Fu et al., 2008; Hahn et al., 2011; Koutsouleris et al., 2009). We used only a single variable (simple 'centroid value' of NIRS signals) and found that the classification rates (unipolar MDD: 74.6% correct classification; the 2 other disorders: 85.5% correct classification (BP, 76.9%; SZ, 90.0%)) were almost equivalent to the rates reported in the previous MRI studies using multivariate statistical methods (which had 80–90% classification rates in the patient group compared with the HC group).

To determine whether a higher disease classification rate could be achieved by using a multivariate pattern analysis (compared with that obtained using one simple variable), which was used in previous MRI studies, we confirmed the results using the multivariate pattern classification analysis described in Supplementary Material (VIII). The leave-one-out cross-validation method revealed that 4 significant variables, or even one variable (the R1 centroid value), could differentiate patients with unipolar MDD from those with either of the 2 other disorders (non-MDD) with a similar degree of mean accuracy (76.8% (unipolar MDD: 73.0% (54/74), non-MDD: 74.8% (83/111))).

Clinical importance and implications

Another clinically valuable feature of our work is that it aimed to facilitate diagnosis among patients with similar depressive symptoms, which psychiatrists often find to be a difficult task. Most BP patients with depressive symptoms are initially diagnosed with and treated for MDD (Akiskal et al., 1995; Goldberg et al., 2001). Therefore, our findings may help differentiate BP with depressive symptoms from MDD. Depressive symptoms and cognitive deficits are also common early signs of SZ (Hafner et al., 2005). Of particular clinical relevance is the observation that SZ patients with concomitant depression have a greater risk of suicide or an unfavourable disease course (an der Heiden et al., 2005). Therefore, sufficient attention must be given to the diagnosis and treatment of depression in SZ patients.

The results of the present study may draw attention to the heterogeneity observed among MDD patients. Rather than simply being misclassified, approximately 25% of patients with unipolar MDD who were classified by the system as having a non-MDD disorder may have a brain pathophysiology that is biologically different from that of the majority of MDD patients. Evidence suggests that 25–50% of individuals with recurrent major depression (particularly those in atypical early-onset or treatment-refractory subgroups) may in fact have broadly defined BP (Angst, 2007). In this study, 74.6% of the patients with MDD were classified correctly; the remaining 25.4% might include either patients who would progress to a diagnosis of one of the 2 other disorders or patients with a broadly defined BP who were diagnosed with MDD according to the DSM criteria. This explanation might be justified by the finding of a correct classification rate of 75% for patients with MDD. For practical purposes, among patients diagnosed clinically with MDD, the early suspicion of the possibility of a diagnosis of a non-MDD disorder with depression would also provide an opportunity to reduce the hazardous effects of the illness on personal, social and occupational aspects; therefore, our results should be of great clinical importance in practical applications. Thus, a prospective study aimed at elucidating the heterogeneity of unipolar MDD is required.

Advantages of the NIRS method

We used the same NIRS system (a non-invasive, portable and user-friendly device) and the same concise measurement procedure at every site; therefore, inter-site compatibility was not an issue here; however, it may be an obstacle in other neuroimaging multi-site studies. Furthermore, we used a high temporal resolution (0.1 s) in the NIRS system for measuring time-specific characteristics of dynamic prefrontal cortical functions; this enabled analyses that included more detailed time-course comparisons of NIRS signal changes. We created and adopted new variables, such as the 'centroid value', to determine the timing of the haemodynamic response (Fig. 1). The high temporal resolution of NIRS might allow not only the detection of functional abnormalities (e.g., hypofrontality), but also the capture of the specific haemodynamic activation time courses of each psychiatric disorder and aid differential diagnosis.

The practical application of biomarkers requires that they be relatively simple. The simplicity of both the test procedure and the associated data analysis is important not only for the participants, but also for their caretakers and clinicians. Therefore, rather than using complicated multivariate statistical methods, we developed a robust classification algorithm for real-time visual evaluation of patients using the simplest, and lowest number of variables on the basis of a ROC analysis. This was important because we sought to develop a psychiatric practice empowered by the initiative of patients by sharing the 'comprehensively visualised' results that can be easily recognisable by patients and caretakers, rather than results from complicated 'black-box' analyses. In addition, using the condensed VFT (<3 min) developed previously by us, we designed a diagnostic support system in a way that the results are available to clinicians in less than 15 min. The availability of such a 'comprehensively visualised' report to clinicians, patients and their caretakers at a first visit, while laying out a future treatment plan, would likely lead to a paradigm shift to a patient-centred approach in clinical psychiatry.

Limitations

The methodological aspects of the present study warrant commentary. First, most of the patients included in the study were taking medications at the time of measurement. To our knowledge, no clear evidence of the effects of medication on NIRS signals has been demonstrated. We found that none of the medications at any dose was significantly correlated with NIRS signals in this study; however, we cannot fully exclude the effects of medication on haemodynamic signals. For confirmation, the application of the algorithm described above (optimal cut-off of the R1 centroid value) to the drug-free patients exclusively, 6 out of 10 patients with MDD patients (60%) and 4 out of 5 patients with SZ (80%) were classified correctly. Second, the size of the sample included in our final analysis was substantially reduced from that initially recruited, because we tried to minimise the confounding factors of age and gender by matching the groups and excluded patients in remission, as well as patients in the manic phase (see Flow diagram). In our confirmatory analysis, we included all non-matched and in-remission patients and found that the results were quite similar, although this analysis had a lower detection power. The optimal threshold of the sample sets before demographical matching was also the same as that calculated originally. These results suggest that the reduction in the total number of study participants after demographical matching did not affect the development of the algorithm [see Supplementary Material (I)]. However, we must consider the possibility that this diagnostic support system is best suited for young and middle-aged patients with moderate or severe symptoms (e.g., aged between 23 and 65 years (mean ± 1.5 SD)). Third, a PCA of haemodynamic response performed to capture a channel cluster led to the identification of 2 cluster regions. Nonetheless, as we thought that pooling many NIRS signals together into only 2 representative regions of interest (R1 (frontopolar and dorsolateral prefrontal regions) and R2 (ventrolateral prefrontal and temporal regions)) might oversimplify the results (see the Discussion of Supplementary Material (II)), we sought to confirm the reliability of the 2 clusters by performing a test–retest analysis in a portion of the samples. We found significant ICCs for both the R1 and R2 integral values and for the R1 centroid value between 2 measurements (see Supplementary Material (IV)). Therefore, we used the two data-derived clusters that reflected a fronto-temporal haemodynamic response during VFT. Fourth, we have controlled some well-known confounders in the analyses. However, the NIRS signal might be affected by the other systemic confounders, such as autonomic function, neuroendocrine function, diet and physical activity. In addition, brain anatomical factors, such as scalp–cortex distance and frontal sinus volume, as well as genetic variants might also be potential confounders. Further studies are required to address the relationship between the NIRS signal and these confounders. If these findings are fully replicated, the development of methods of integrating confounding factors into NIRS signal in the future will be ideal. Fifth, we did not use the exclusion criterion of first-degree relatives with axis I psychiatric disorders for healthy controls. This could give a bias to the data in healthy controls, which means that some of the first-degree relatives of persons with axis I psychiatric disorders might have been included as a healthy control in the present study. However, as the same situations are assumed in real clinical settings, we daringly recruited healthy controls without applying that strict exclusion criterion.

Conclusions and future implications

In conclusion, this multi-site study provided evidence that the fronto-temporal NIRS signal may be used as a tool in assisting the diagnosis of major psychiatric disorders with depressive symptoms. Future NIRS research should be performed to study the applicability of this method to (1) the identification of a need for therapy, (2) the assessment of the efficacy of various treatments, (3) the establishment of prognostic predictions that may be clarified by longitudinal follow-up assessments of patients in various clinical stages and (4) the examination of the use of NIRS as a screening tool.

