Evaluation of group homogeneity during acupuncture stimulation in fMRI studies

Journal:	*Journal of Magnetic Resonance Imaging*
Manuscript ID:	JMRI-09-0777.R1
Wiley - Manuscript type:	Original Research
Classification:	Functional neuro-imaging < Neuro-imaging < Clinical Science
Manuscript Keywords:	homogeneity, individual differences, outlier, mean distances, MDS, acupuncture

Evaluation of group homogeneity during acupuncture stimulation in fMRI studies

Jinbo Sun, PhD[1], Wei Qin, PhD[1], Minghao Dong, PhD[1],Kai Yuan, PhD[1], Jixin Liu, PhD[1], Peng Liu, PhD[1], Yi Zhang, PhD[1], Karen M. von Deneen, MD[3], Jie Tian, PhD[1,2]

[1]*Life Sciences Research Center, School of Life Sciences and Technology, Xidian University, Xi'an, Shaanxi 710071, China*

[2]*Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China*

[3]*Departments of Psychiatry and Neuroscience, McKnight Brain Institute, University of Florida, Gainesville, FL 32610, USA*

Corresponding author:

Name: Jie Tian

Address: Institute of Automation, Chinese Academy of Science P.O. Box 2728, Beijing, 100190, China

Tel: +86-010-82618465; **FAX:** 010-62527995; **E-mail:** tian@ieee.org

Grant Support

This paper is supported by Changjiang Scholars and Innovative Research Team in University（PCSIRT）under Grant No. IRT0645, the Chair Professors of Cheung Kong Scholars Program of Ministry of Education of China, CAS Hundred Talents Program, the National Natural Science Foundation of China under Grant Nos. 30970774, 60901064, 30873462, 90209008, 30870685, 30672690, 30600151, 60532050, 60621001, the Joint Research Fund for Overseas Chinese Young Scholars under Grant No. 30528027, the Beijing Natural Science Fund under Grant No. 4071003，the

Knowledge Innovation Program of the Chinese Academy of Sciences under Grant No. KGCX2-YW-129, the Project for the National Key Basic Research and Development Program (973) under Grant No. 2006CB705700, and 863 Program under Grant No. 2008AA01

ABSTRACT

Purpose： To evaluate the extent of individual differences of fMRI signal changes induced by acupuncture stimulation.

Materials and Methods： Acupuncture at ST36 and checkerboard stimulation was applied to 16 subjects. We calculated the mean distance using beta values in a generalized linear model (GLM) analysis and employed it to study the group homogeneity by detecting the outliers.

Results： A more significant individual difference was presented in acupuncture stimulation compared with visual stimulation through evaluation of the mean distance. From the group results, we found that the activation were more significant in the homogeneous group results. Combining the behavior and fMRI results, there was no direct correlation between deqi index and mean distance in acupuncture stimulation. The deqi index of the outlier was in the normal range and did not significantly differing from others.

Conclusion： Traditional group results without removing outliers were not sensitive enough to detect the real acupuncture effect. We suggest that individual difference should be taken into consideration for future acupuncture studies. Also, group analysis paralleled with individual analysis is critical for full understanding of acupuncture effects.

Key Words： homogeneity; individual difference; outlier; mean distances; MDS; acupuncture

INTRODUCTION

In the past several years, fMRI has been used to explore the neurobiological mechanism of acupuncture (1-8). Most of these studies employed the BLOCK paradigm and the statistical inference method of traditional generalized linear models (GLM), which derived from modern cognitive psychology studies (9).

Recently, test–retest reliability of fMRI blood oxygenation level dependent (BOLD) signal changes has been investigated in multiple tasks, e.g. finger tapping, visual simulation, auditory oddball, working memory and learning related tasks (10-14). More specifically, the reliable pattern that fMRI signal increases in the primary motor cortex, supplementary motor area (SMA), and cerebellum (14-16) was found in finger tapping task, despite of variant analysis methods adopted. However, a previous acupuncture study applied test-retest reliability analysis of BOLD signal changes to compare the results between the finger tapping task and electro-acupuncture stimulation, revealing that fMRI signal changes evoked by electro-acupuncture stimulation were significantly more variable than those from the finger tapping task (17). We also compared the several acupuncture studies (4, 6-8, 18-21) and found disagreement in BOLD patterns. Moreover, we found discrepancies between individual activation and group results within the same research. Conclusively, individual difference in acupuncture could normally be interpreted as two distinct modalities. Individual activation, which was commonly identified in large proportion of subjects, was not presented in group results, and *vice versa*. E.g. Napadow et al. showed SI activation in 75% subjects without activation at group level (8). Hui et al.