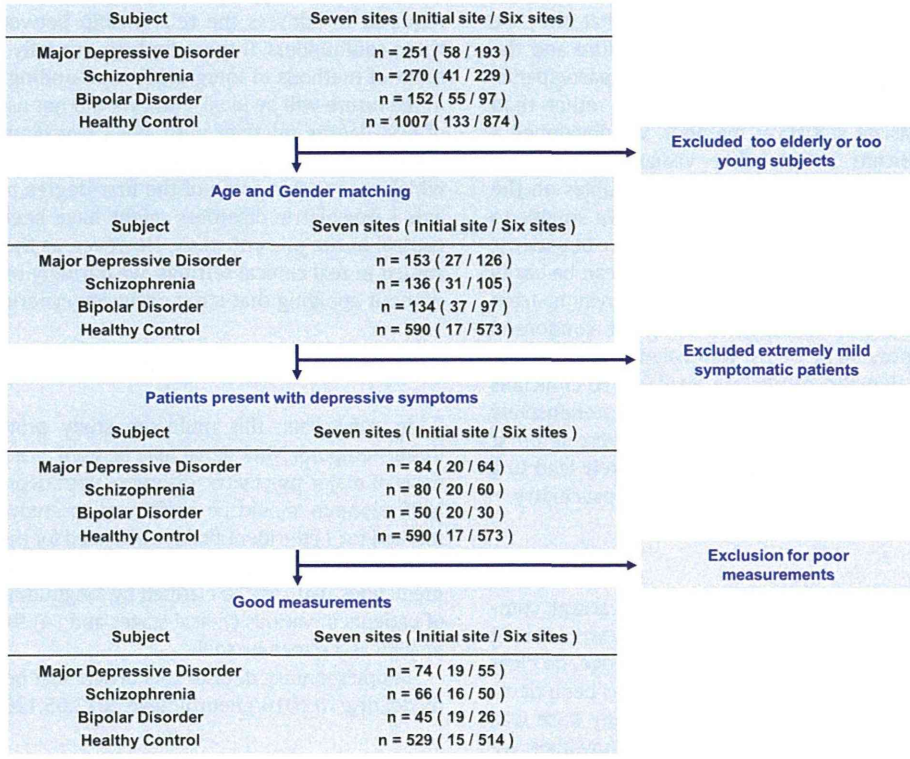
Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2013.05.126.

Acknowledgments

The authors would like to thank all the participants in this study. The authors also thank Makoto Ito, M.D., Ph.D., Tomohiro Suto, M.D., Ph.D., Yutaka Yamagishi, M.D., Naoki Hanaoka, M.D., Ph.D., Toshimasa Sato, Ed.M. and Noriko Sakurai of Gunma University; Kohei Marumo, M.D., Ph.D. and Yuki Kawakubo, Ph.D., of The University of Tokyo; Hitomi Kobayashi, Ph.D. and Junko Motoki, Ph.D., of Showa University; Osamu Saitoh, M.D., Ph.D., Kimitaka Anami, M.D., Ph.D., Yohtaro Numachi, M.D., Ph.D., Yuji Sugimura and Masaru Ogawa of the National Centre of Neurology & Psychiatry Hospital; Daisuke Gotoh, M.D., Yojiro Sakai, M.D., Ph.D. and Emi Yoshida, M.D., of Fukushima Medical University, for technical assistance during data collection.

Role of the funding source

This study was supported in part by grants-in-aid for scientific research from the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan (Nos. 18390318, 19659289, 20390310 and 22659209 to MF; Nos. 21249064 and 221S003 [Innovative Areas (Comprehensive Brain Science Network)] to KK; No. 17019029 [Priority Areas (Applied Genomics)] to YO; and No. 23791309 to RT), and by grants-in-aid from the Ministry of Health, Labour, and Welfare (MHLW) of Japan (H20-kokoro-ippan-001, H20-3, and H22-seishin-ippan-015 to KK). This study was also supported in part by Health and Labour Sciences Research Grants for Comprehensive Research on Disability Health and Welfare (previously, Health and Labour Science Research Grant for Research on Psychiatric and Neurological Disease and Mental Health; H20-001 to MF and H23-seishin-ippan-002 to RT) and by the Intramural Research Grants for Neurological and Psychiatric Disorders of The National Centre of Neurology and Psychiatry (previously, The Research Grant for Nervous and Mental Disorders from the MHLW; 21B-1 to MF and 20B-3 to TN). In addition, this study was supported in part by an Intramural Research Grant for Neurological and Psychiatric Disorders of NCNP (No. 23-10 to RT). A part of this study was also the result of a project entitled 'Development of biomarker candidates for social behavior', which was carried out under the Strategic Research Program for Brain Sciences by MEXT, Japan. This study was also supported in part



Flow diagram. Main analyses were based on matched samples according to the flow diagram.

by grants from the Japan Research Foundation for Clinical Pharmacology (to RT).

The funders had no role in the design of the study, data collection, data analysis, decision to publish, or preparation of the manuscript. The corresponding author had full access to all of the data in this study and takes responsibility for the integrity of the data, the accuracy of the data analysis and the decision to submit the manuscript for publication.

Author contributions

Masato Fukuda designed the experiments and organized the multi-site collaborative study. Ryu Takizawa, Masato Fukuda, Shingo Kawasaki, and Kiyoto Kasai analysed the data and wrote the first draft of the paper. The other contributors performed data acquisition and revised the first draft critically for important intellectual content. All contributors have approved the final version of the manuscript.

Contributors from the Joint Project for Psychiatric Application of Near-Infrared Spectroscopy (JPSY-NIRS) Group

Gunma University: Masato Fukuda, M.D., Ph.D., Masashi Suda, M.D., Ph.D., Yuichi Takei, M.D., Ph.D., Yoshiyuki Aoyama, M.D., Ph.D., Kosuke Narita, M.D., Ph.D., Masahiko Mikuni, M.D., Ph.D., Masaki Kameyama, M.D., Ph.D., and Toru Uehara, M.D., Ph.D.

The University of Tokyo: Ryu Takizawa, M.D., Ph.D., Kiyoto Kasai, M.D., Ph.D., Masaru Kinou, M.D., Ph.D., Shinsuke Koike, M.D. and Ayaka Ishii-Takahashi, M.D.

Hitachi Medical Corporation: Shingo Kawasaki, MS, Noriyoshi Ichikawa, BS, and Michiyuki Fujiwara, AS.

Showa University: Haruhisa Ohta, M.D., Ph.D., Hiroi Tomioka, M.D., Ph.D., Bun Yamagata, M.D., Ph.D., and Kaori Yamanaka, M.D.

Keio University: Masaru Mimura, M.D., Ph.D.

Tottori University: Shenghong Pu, Ph.D.

The National Centre of Neurology and Psychiatry: Kazuyuki Nakagome, M.D., Ph.D., Takamasa Noda, M.D., Taro Matsuda, M.D., and Sumiko Yoshida, M.D., Ph.D.

Fukushima Medical University: Shin-Ichi Niwa, M.D., Ph.D., Soichi Kono, M.D., Hirooki Yabe, M.D., Ph.D., and Sachie Miura, M.D.

Mie University and Tokyo Metropolitan Matsuzawa Hospital: Yuji Okazaki, M.D., Yukika Nishimura, Ph.D., Hisashi Tanii, M.D., Ph.D., Ken Inoue, M.D., Ph.D., Chika Yokoyama, MA, Yoichiro Takayanagi, M.D., Ph.D., Katsuyoshi Takahashi, M.D., and Mayumi Nakakita, M.D.

Conflict of interest

We would like to disclose potential conflicts of interest regarding all financial and material support for the present study. The principal investigators of each site (Masato Fukuda of Gunma University, Kiyoto Kasai of The University of Tokyo, Masaru Mimura of Keio university and Showa University, Kazuyuki Nakagome of The National Center of Neurology and Psychiatry and Tottori University, Shin-ichi Niwa of Fukushima Medical University, Yuji Okazaki of both Mie University and Tokyo Metropolitan Matsuzawa Hospital and Takamasa Noda of The National Center of Neurology and Psychiatry) have potential conflicts of interest in the submitted work. Each site has had an official contract with the Hitachi Group (Advanced Research Laboratory, Hitachi, Ltd., and The Research and Developmental Center, Hitachi Medical Corporation) for a collaborative study of the clinical application of NIRS in psychiatric disorders. For this study, the Hitachi Group provided a project grant (JPY 300,000–2,000,000 per year) and material support (temporary rental of a NIRS (Optical Topography) ETG-4000 system) for each site. Shingo Kawasaki, Noriyoshi Ichikawa and Michiyuki Fujiwara are employees of Hitachi Medical Corporation.

The other authors have no financial relationships with any organisations that might have an interest in the submitted work in previous years and no other relationships or activities that could appear to have influenced the submitted work.

References

Akiskal, H.S., Maser, J.D., Zeller, P.J., Endicott, J., Coryell, W., Keller, M., Warshaw, M., Clayton, P., Goodwin, F., 1995. Switching from 'unipolar' to bipolar II. An 11-year

prospective study of clinical and temperamental predictors in 559 patients. Arch. Gen. Psychiatry 52, 114–123.
Almeida, J.R., Versace, A., Mechelli, A., Hassel, S., Quevedo, K., Kupfer, D.J., Phillips, M.L., 2009. Abnormal amygdala–prefrontal effective connectivity to happy faces differentiates bipolar from major depression. Biol. Psychiatry 66, 451–459.
an der Heiden, W., Konnecke, R., Maurer, K., Ropeter, D., Hafner, H., 2005. Depression in the long-term course of schizophrenia. Eur. Arch. Psychiatry Clin. Neurosci. 255, 174–184.
Angst, J., 2007. The bipolar spectrum. Br. J. Psychiatry 190, 189–191.
Barch, D.M., Sheline, Y.I., Csernansky, J.G., Snyder, A.Z., 2003. Working memory and prefrontal cortex dysfunction: specificity to schizophrenia compared with major depression. Biol. Psychiatry 53, 376–384.
Barrett, S.L., Mulholland, C.C., Cooper, S.J., Rushe, T.M., 2009. Patterns of neurocognitive impairment in first-episode bipolar disorder and schizophrenia. Br. J. Psychiatry 195, 67–72.
Costafreda, S.G., Fu, C.H., Lee, L., Everitt, B., Brammer, M.J., David, A.S., 2006. A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus. Hum. Brain Mapp. 27, 799–810.
Curtis, V.A., Dixon, T.A., Morris, R.G., Bullmore, E.T., Brammer, M.J., Williams, S.C., Sharma, T., Murray, R.M., McGuire, P.K., 2001. Differential frontal activation in schizophrenia and bipolar illness during verbal fluency. J. Affect. Disord. 66, 111–121.
Davatzikos, C., Shen, D., Gur, R.C., Wu, X., Liu, D., Fan, Y., Hughett, P., Turetsky, B.I., Gur, R.E., 2005. Whole-brain morphometric study of schizophrenia revealing a spatially complex set of focal abnormalities. Arch. Gen. Psychiatry 62, 1218–1227.
Ferrari, M., Quaresima, V., 2012. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. Neuroimage 63 (2), 921–935.
First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 1997. Structured Clinical Interview for DSM-IV Axis I Disorders. Biometric Research Department, New York State Psychiatric Institute, New York, U.S.A.
Fu, C.H., Mourao-Miranda, J., Costafreda, S.G., Khanna, A., Marquand, A.F., Williams, S.C., Brammer, M.J., 2008. Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. Biol. Psychiatry 63, 656–662.
Goldberg, J.F., Harrow, M., Whiteside, J.E., 2001. Risk for bipolar illness in patients initially hospitalized for unipolar depression. Am. J. Psychiatry 158, 1265–1270.
Gur, R.E., Keshavan, M.S., Lawrie, S.M., 2007. Deconstructing psychosis with human brain imaging. Schizophr. Bull. 33, 921–931.
Hafner, H., Maurer, K., Trendler, G., an der Heiden, W., Schmidt, M., 2005. The early course of schizophrenia and depression. Eur. Arch. Psychiatry Clin. Neurosci. 255, 167–173.
Hahn, T., Marquand, A.F., Ehlis, A.C., Dresler, T., Kittel-Schneider, S., Jarczok, T.A., Lesch, K.P., Jakob, P.M., Mourao-Miranda, J., Brammer, M.J., Fallgatter, A.J., 2011. Integrating neurobiological markers of depression. Arch. Gen. Psychiatry 68, 361–368.
Hahn, T., Marquand, A.F., Plichta, M.M., Ehlis, A.C., Schecklmann, M.W., Dresler, T., Jarczok, T.A., Eirich, E., Leonhard, C., Reif, A., Lesch, K.P., Brammer, M.J., Mourao-Miranda, J., Fallgatter, A.J., 2013. A novel approach to probabilistic biomarker-based classification using functional near-infrared spectroscopy. Hum. Brain Mapp. 34, 1102–1114.
Hamilton, M., 1960. A rating scale for depression. J. Neurol. Neurosurg. Psychiatry 23, 56–62.
Holsboer, F., 2008. How can we realize the promise of personalized antidepressant medicines? Nat. Rev. Neurosci. 9, 638–646.
Hoshi, Y., Kobayashi, N., Tamura, M., 2001. Interpretation of near-infrared spectroscopy signals: a study with a newly developed perfused rat brain model. J. Appl. Physiol. 90, 1657–1662.
Kameyama, M., Fukuda, M., Yamagishi, Y., Sato, T., Uehara, T., Ito, M., Suto, T., Mikuni, M., 2006. Frontal lobe function in bipolar disorder: a multichannel near-infrared spectroscopy study. NeuroImage 29, 172–184.
Kay, S.R., Opler, L.A., Fiszbein, A., 1991. Positive and Negative Syndrome Scale (PANSS) Rating Manual. Multi-health Systems Inc., Toronto.
Koike, S., Takizawa, R., Nishimura, Y., Takano, Y., Takayanagi, Y., Kinou, M., Araki, T., Harima, H., Fukuda, M., Okazaki, Y., Kasai, K., 2011. Different hemodynamic response patterns in the prefrontal cortical sub-regions according to the clinical stages of psychosis. Schizophr. Res. 132, 54–61.
Koutsouleris, N., Meisenzahl, E.M., Davatzikos, C., Bottlender, R., Frodl, T., Scheuerecker, J., Schmitt, G., Zetzsche, T., Decker, P., Reiser, M., Moller, H.J., Gaser, C., 2009. Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. Arch. Gen. Psychiatry 66, 700–712.
Mathers, C.D., Loncar, D., 2006. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 3, e442.
McGorry, P.D., Killackey, E., Yung, A., 2008. Early intervention in psychosis: concepts, evidence and future directions. World Psychiatry 7, 148–156.
Obrig, H., Villringer, A., 2003. Beyond the visible—imaging the human brain with light. J. Cereb. Blood Flow Metab. 23, 1–18.
Okada, E., Delpy, D.T., 2003. Near-infrared light propagation in an adult head model. II. Effect of superficial tissue thickness on the sensitivity of the near-infrared spectroscopy signal. Appl. Opt. 42, 2915–2922.
Phillips, M.L., Vieta, E., 2007. Identifying functional neuroimaging biomarkers of bipolar disorder: toward DSM-V. Schizophr. Bull. 33, 893–904.
Prince, M., Patel, V., Saxena, S., Maj, M., Maselko, J., Phillips, M.R., Rahman, A., 2007. No health without mental health. Lancet 370, 859–877.
Schecklmann, M., Ehlis, A.C., Plichta, M.M., Fallgatter, A.J., 2008. Functional near-infrared spectroscopy: a long-term reliable tool for measuring brain activity during verbal fluency. Neuroimage 43 (1), 147–155.
Shattuck, D.W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K.L., Poldrack, R.A., Bilder, R.M., Toga, A.W., 2008. Construction of a 3D probabilistic atlas of human cortical structures. NeuroImage 39, 1064–1080.
Strangman, G., Boas, D.A., Sutton, J.P., 2002a. Non-invasive neuroimaging using near-infrared light. Biol. Psychiatry 52, 679–693.
Strangman, G., Culver, J.P., Thompson, J.H., Boas, D.A., 2002b. A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. NeuroImage 17, 719–731.
Suto, T., Fukuda, M., Ito, M., Uehara, T., Mikuni, M., 2004. Multichannel near-infrared spectroscopy in depression and schizophrenia: cognitive brain activation study. Biol. Psychiatry 55, 501–511.
Takizawa, R., Kasai, K., Kawakubo, Y., Marumo, K., Kawasaki, S., Yamasue, H., Fukuda, M., 2008. Reduced frontopolar activation during verbal fluency task in schizophrenia: a multi-channel near-infrared spectroscopy study. Schizophr. Res. 99, 250–262.
Tsuzuki, D., Jurcak, V., Singh, A.K., Okamoto, M., Watanabe, E., Dan, I., 2007. Virtual spatial registration of stand-alone fNIRS data to MNI space. NeuroImage 34, 1506–1518.
Yamashita, Y., Maki, A., Ito, Y., Watanabe, E., Koizumi, H., 1996. Noninvasive near-infrared topography of human brain activity using intensity modulation spectroscopy. Opt. Eng. 35, 1046–1049.
Young, R.C., Biggs, J.T., Ziegler, V.E., Meyer, D.A., 1978. A rating scale for mania: reliability, validity and sensitivity. Br. J. Psychiatry 133, 429–435.
Zanelli, J., Reichenberg, A., Morgan, K., Fearon, P., Kravariti, E., Dazzan, P., Morgan, C., Zanelli, C., Demjaha, A., Jones, P.B., Doody, G.A., Kapur, S., Murray, R.M., 2010. Specific and generalized neuropsychological deficits: a comparison of patients with various first-episode psychosis presentations. Am. J. Psychiatry 167, 78–85.
Zimmermann, P., Bruckl, T., Nocon, A., Pfister, H., Lieb, R., Wittchen, H.U., Holsboer, F., Angst, J., 2009. Heterogeneity of DSM-IV major depressive disorder as a consequence of subthreshold bipolarity. Arch. Gen. Psychiatry 66, 1341–1352.