found significant bilateral signal decreases in nucleus accumbens within 71% subjects while only a contralateral response at group level (19). On the other hand, Li et al. showed activation in several regions of occipital lobe at group level while only 39% subjects had significant activations (20). Cho et al. presented that among the 16 volunteers, only 8 volunteers showed cortical activation during acupuncture stimulation (18). Considering the divergence between individual response patterns and group response patterns to acupuncture stimulation, we suggested that the statistical methods should be used prudently in acupuncture studies.

Random effects analysis, defined as the gold standard for group analysis in neuroimaging studies (22), is used to locate consistently activated regions by incorporating intra- and inter-subjects variability (9, 23, 24). Random effects analysis relies on the hypothesis that subjects are independently identically distributed and drawn from the same homogeneous population (24). In this frame, subjects are expected to have the same common mean which is close to the true mean of the population (25), i.e. the degree of individual difference (DID) should be limited to a certain range. The evaluation of group homogeneity drew attention in some researches, e.g. the mental calculation task (26). The likelihood also exists in acupuncture research. Therefore, it is necessary to evaluate DID in acupuncture research due to prominent individual difference across subjects. However, to our knowledge, few studies evaluated the homogeneity of fMRI signal changes evoked by acupuncture. The current study, for the first time in acupuncture research, presents a novel measure to evaluate the similarity among subjects by calculating the mean distance using beta values of GLM analysis. Furthermore, we study the group homogeneity and detect outliers with the help of this method. To verify our method, we compared the group homogeneities of the acupuncture paradigm and checkerboard stimulation task, which were of dissimilar DID and similar DID. Finally, we furthered our results by comparing group results with and without the detected outlier in the acupuncture group.

MATERIALS AND METHODS

To reduce non-interested inter-subject variability, participants were recruited from a group of sixteen college students (8 males and 8 females; 21.9±1.2 years). All subjects were right-handed with normal or corrected-to-normal vision. The subjects were acupuncture naive, have no history of major medical illness, head trauma, or neuropsychiatric disorders, used no prescription medications within the last month, and have no contraindication to exposure to a high magnetic field. All subjects were given written, informed consent after the experimental procedure had been fully explained, and all research procedures were approved by a local Subcommittee on Human Studies and were conducted in accordance with the Declaration of Helsinki.

Subjects underwent two experimental sessions: a checkerboard session and an acupuncture session (ACUP). The ACU sessions were conducted three days after visual session.

A visual fMRI experiment was adopted (Fig 1B) to assist the manifestation of individual difference in acupuncture. A checkerboard pattern flashing at 8 Hz was used as the visual stimulus. The whole 5 min run had two cycles consisting of a 1 min visual stimulus and a 1 min of resting scan, preceded by a 1 min resting period (see Bai et al. 2009 for details).

In ACUP, we employed a non-repeated event-related (NRER) design introduced by Qin et al. 2008 (Fig 1A). An acupuncture needle was inserted at ST36 on the right leg, and after resting for 1 min, the needle was manipulated for 1.5 minutes. The needle was kept inserted at the acupoint for 12.5 minutes. ST36 is one of the most frequently used acupoints for analgesia and has been studied in both animal and human acupuncture researches (4, 27). The stimulation was administered by a balanced 'tonifying and reducing' technique using a sterile disposable 38 gauge stainless steel acupuncture needle (0.2 mm × 40 mm). The needle was inserted perpendicularly to a depth of 2-3 cm, and rotated manually clockwise and counterclockwise for 1.5 min with frequency of 1 Hz. The entire acupuncture procedure was conducted by the same licensed acupuncturist.

At the end of the acupuncture run, the subjects were asked to quantify their sensations using a 10-point 11-elicitor visual analogue scale (VAS) including aching, pressure, soreness, heaviness, fullness, warmth, coolness, numbness, tingling, dull or sharp pain they felt during the scan (2). All these sensations are used to rate their *deqi* experience which is crucial to acupuncture therapeutic effect.