Ishijima et al. Arthritis Research & Therapy 2014, 16:R18
http://arthritis-research.com/content/16/1/R18

arthritis research&therapy

RESEARCH ARTICLE Open Access

Intra-articular hyaluronic acid injection versus oral non-steroidal anti-inflammatory drug for the treatment of knee osteoarthritis: a multi-center, randomized, open-label, non-inferiority trial

Muneaki Ishijima[1*], Toshitaka Nakamura[2], Katsuji Shimizu[3], Kunihiko Hayashi[4], Hiraku Kikuchi[5], Satoshi Soen[6], Go Omori[7], Toshihiko Yamashita[8], Yuji Uchio[9], Junji Chiba[10], Yuki Ideno[4], Mitsuaki Kubota[1], Hisashi Kurosawa[1], Kazuo Kaneko[1] and for the Research Group of Cartilage Metabolism

Abstract

Introduction: While many of the commonly used conservative treatments for knee osteoarthritis (OA) have been recognized to be effective, there is still insufficient evidence available. Among the pharmacological treatments for knee OA, oral non-steroidal anti-inflammatory drugs (NSAIDs) act rapidly and are recommended for the management of OA. However, frequent and serious adverse effects of NSAIDs have been recognized. Intra-articular injections of hyaluronic acid (IA-HA) for the treatment of knee OA have been shown to reduce pain and improve joint function. However, there has been no qualified direct comparison study of the efficacy and safety between IA-HA and NSAIDs for patients with knee OA. The aim of this study was to clarify the efficacy and safety of early-phase IA-HA in comparison to those of NSAIDs for patients with knee OA.

Methods: This multicenter, randomized, open-label, parallel-group, non-inferiority comparison study with an oral NSAID involved a total of 200 patients with knee OA. An independent, computer-generated randomization sequence was used to randomly assign patients in a 1:1 ratio to NSAIDs three times per day for five weeks (n = 100) or IA-HA once a week for five weeks (n = 100). The primary endpoint was the percentage change in the patient-oriented outcome measure for knee OA, the Japanese Knee Osteoarthritis Measure (JKOM) score. All patients were questioned regarding any adverse events during treatment. The full analysis set (FAS) was used for analysis. The margin of non-inferiority was 10%.

Results: The analyses of primary endpoint included 98 patients in the IA-HA group and 86 patients in the NSAID group. The difference in the percentage changes of the JKOM score between the two intervention arms (IA-HA; -34.7% ($P<0.001$), NSAID; -32.2% ($P<0.001$)) was −2.5% (95% confidence interval (CI): -14.0 to 9.1), indicating IA-HA was not inferior to NSAID. The frequency of both withdrawal and adverse events in the IA-HA group were significantly lower than those in the NSAID group ($P = 0.026$ and 0.004, respectively).

Conclusions: The early efficacy of IA-HA is suggested to be not inferior to that of NSAIDs, and that the safety of the early phase of IA-HA is superior to that of NSAIDs for patients with knee OA.

Trial registration: UMIN Clinical Trials Registry (UMIN-CTR), UMIN000001026.

* Correspondence: ishijima@juntendo.ac.jp
[1]Department of Medicine for Orthopaedics and Motor Organ, Juntendo University Graduate School of Medicine, 2-1-1, Hongo, Bunkyo-ku, Tokyo 113-8421, Japan
Full list of author information is available at the end of the article

Introduction

Osteoarthritis (OA) is an increasingly important public-health problem [1]. The total societal cost of the treatment of OA has been estimated to increase worldwide because of its dramatic growth in morbidity [2]. The current treatment for knee OA consists of conservative treatment, such as exercise, physical therapy, pharmacological agents and, in some cases, surgical treatment [3,4]. While many of the commonly used conservative treatments have been recognized to be effective [5], there is still insufficient evidence available.

Among the pharmacological treatments for knee OA, oral non-steroidal anti-inflammatory drugs (NSAIDs) act rapidly and are recommended for the management of OA, although frequent and serious adverse effects of NSAIDs have been recognized [5]. Hyaluronic acid (HA) is a natural constituent of joint fluid. Intra-articular injections of HA (IA-HA) for the treatment of knee OA have been shown to reduce the pain and improve joint function [5-7]. Although IA-HA is also recommended, it acts relatively slowly and there was considerable heterogeneity in the outcomes between trials [8-11]. In addition, there has been no qualified direct comparison study of efficacy and safety between IA-HA and NSAIDs for patients with knee OA.

The aim of this multicenter, randomized, parallel-group, open-label, non-inferiority trial was to compare the early efficacy and safety of IA-HA and NSAIDs in patients with knee OA.

Methods

Study design and participants

The trial was planned by the Cartilage Metabolism Research Group, consisting mainly of Japanese orthopedists, to clarify the early efficacy and safety of IA-HA (high molecular weight 2,700 kDa HA, Chugai Pharmaceutical Co. Ltd., Tokyo, Japan) in comparison to an NSAID (loxoprofen sodium, Daiichi Sankyo Pharmaceuticals Co. Ltd., Tokyo, Japan), in a multicenter, randomized, open-label, parallel-group, non-inferiority trial. The protocol was reviewed and approved by the ethics committee of Juntendo University, Tokyo, Japan, and was also reviewed by the institutional review board of each participating institution. This study was undertaken at 20 hospitals throughout Japan between February, 2008 and December, 2010 (see Acknowledgements), in accordance with the Declaration of Helsinki, and with the Ethical Guidelines for Clinical Studies of the Japanese Ministry of Health, Labor, and Welfare. This trial was registered at UMIN-CTR [12], UMIN000001026.

Subjects

All patients provided written informed consent before enrollment in this trial. The inclusion criteria for the present study included (1) subjects who were able to walk with painful knee OA and fulfilled the criteria for knee OA of the medial femorotibial joint as defined by the American College of Rheumatology (ACR) [13], (2) the age of the subjects ranged from 50 to 80, (3) female subjects were required to be postmenopausal, and (4) all subjects had radiographic knee OA with Kellgren-Lawrence (K/L) grade 1 to 3 [14] evaluated by the weight-bearing anteroposterior X-rays of the tibiofemoral joint using the bilateral standing extended view.

The exclusion criteria included (1) patients who had received either an oral, topical or intra-articular steroid during the four weeks before the study, (2) patients who had received IA-HA within four weeks before the study, (3) patients who had received either an oral, topical or suppository NSAID within two weeks before the study, (4) patients who had secondary knee OA, (5) patients with patellofemoral OA with a K/L grade of 3 or higher, (6) patients with severe OA (K/L grade 3 or higher) in a location other than the knee joint, (7) patients with rheumatoid arthritis, (8) patients who had received joint replacement surgery in either knee or/and a hip, (9) patients who had allergies to either HA or NSAIDs, (10) patients who had either hematological, cardiac, hepatic or renal disorders, (11) patients who had experienced an asthma attack induced by NSAIDs, and (12) patients whom the physician recognized as not suitable for enrollment in the study for other reasons.

Randomization and masking

A centralized, computer-generated randomization was conducted to randomly assign patients in a 1:1 ratio to the IA-HA or NSAID groups. Investigators were masked to assignment before, but not after, randomization. The website for patient registration and randomization was prepared and controlled by the coordinating data center (Gunma University, Maebashi, Japan). The blocked randomization was stratified by the participating medical center and the K/L grade of knee OA.

Treatment procedures

A total of 200 patients with symptomatic knee OA were registered from 20 hospitals and randomized for treatment with the NSAID or IA-HA, as described above. For patients treated with the NSAID, they received three daily 60 mg NSAID tablets (total 180 mg)/day, one after each meal, for five weeks. Additional use of gastro-protective drugs, such as a proton pump inhibitor (PPI), in combination with the NSAID was allowed for those in the NSAID group. For patients treated with IA-HA, an intra-articular injection of high-molecular-weight 2,700 kDa HA (25 mg) was administered into the affected joint five times, at weekly intervals in the morning. Concomitant use of other drugs for the treatment of OA and drugs that affect bone and cartilage metabolism were not allowed during the trial.

Outcome measures for the assessment of efficacy and safety

The patients were evaluated for their (1) baseline characteristics, (2) radiographic analysis of the knee, (3) compliance with the treatment, (4) clinical manifestations, and (5) safety.