To quantitatively summarize the full multivariate breadth and depth of deqi sensation in each individual, we employed the VAS index, defined as a weighted average of all sensations using an exponential smoothing (28). This index is convenient to devise a single value. In a further analysis, we investigated the relation between the individual deqi composite (VAS index) and the mean distance.

Functional images were acquired on a 3T GE Signa scanner using a standard whole head coil (LX platform, gradients 40mT/m, 150T/m/s, GE Medical Systems, Milwaukee, WI). A custom-built head holder was used to prevent head movements. Thirty-two axial slices (FOV = 240mm×240mm, matrix = 64×64, thickness = 5 mm), parallel to the AC-PC plane and covering the whole brain, were obtained using a T2*-weighted single-shot, gradient-recalled echo planar imaging (EPI) sequence (TR = 1,500 ms, TE = 30 ms, flip angle = 90°). The scan covered the entire brain including the cerebellum and brainstem. Prior to the functional run, high-resolution structural information on each subject was acquired using 3D MRI sequences with a voxel size of 1mm^3 for anatomical localization (TR = 2.7s, TE = 3.39ms, matrix = 256×256, FOV = 256mm×256mm, flip angle = 7°, in-plane resolution = 1mm×1mm, slice thickness = 1mm).

Pre-processing and statistical analysis were performed using the Statistical Parametric Mapping software (SPM5, http://fil.ion.ac.uk/spm). The first 5 time points were discarded to obtain a stable state. In acupuncture research, time-varied characteristic of acupuncture effect is incompatible with estimation of statistical parameters in the GLM analysis (1, 2). Thus, only data before and during manipulation for ACUP runs were selected (total of 100 time points). Based on our hypothesis, we believed that individual difference during acupuncture manipulation would be relatively smaller than that during rest after acupuncture manipulation. In addition, smaller motion artifacts would be introduced shortly. The selected images were realigned to the first volume acquired in the scan session. All subjects satisfied our excessive motion threshold of less than 1mm spatial displacement in any direction. The images for each run were then normalized to the range of [1, 1000] to compensate for any intensity variations across runs. Subsequently, the images were normalized to the standard EPI template and re-sampled to a voxel size of 3 x 3 x 3 mm^3, and then smoothed spatially using a 6 mm full-width-at-half maximum (FWHM) isotropic Gaussian kernel to decrease spatial noise.

For each subject in both sessions, the fixed effect analysis was applied based on the GLM for the first step. The estimated parameter of each voxel was extracted to draw a Beta-map. The similarity degree was then calculated using the Beta-maps generated above. If a_i and b_i respectively represent the beta value in corresponding positions, then the Euclidean distance for subject A and subject B is determined by

$$D_{ab} = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2},$$

where n is the number of voxels in the whole brain. The smaller Euclidean distance between subjects A and B indicates larger similarity between them. Both the ACUP stimulation and checkerboard task, a symmetrical matrix (element 0 in main diagonal) was constructed from the Euclidean distances between subjects in each group respectively. This matrix is similar to the distance matrix derived from the RV coefficient, but it is more convenient and efficient (26).

Multidimensional scaling (MDS) was used to provide a visual representation of the distance matrix. MDS ploted the subjects as points on a map so that the distance on this map were similar to the original distance. It is a simple visual representation tool that shows whether there is a specific pattern to the distance between subjects, such as the presence of outliers or subgroups (26). In this study, we plotted the MDS maps for 1) ACUP stimulation with the outliers, 2) ACUP stimulation without the outliers, and 3) checkerboard task. In the results, maps 2) and 3) showed the different group homogeneity in the different tasks.

To evaluate the DID and detect the possible outliers, appropriate measurement and statistical tests should be incorporated. First, the mean of each row of the distance matrix was calculated. Each mean which represented the average distance between the subject and the rest in the group was defined as DID. Then, the mean was used to judge whether this subject was an outlier. If the group is homogeneous, these mean values will be approximately equal. The value (or values) would be larger than the rest if there is an outlier or several. Many statistical tests have been developed to detect outliers, such as Grubb's or Dixon tests, Cook distances (29), frequency distribution, Chauvenet's criterion, Q test and BOXPOLT. In this study, we selected the well known BOXPOLT to determine "extreme" outliers. BOXPOLT is useful in displaying difference between populations without making any assumptions on the underlying statistical distribution. The space between different parts of the box indicates the degree of dispersion (spread) and skewness in data.