Evaluation of the response to treatment (efficacy)

Pain was evaluated by a visual analog scale (VAS, 0 to 100). The clinical manifestations were evaluated by the Japanese Knee Osteoarthritis Measure (JKOM) score [15]. The JKOM is a patient-based, self-answered evaluation score that includes four subcategories: pain and stiffness (0 to 32), activities of daily living (0 to 40), social activities (0 to 20), and general health conditions (0 to 8) with 100 points as the maximum score. The JKOM score is higher in patients with more pain and physical disability, and this evaluation modality is considered to have sufficient reliability and validity for studies of the clinical outcomes of Japanese subjects with knee OA [15]. The measure has also been shown to have reliability and validity by means of statistical evaluation and comparison with other health-related scales, such as the Western Ontario and McMaster Universities Arthritis Index (WOMAC) and the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) [15].

The primary endpoint was to compare the percentage change from baseline in the JKOM score at five weeks. The secondary endpoint was to compare the percentage change from baseline in the pain VAS score.

The definition of a response to treatment was made following the criteria defined by the Outcome Measures in Rheumatology Clinical Trials (OMERACT) and Osteoarthritis Research Society International (OARSI) [16]. This measure consists of both absolute and relative changes in scales, including both pain and function, to evaluate the affected knee. Relative change means the percentage of change during the study (final minus baseline over baseline × 100), whereas absolute change indicates the absolute change during the study (final minus baseline on an interval scale of 0 to 100). Before assessing patients based on this scale, we partly modified it for this study by using the JKOM score, as already reported [17,18]. The response was defined as relief of joint pain or improvement in function (at least 50% reduction of the score) and a decrease of at least 20 mm on the VAS, or clinical improvement meeting at least two of the following three conditions: a decrease in joint pain of at least 20% and at least 10 mm on the VAS; an improvement in function of at least 20% and a decrease of at least 4 points from a total 40 points (equal to an absolute change of 10%) on the JKOM functional subcategory scale; and a decrease in the patient's global assessment score by at least 20% and at least 10 points from a total of 100 on the total JKOM scale.

Assessment of adverse events induced according to the treatment modality (safety)

Safety was monitored by recording all adverse events, evaluating the laboratory data and assessing vital signs. This was performed for all participants in both groups at each weekly visit.

Statistical analysis

Sample size determination

The trial was designed to establish whether the symptom-modified effect of IA-HA was non-inferior to that of NSAID (Δ10%). The sample size of this non-inferiority trial was calculated to require a total of 194 patients (97 per treatment group) based on the results of our pilot study with a 5% dropout rate, 10% non-inferiority margin, 27% standard deviation (SD), 5% one-sided alpha level, and power = 0.8 (pilot study: Toshitaka Nakamura, unpublished data, 2007). The 10% margin was set as the smallest value that would be clinically important, assuming a reduction of 30% in the mean percentage change of JKOM score in patients with both IA-HA and NSAID treatment and a reduction of 10% those receiving a placebo treatment.

Data analysis

The primary statistical analyses of efficacy and safety were performed on the full analysis set (FAS), which included all patients treated at least once. For the primary endpoint of the study, a two-sided 95% confidence interval (CI) for the group difference 'test treatment minus reference treatment' was calculated for the percentage change from baseline in the JKOM score as non-inferiority analysis. The non-inferiority of the test treatment was confirmed if the upper limit of the CI was ≤ margin of non-inferiority delta (10%). For the secondary endpoint, the group difference and its 95% CI was calculated for the percentage change from baseline in the VAS pain score.

Quantitative variables were described using the mean, standard deviation and range. The efficacy of treatment was examined by a paired *t* test for both JKOM score and pain VAS score. A multiple logistic regression analysis was used to estimate odds ratios and their 95% CIs between the IA-HA and NSAID treatments in models adjusted for age, K/L grade, body mass index (BMI) and the participating medical centers.

All analyses were performed using the SAS System Release 9.1 software program (SAS Institute, Cary, NC, USA). The registration number of this trial is UMIN000001026, and information on the trial can be found online at [12].

Results

Patient baseline characteristics

A flow chart of this trial is shown in Figure 1. When 200 patients were enrolled, half (100) of the patients were

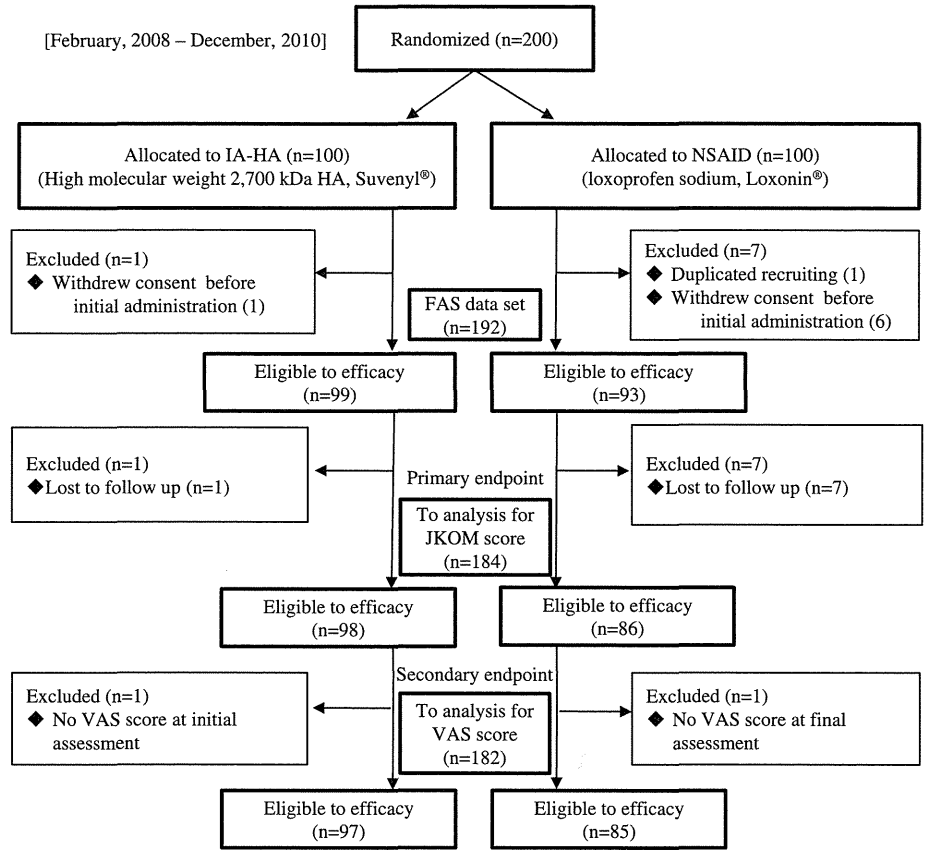


Figure 1 A flow chart of the study. FAS, full analysis set; HA, hyaluronic acid; IA-HA, intra-articular injections of HA; ITT, intention-to-treat; JKOM, Japanese Knee Osteoarthritis Measure (JKOM); NSAID, non-steroidal anti-inflammatory drug; VAS; visual analog scale.

randomly allocated into the NSAID group, and the other half allocated into the IA-HA group. Two patients in the IA-HA group and 14 patients in the NSAID group were excluded; therefore, the remaining 184 patients were included in the analyses of the primary endpoint.

The baseline patient characteristics are shown in Table 1. No significant statistical differences between the baseline characteristics of both groups were found.

Efficacy analyses (primary and secondary endpoints)

For the primary endpoint analysis, the JKOM score of the patients in both the NSAID group and in the IA-HA group was significantly reduced by the treatment ($P<0.001$, Table 2), and the percentage change from baseline in the JKOM score for the two groups was −32.2% and −34.7%, respectively. The difference in the percentage changes of the JKOM score between the two intervention arms (primary endpoint) was −2.5% (95% CI: -14.0 to 9.1%).

In a multiple regression analysis performed taking into consideration the factors considered to stratify the study design, the difference in the primary endpoint between the two intervention arms was also less than 10% (data not shown). These results demonstrate that the IA-HA treatment was non-inferior to the NSAID treatment for the percentage reduction in the clinical symptoms evaluated by the JKOM.

For the secondary endpoint analysis, the pain VAS score of the patients in the NSAID group was significantly

Table 1 Baseline characteristics of the patients in the study

		IA-HA (n = 99)	NSAID (n = 93)
Age (y)	Mean (SD)	68.2 (7.1)	68.5 (7.0)
Gender	Male	27	22
	Female	72	71
BMI	Mean (SD)	23.8 (3.4)	24.4 (3.6)
K/L grade	1	16	15
	2	48	50
	3	35	28
JKOM score (Min:0 - Max:100)	Mean (SD)	33.8 (15.8)	31.6 (14.1)
Pain VAS (Min:0 - Max:100)	Mean (SD)	60.3 (22.4)	55.1 (21.9)

IA, intra-articular; HA, hyaluronic acid; NSAID, non-steroidal anti-inflammatory drug; BMI, body mass index; K/L, Kellgren-Laurence grade; JKOM, Japanese Knee Osteoarthritis Measure; VAS, visual analog scale.

Table 2 Results of the primary endpoint of the study

	JKOM score			% change of JKOM score		
		Mean	SD	P (post vs. pre)	Mean (%)	SD	Difference (%) [IA-HA - NSAID] (95% CI)
IA-HA (n = 98)	Pre-treatment	33.8	15.9	<0.001	−34.7	39.6	−2.5 (−14.0 to 9.1)
	Post-treatment	21.5	14.6				
NSAID (n = 86)	Pre-treatment	32.0	14.0	<0.001	−32.2	39.8	
	Post-treatment	22.0	15.5				

The effect of the treatment of either IA-HA or NSAID for the patients with knee OA evaluated by JKOM score (left) and percentage (%) change of JKOM score (right). A *P* value less than 0.05 was considered to be significant. JKOM, Japanese Knee Osteoarthritis Measure; IA, intra-articular; HA. hyaluronic acid; NSAID, non-steroidal anti-inflammatory drug.

reduced by the treatment ($P<0.001$, Table 3). The percentage change from baseline in the VAS score in the NSAID group was −36.0%. The pain VAS in the IA-HA group was also significantly reduced by the treatment, with a percentage change from baseline in the VAS score of −41.2% ($P<0.001$). The difference in the percentage changes in the pain VAS score between the two intervention arms (secondary endpoint) was −5.2% (95% CI: -23.8 to 13.4%).

Subanalyses

When the patients were divided into two groups (responders or non-responders) by the OMERACT-OARSI response criteria [16], 69.7% (69/99) of the patients in IA-HA group were classified as 'responders', while 62.4% responders were found (58/93) in the NSAID group. Again, there were no significant differences in the frequency of 'responders' between these two groups ($P = 0.283$).

A multiple logistic regression analysis, which was adjusted for age, K/L grade, BMI and the participating medical centers, confirmed the lack of significant differences in the odds ratio of responders between those who received IA-HA treatment and those who received NSAID treatment (odds ratios: 1.47 (95% CI: 0.761 to 2.83)).

We further investigated whether IA-HA is broadly effective from very early (K/L grade of 1) to moderate stages of knee OA (K/L grade of 3) (Table 4). Both IA-HA and NSAID groups significantly reduced the patient-oriented outcome measure evaluated by the JKOM score in the patients with both a K/L grade of 2 and 3. In patients with a K/L grade of 1, IA-HA treatment also reduced the JKOM score, but this reduction was not significant ($P = 0.058$). On the other hand, NSAID treatment of this group significantly reduced the JKOM score ($P = 0.001$).

Safety analyses

During the five weeks of examination, nine of ninety-nine patients (9.1%) in the IA-HA group were withdrawn from the study (one patient's symptoms improved and eight patients were lost to follow-up). Nineteen of ninety-three patients (20.4%) of NSAID group were withdrawn from the study (five patients experienced side effects, four withdrew consent, two patient's symptoms improved, and eight were lost to follow-up). The frequency of the withdrawal rate in the IA-HA group was significantly lower than that in the NSAID group ($P = 0.026$, Table 5).

Serious adverse events, including gastrointestinal (GI) hospitalization, were not observed in both groups during this study. As one patient complained of stiffness in the affected knee after injection, the frequency of adverse events in patients treated with the IA-HA was 1.0%. Ten (symptom related to GI tract disorder, seven; drug allergy, three) of ninety-three patients (10.8%) exhibited adverse events in those treated with the NSAID. The frequency of adverse events in the IA-HA group was significantly lower than that of those in NSAID group ($P = 0.004$, Table 5).

Discussion

This short-term trial clearly demonstrated that both the IA-HA at weekly intervals and daily oral NSAID over five weeks significantly improved both the clinical symptoms evaluated by the patient-oriented outcome measure

Table 3 Results of the secondary endpoint of the study

	Pain VAS			% change of VAS score		
		Mean	SD	P (post vs. pre)	Mean (%)	SD	Difference (%) [IA-HA - NSAID] (95% CI)
IA-HA (n = 97)	Pre-treatment	60.1	22.4	<0.001	−41.2	52.7	−5.2 (−23.8 to 13.4)
	Post-treatment	31.8	24.1				
NSAID (n = 85)	Pre-treatment	55.5	21.8	<0.001	−36.0	73.8	
	Post-treatment	31.9	23.9				

The effect of the treatment of either IA-HA or NSAID for the patients with knee OA evaluated by pain VAS score (left) and percentage (%) change of VAS score (right). A *P* value less than 0.05 was considered to be significant. VAS, visual analog scale; IA, intra-articular; HA, hyaluronic acid; NSAID, non-steroidal anti-inflammatory drug.

Table 4 JKOM score and percentage change of JKOM score by K/L grade subgroup

		JKOM score			% change of JKOM score		
		Mean	SD	*P* (post vs. pre)	Mean (%)	SD	Difference (%) [IA-HA - NSAID] (95% CI)
K/L grade 1							
IA-HA (n = 15)	Pre-treatment	24.8	13.0	0.058	−9.3	78.0	25.7 (−19.3 to 70.7)
	Post-treatment	18.7	12.6				
NSAID (n = 14)	Pre-treatment	35.9	15.4	0.001	−34.9	26.0	
	Post-treatment	23.4	15.7				
K/L grade 2							
IA-HA (n = 48)	Pre-treatment	33.1	14.7	<0.001	−43.8	27.1	−9.2 (−24.4 to 6.0)
	Post-treatment	18.8	12.6				
NSAID (n = 45)	Pre-treatment	30.8	13.9	<0.001	−34.6	44.9	
	Post-treatment	20.4	14.5				
K/L grade 3							
IA-HA (n = 35)	Pre-treatment	38.6	17.1	<0.001	−33.1	23.6	−6.2 (−21.6 to 9.3)
	Post-treatment	26.4	16.9				
NSAID (n = 27)	Pre-treatment	31.9	13.6	0.003	−26.9	37.0	
	Post-treatment	23.7	17.3				

A *P* value less than 0.05 was considered to be significant. JKOM, Japanese Knee Osteoarthritis Measure; IA, intra-articular; HA, hyaluronic acid; NSAID, non-steroidal anti-inflammatory drug; K/L, Kellgren-Laurence grade.

and the pain severity evaluated by a VAS. No significant differences in the symptom-modifying effects were observed during this short period. In addition, the safety of the early phase of IA-HA treatment was superior to that of the NSAID in the patients with knee OA.

HA is a large glycosaminoglycan composed of repeating disaccharides of glucuronic acid and N-acetyl glucosamine that is naturally present in synovial fluid. Several protective properties of HA have been reported including shock absorption, traumatic energy dissipation, protective coating of the articular cartilage surface, and lubrication [19]. Numerous clinical trials, meta-analyses and systematic reviews have indicated its clinical efficacy for knee OA [5,9,10,20]. Based on these previous findings, the OARSI recommendations that were revised in 2010 summarized the effect size (ES) of IA-HA at 0.60 (95% CI; 0.37, 0.83). However, as the ES declined to 0.22 (95% CI; -0.11, 0.54) when only the high-quality trials were selected [5], controversy remains regarding the efficacy of HA in treating knee OA [8]. A recent meta-analysis concluded that the pain reduction by IA-HA is observed later than that of intra-articular corticosteroids [9]. In addition, the effects of IA-HA for knee OA pain continued over six months post-intervention [10]. However, few studies have been conducted to clarify the early effects and safety of IA-HA in comparison to those of NSAIDs. The results of this study clearly indicated that the early efficacy of IA-HA was not inferior in comparison to that of the NSAID.

Table 5 Withdrawal and harmful events during the study

Withdrawn	Completed	Withdrawn	Frequency (%)	*P*
IA-HA (n = 99)	90	9	9.1	0.026
NSAID (n = 93)	74	19	20.4	
Harmful events	**Not occurred**	**Occurred**	**Frequency (%)**	***P***
IA-HA (n = 99)	98	1	1.0	0.004
NSAID (n = 93)	83	10	10.8	

IA, intra-articular; HA, hyaluronic acid; NSAID, non-steroidal anti-inflammatory drug.

A number of HA products with a variety of the molecular weights, ranging from approximately 600 to 6,000 kDa, have been developed as IA-HA for the treatment of OA [8]. The considerable heterogeneity of outcomes between trials may be due in part to differences in HA products [5]. High-molecular-weight HA (>6,000 kDa) is suggested to have greater effects in comparison to lower-molecular-weight HA [8]. On the other hand, the intra-articular injection of high-molecular-weight HA (>6,000 kDa) showed a greater frequency of adverse events, such as pain flares, and hot and swollen knees, which typically occurred 24 to 72 hours after injection [21]. There were no cases of painful, hot or swollen knees during the study.