Finally, second level group analysis was performed using a random effect model with and without the suspected the outlier.

RESULTS

Fig 2A/2B showed 2D MDS representation of distance between subjects in the checkerboard task and ACUP stimulation (without a possible outlier). The drawing illustrated that subjects in the ACUP stimulation showed a more scattered distribution, indicating a more significant individual difference compared to the checkerboard task. The mean distances of both tasks were displayed in Fig 2C. The mean distance of ACUP stimulation was significantly larger than that of checkerboard task ($P < 0.05$). A 2D MDS plot of the ACUP stimulation before detecting the outlier was generated (figure not shown). In the outlier detection step, the subject which was distant from others was confirmed to be an outlier (see Fig 3A and Fig 3B for more detail).

We detected the outlier in the checkerboard task and ACUP stimulation respectively. As depicted in Fig 2A, the checkerboard task was homogeneous and no outliers were found. The mean distances of each subject in the ACUP stimulation were shown in descending order in Fig 3A. Subject 5 had larger mean distance than the others. The BOXPLOT of these mean distances was presented in Fig 3B, demonstrating that subject 5 was an outlier at this statistical benchmark.

Fig 3C showed random effect analysis of two different acupuncture groups: the entire group (second row) and the homogeneous group (first row) from which subject 5 (the outlier) was excluded. Results were thresholded at a level corresponding to a *P*-value of 0.005 (uncorrected). In the homogeneous group presented in Fig 3C, significant activation was found in the right orbital part of the prefrontal gyrus, the right medial part of the superior frontal gyrus, the right triangular part of the inferior frontal gyrus, the right opercular part of the inferior frontal gyrus and the right anterior insula. These regions were also activated in the entire group. However, these activations were less significant compared with the homogeneous group. It is noteworthy that the right orbital part of the inferior frontal gyrus, the left anterior insula, brainstem and the right thalamus were only activated in the homogeneous group.

The Pearson correlation coefficient between the VAS index and the mean distance of each subject was calculated and the regression plots were shown in Fig 4. Results indicated that the correlation between the VAS index and the mean distance did not exceed the significant level (r=0.20, P=0.47). Furthermore, the deqi degree of subject 5 (the outlier) was in the normal range and was not significantly different from others.

DISCUSSION

In this study, we presented a novel measure to evaluate the group homogeneity straightforwardly for the first time in acupuncture research. Using this measure, we evaluated the DID and detected outliers of acupuncture stimulation. The more sensitive and meaningful results of the homogeneous group were concluded to be more valuable in detecting the real acupuncture effect. Interestingly, the deqi degree of subject 5 (the outlier) was not significantly different from others, indicating that the outliers in this paper were not caused by deqi sensation. On the other hand, deqi sensation is an essential factor to determine whether acupuncture takes effect in clinical use, we suggested the outlier should not be erroneously ignored. Therefore, group analysis in conjunction with individual analysis is likely to be critical for a better objective assessment. Moreover, the general applicability of this approach should be prevailingly extended to other fMRI researches regarding to individual difference issues, e.g. memory studies.

Test-retest analysis confirmed reliable fMRI BOLD response patterns in basic cognitive tasks, such as finger tapping, visual simulation, auditory oddball, working memory and learning indicating that individual difference was not significant across subjects (10-14). However, several acupuncture studies inferred that the influence of individual difference should be taken into consideration (17). This study detected the differences of DID between the visual and the acupuncture stimulation. Fig 2A and Fig 2B showed subjects under visual stimulation had better agglomeration compared with those under acupuncture stimulation. Fig 2C showed subjects under visual condition has smaller mean distance than that under acupuncture condition. Both evidence indicated that the homogeneity among the ACUP subjects was lower compared with that of the checkerboard task. This representation might be due to the unique characteristic of acupuncture. As Traditional Chinese Medicine (TCM) theory implies, subjects acts differently to acupuncture stimulation (30, 31), which could explain the individual difference in our results.