The molecular mechanisms underlying the efficacy of IA-HA for OA remain unclear. OA is frequently associated with the signs and symptoms of inflammation, including joint pain, swelling and stiffness leading to significant functional impairment and disability [2]. Synovitis plays an important role in inducing the pain, swelling and stiffness in OA [22], and the severity of synovitis is well correlated with the JKOM score of the patients with knee OA [23]. It has recently been reported that HA inhibits the activities of matrix

metalloproteinases and aggrecanases which are, at least in part, involved in OA cartilage degradation as a result of their induction by proinflammatory cytokines, such as interleukin (IL)-1 [19,24-26]. Therefore, HA is speculated to modify the structural damage of joints and the rate of OA progression in addition to the symptom-modifying effect [27], although further studies are required.

In this trial, the early efficacy of IA-HA was compared with that of NSAID for the treatment of knee OA. NSAIDs have also been proven to be an effective conservative treatment for knee OA [5]. However, a high incidence of serious GI tract adverse events associated with the use of oral NSAIDs was also demonstrated in a population-based cohort study of older patients [28]. In addition, the hospitalization due to GI tract side effects in patients receiving non-selective NSAIDs was twice as high as that in those given the cyclooxygenase (Cox)-2 selective agent, celecoxib, or a non-selective NSAID together with a PPI [28]. Although a PPI was not routinely used in addition to the NSAID (loxoprofen sodium) in this study, no serious GI events were noted.

Since chronic kidney disease (CKD), which is similar to knee OA, is also a prevalent disease especially in older populations, knee OA patients with CKD may have a different risk profile and treatment response than knee OA patients without CKD. However, as patients with renal disorders were excluded in the present study, as described in the Methods section, whether the presence of CKD has any effect on the efficacy and safety of either the IA-HA or NASIDs remains unclear.

The efficacy of IA-HA for knee OA has been debated for over a decade. Although it has been systemically evaluated in meta-analyses, most previous studies have focused on comparing the findings with either placebo or intra-articular corticosteroids [9,10,29]. No previous studies have undertaken a meta-analysis with NSAIDs, which is one of the most efficacious and widely used treatments for knee OA [5]. The present study clearly shows that IA-HA is as effective as continuous NSAID use at five weeks of treatment, and, in addition, it showed a more favorable safety profile of IA-HA over NSAIDs for knee OA. The present study suggests that future randomized trials should thus be carried out with a longer duration of follow-up and larger samples, in order to identify optimal knee OA treatment alternatives. Furthermore, it would also be interesting to evaluate whether any synergistic effect of these two combined treatments exists when they are combined.

The current study does have some limitations. First, this investigation was an open-label randomized trial and not a double-blind controlled trial. Therefore, the design may have introduced certain bias into the results. Second, the trial's size was calculated to have sufficient power to exclude a 10% between-group percentage change of JKOM score, which can be debated. This margin was supported by our pilot study, as described previously. Third, in subgroup analysis for the patients with a K/L grade of 1, IA-HA treatment reduced the JKOM score. However, this reduction was not statistically significant ($P = 0.058$). Although the reason for this is unclear, the interpretation of the result was limited by the small number of patients (n = 15) and, therefore, it may be one of the limitations. Even though some subjects had a K/L grade of 1, some have an increased risk for rapid progression of the disease [30]. Unfortunately, we cannot predict radiographically who is at risk for progression [4].

Conclusions

The early efficacy of IA-HA is suggested to be not inferior to that of a NSAID, and the safety of the early phase of IA-HA is superior to that of a NSAID for patients with knee OA.

Abbreviations

ACR: American College of Rheumatology; BMI: body mass index; CKD: chronic kidney disease; FAS: full analysis set; GI: gastrointestinal; HA: hyaluronic acid; IA-HA: intra-articular injections of HA; JKOM: Japanese Knee Osteoarthritis Measure; K/L: Kellgren-Lawrence; NSAIDs: non-steroidal anti-inflammatory drugs; OA: osteoarthritis; OARSI: Osteoarthritis Research Society International; OMERACT: Outcome Measures in Rheumatology Clinical Trials; PPI: proton pump inhibitor; SF-36: Medical Outcomes Study 36-Item Short-Form Health Survey; VAS: visual analog scale; WOMAC: Western Ontario and McMaster Universities Arthritis Index.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

As principal investigators of this study, all authors of this study had full access to all data, and take responsibility for their integrity and the accuracy of their analysis. TN, KS, KH, HKu and KK participated in the study design. HKu and KK supervised the study. MI, TN, KS, HKi, SS, GO, TY, YU, JC, MK and HKu collected the data. MI, YI and KH analyzed the data. YI and KH provided statistical expertise. MI and KH drafted the manuscript, and the manuscript was revised for content by MI, TN, KS, KH, HKi, SS, GO, TY, YU, JC, YI, MK, HKu and KK. All authors read and approved the final manuscript.

Acknowledgements

The authors thank all investigators involved in this study at the following medical facilities listed in alphabetical order: Gifu University Hospital (Gifu, Japan), Heisei Memorial Hospital (Shimane, Japan), Izumo Citizens Hospital (Shimane, Japan), Izumo General Medical Center (Shimane, Japan), Juntendo University Medical Center (Tokyo, Japan), Juntendo University Nerima Hospital (Tokyo, Japan), Ken-ai Memorial Hospital (Fukuoka, Japan), Kinki University Sakai Hospital (Osaka, Japan), Kitanarashino Hanawa Hospital (Chiba, Japan), Nara Hospital Kinki University Faculty of Medicine (Ikoma, Japan), Niigata University Hospital (Niigata, Japan), Niigata Rehabilitation Hospital (Niigata, Japan), Nishimino-Kosei Hospital (Gifu, Japan), Nishi-Niigata Chuo National Hospital (Niigata, Japan), Occupational and Environmental Health University Hospital (Kitakyusyu, Japan), Shimane University Hospital (Shimane, Japan), Sapporo Daiichi Hospital (Hokkaido, Japan), Sapporo Medical University Hospital (Hokkaido, Japan), Tokyo Women's Medical University Medical Center East (Tokyo, Japan), and Tama Nannbu Chiiki Hospital (Tokyo, Japan).

The authors also express sincere thanks to Dr. Fumie Tokuda and Dr. Chizuko Wakabayashi at the coordinating data center (Gunma University, Gunma, Japan) for their help with data management and data analysis. We also express our appreciation to Dr. Yoshihiko Yamada, Dr. A. Robin Poole and Dr. Joe Matsuoka for their helpful comments and suggestions.

Author details
[1]Department of Medicine for Orthopaedics and Motor Organ, Juntendo University Graduate School of Medicine, 2-1-1, Hongo, Bunkyo-ku, Tokyo 113-8421, Japan. [2]Department of Orthopaedic Surgery, University of Occupational and Environmental Health, 1-1, Iseigaoka, Yahatanishi-ku, Kitakyushu 807-8555, Japan. [3]Department of Orthopaedic Surgery, Gifu University, School of Medicine, 1-1, Yanagido, Gifu City 501-1193, Japan. [4]Department of Basic Medical Sciences, School of Health Sciences, Faculty of Medicine, Gunma University, 3-39-22, Showa-machi, Maebashi, Gunma 371-8511, Japan. [5]Department of Orthopaedic Surgery, Sakai Hospital Kinki University Faculty of Medicine, 2-7-1, Harayamadai, Minami-ku, Sakai-city 590-0132, Osaka, Japan. [6]Department of Orthopaedic Surgery and Rheumatology, Nara Hospital Kinki University Faculty of Medicine, 1481-1, Otodacho, Ikoma, Nara 630-0293, Japan. [7]Department of Health and Sports, Niigata University of Health and Welfare, 1398, Shimamicho, Niigata Kita-ku, Niigata 950-3198, Japan. [8]Department of Orthopaedic Surgery, Sapporo Medical University School of Medicine, South-1, West-16, Chuo-ku, Sapporo 060-8543, Japan. [9]Department of Orthopaedics, Shimane University, School of Medicine, 89-1, Enya-cho, Izumo-shi, Shimane 693-8501, Japan. [10]Department of Orthopaedic Surgery, Tokyo Women's Medical University, Medical Center East, 2-1-10, Nishiogu, Arakawa, Tokyo 116-8567, Japan.

Received: 2 June 2013 Accepted: 8 January 2014
Published: 21 January 2014

References

1. Dieppe PA, Lohmander LS: **Pathogenesis and management of pain in osteoarthritis.** *Lancet* 2005, **365**:965–973.
2. Felson DT: **Clinical practice, Osteoarthritis of the knee.** *N Engl J Med* 2006, **354**:841–848.
3. Carr AJ, Robertsson O, Graves S, Price AJ, Arden NK, Judge A, Beard DJ: **Knee replacement.** *Lancet* 2012, **379**:1331–1340.
4. Ishijima M, Watari T, Naito K, Kaneko H, Futami I, Yoshimura-Ishida K, Tomonaga A, Yamaguchi H, Yamamoto T, Nagaoka I, Kurosawa H, Poole RA, Kaneko K: **Relationships between biomarkers of cartilage, bone, synovial metabolism and knee pain provide insights into the origins of pain in early knee osteoarthritis.** *Arthritis Res Ther* 2011, **13**:R22.
5. Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, Bierma-Zeinstra S, Brandt KD, Croft P, Doherty M, Dougados M, Hochberg M, Hunter DJ, Kwoh K, Lohmander LS, Tugwell P: **OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009.** *Osteoarthritis Cartilage* 2010, **18**:476–499.
6. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, Bierma-Zeinstra S, Brandt KD, Croft P, Doherty M, Dougados M, Hochberg M, Hunter DJ, Kwoh K, Lohmander LS, Tugwell P: **OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence.** *Osteoarthritis Cartilage* 2007, **15**:981–1000.
7. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, Bierma-Zeinstra S, Brandt KD, Croft P, Doherty M, Dougados M, Hochberg M, Hunter DJ, Kwoh K, Lohmander LS, Tugwell P: **OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines.** *Osteoarthritis Cartilage* 2008, **16**:137–162.
8. Lo GH, LaValley M, McAlindon T, Felson DT: **Intra-articular hyaluronic acid in treatment of knee osteoarthritis: a meta-analysis.** *JAMA* 2003, **290**:3115–3121.
9. Bannuru RR, Natov NS, Obadan IE, Price LL, Schmid CH, McAlindon TE: **Therapeutic trajectory of hyaluronic acid versus corticosteroids in the treatment of knee osteoarthritis: a systematic review and meta-analysis.** *Arthritis Rheum* 2009, **61**:1704–1711.
10. Bannuru RR, Natov NS, Dasi UR, Schmid CH, McAlindon TE: **Therapeutic trajectory following intra-articular hyaluronic acid injection in knee osteoarthritis–meta-analysis.** *Osteoarthritis Cartilage* 2011, **19**:611–619.
11. Rutjes AW, Juni P, da Costa BR, Trelle S, Nuesch E, Reichenbach S: **Viscosupplementation for osteoarthritis of the knee: a systematic review and meta-analysis.** *Ann Intern Med* 2012, **157**:180–191.
12. **UMIN Clinical Trials Registry.** [http://www.umin.ac.jp/ctr/index.htm]
13. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, Christy W, Cooke TD, Greenwald R, Hochberg M, Howell D, Kaplan D, Koopman W, Longley S III, Mankin H, McShane DJ, Medsger T Jr, Meenan R, Mikkleson W, Moskowitz R, Murphy W, Rothschild B, Segal M, Sokoloff L, Wolfe F: **Development of criteria for the classification and reporting of osteoarthritis, Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association.** *Arthritis Rheum* 1986, **29**:1039–1049.
14. Kellgren JH, Lawrence JS: **Radiological assessment of osteo-arthrosis.** *Ann Rheum Dis* 1957, **16**:494–502.
15. Akai M, Doi T, Fujino K, Iwaya T, Kurosawa H, Nasu T: **An outcome measure for Japanese people with knee osteoarthritis.** *J Rheumatol* 2005, **32**:1524–1532.
16. Pham T, van der Heijde D, Altman RD, Anderson JJ, Bellamy N, Hochberg M, Simon L, Strand V, Woodworth T, Dougados M: **OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited.** *Osteoarthritis Cartilage* 2004, **12**:389–399.
17. Kawasaki T, Kurosawa H, Ikeda H, Takazawa Y, Ishijima M, Kubota M, Kajihara H, Maruyama Y, Kim SG, Kanazawa H, Doi T: **Therapeutic home exercise versus intraarticular hyaluronate injection for osteoarthritis of the knee: 6-month prospective randomized open-labeled trial.** *J Orthop Sci* 2009, **14**:182–191.
18. Kubota M, Ishijima M, Kurosawa H, Liu L, Ikeda H, Osawa A, Takazawa Y, Kawasaki T, Saita Y, Kimura Y, Kaneko K: **A longitudinal study of the relationship between the status of bone marrow abnormalities and progression of knee osteoarthritis.** *J Orthop Sci* 2010, **15**:641–646.
19. Julovi SM, Ito H, Nishitani K, Jackson CJ, Nakamura T: **Hyaluronan inhibits matrix metalloproteinase-13 in human arthritic chondrocytes via CD44 and P38.** *J Orthop Res* 2011, **29**:258–264.
20. Conrozier T, Balblanc JC, Richette P, Mulleman D, Maillet B, Henrotin Y, Rannou F, Piroth C, Hilliquin P, Mathieu P, Walliser-Lohse A, Rousselot I, Plattner V, Maillefert JF, Vignon E, Chevalier X: **Early effect of hyaluronic acid intra-articular injections on serum and urine biomarkers in patients with knee osteoarthritis: an open-label observational prospective study.** *J Orthop Res* 2012, **30**:679–685.
21. Reichenbach S, Blank S, Rutjes AW, Shang A, King EA, Dieppe PA, Juni P, Trelle S: **Hylan versus hyaluronic acid for osteoarthritis of the knee: a systematic review and meta-analysis.** *Arthritis Rheum* 2007, **57**:1410–1418.
22. Goldring MB, Goldring SR: **Osteoarthritis.** *J Cell Physiol* 2007, **213**:626–634.
23. Liu L, Ishijima M, Futami I, Kaneko H, Kubota M, Kawasaki T, Matsumoto T, Kurihara H, Ning L, Xu Z, Ikeda H, Takazawa Y, Saita Y, Kimura Y, Xu S, Kaneko K, Kurosawa H: **Correlation between synovitis detected on enhanced-magnetic resonance imaging and a histological analysis with a patient-oriented outcome measure for Japanese patients with end-stage knee osteoarthritis receiving joint replacement surgery.** *Clin Rheumatol* 2010, **29**:1185–1190.
24. Yatabe T, Mochizuki S, Takizawa M, Chijiiwa M, Okada A, Kimura T, Fujita Y, Matsumoto H, Toyama Y, Okada Y: **Hyaluronan inhibits expression of ADAMTS4 (aggrecanase-1) in human osteoarthritic chondrocytes.** *Ann Rheum Dis* 2009, **68**:1051–1058.
25. Plaas A, Li J, Riesco J, Das R, Sandy JD, Harrison A: **Intraarticular injection of hyaluronan prevents cartilage erosion, periarticular fibrosis and mechanical allodynia and normalizes stance time in murine knee osteoarthritis.** *Arthritis Res Ther* 2011, **13**:R46.
26. Ning L, Ishijima M, Kaneko H, Kurihara H, Arikawa-Hirasawa E, Kubota M, Liu L, Xu Z, Futami I, Yusup A, Miyahara K, Xu S, Kaneko K, Kurosawa H: **Correlations between both the expression levels of inflammatory mediators and growth factor in medial perimeniscal synovial tissue and the severity of medial knee osteoarthritis.** *Int Orthop* 2011, **35**:831–838.
27. Goldberg VM, Buckwalter JA: **Hyaluronans in the treatment of osteoarthritis of the knee: evidence for disease-modifying activity.** *Osteoarthritis Cartilage* 2005, **13**:216–224.
28. Rahme E, Barkun A, Nedjar H, Gaugris S, Watson D: **Hospitalizations for upper and lower GI events associated with traditional NSAIDs and acetaminophen among the elderly in Quebec, Canada.** *Am J Gastroenterol* 2008, **103**:872–882.
29. McAlindon TE, Bannuru RR: **Osteoarthritis: Is viscosupplementation really so unsafe for knee OA?** *Nat Rev Rheumatol* 2012, **8**:635–636.
30. Hart DJ, Spector TD: **Kellgren & Lawrence grade 1 osteophytes in the knee–doubtful or definite?** *Osteoarthritis Cartilage* 2003, **11**:149–150.

doi:10.1186/ar4446
Cite this article as: Ishijima *et al.*: Intra-articular hyaluronic acid injection versus oral non-steroidal anti-inflammatory drug for the treatment of knee osteoarthritis: a multi-center, randomized, open-label, non-inferiority trial. *Arthritis Research & Therapy* 2014 **16**:R18.