Fig 3A and Fig 3B clearly presented that the subject 5 was distant from the others and thus considered to be the outlier. (Subject 5 in Fig 3A and Fig 2B are not the same one, since in Fig 2B the outlier was excluded.) The differences between the two rows in Fig 3C were significant ($P < 0.05$). In group analysis, right opercular part of the inferior frontal gyrus and right anterior insula were found to be greater activated without the outlier, namely the homogeneous group, than with the outlier. The right orbital part of the inferior frontal gyrus, left anterior insula, brainstem, right thalamus and the secondary somatosensory cortex (SII) were activated only in the homogeneous group. All these activated differences suggest that the results obtained from all subjects were less sensitive than those in the homogeneous group. Activation of the insula has been reported in previous acupuncture studies (32-34). It was consistent with electrophysiological studies and clinical investigations (35) showing the insula's involvement in emotional processing (fear, uneasiness, and so forth) and ascending visceral symptoms (36). This also coincided with another observation, which suggested that acupuncture modulated the CEN and DMN networks anchored by the anterior insula (1). The involvement of the insula during acupuncture stimulation is also consistent with TCM's viewpoint that the acupuncture may mediate physical disorders through the insular visceral feedback pathway (1, 2, 27). Activation in the brainstem in the homogeneous group was evident. Several acupuncture studies have already emphasized the importance brainstem in the endogenous monoaminergic and opioidergic systems (1, 4). The activation of SII and the thalamus was expected under somatosensory stimulation at somatic points (37), and the thalamus is also part of the reported neuromatrix of pain (38, 39). In conclusion, the group results became more sensitive when we removed the outlier and the results had their biological significance and relevance to acupuncture.

Individual difference is a common issue in acupuncture research. Some methods were proposed to prevent the individual difference from corrupting results. Several researchers utilized the fixed-effect model to examine the group results (7, 8).

Nevertheless, the fixed effect model cannot infer the results to the parent population. Several groups proposed larger group size and multiple sessions should be taken to obtain more statistically powerful and reliable results (8, 17). However, this is not always the most practical and economical strategy. In addition, researchers averaged the BOLD signal oscillation across all the subjects regardless of their response patterns (6). This was suspicious because no mathematical evidence showed the response patterns of all the samples follow the hypothesis of the group analysis. The other researchers empirically excluded the subject, which show no activation in the individual results (21). We suggested that the method was not rational due to the lack of quantitative measurements. Overall, the existing strategies were unsuitable to solve the problem of individual variability. The method introduced in the present study straightforwardly evaluated the individual differences, easily determined the homogeneity of the group and effectively located the outliers. Moreover, the general applicability of this approach should be prevailingly extended to other fMRI researches regarding to individual difference issues, e.g. memory studies.

Considering deqi is associated with a remedial mechanism of the acupuncture, one question should be raised, what the relationship between individual difference and deqi sensation, i.e. were there different factors in VAS and deqi sensation between subject5 and the other subjects? We analyzed the VAS index and mean distance, finding that there was no direct correlation between the parameters. It indicated that the mean distance is irrelevant with deqi sensation. Therefore, we would like to argue that the outliers in this paper were not determined by deqi sensation. Meanwhile, It is worthwhile to note that the deqi degree of subject 5 (the outlier) was not significantly different from others. Considering deqi sensation has its own clinical effect when judging whether acupuncture takes effect or not, thus, the outlier cannot be erroneously ignored. Therefore, group analysis in conjunction with individual analysis is likely to be critical for a better objective assessment. However, to our knowledge, what exact factors cause the appearance of outlier remains to be answered and need quite a lot of efforts. To understand the acupuncture effect scientifically, individual information such as behavioral data and biochemical indexes should be added to future acupuncture studies.

In conclusion, we focused on individual difference under acupuncture stimulation, and presented a novel method to quantify the issue. We suggest that group analysis in conjunction with individual analysis should be considered in future acupuncture studies.

REFERENCE

1. Bai L, Qin W, Tian J, et al. Acupuncture modulates spontaneous activities in the anticorrelated resting brain networks. Brain Research 2009;1279:37-49.
2. Bai L, Qin W, Tian J, et al. Time-varied characteristics of acupuncture effects in fMRI studies. Human Brain Mapping 2009;30(11):3445-3460.
3. Qin W, Tian J, Bai L, et al. FMRI connectivity analysis of acupuncture effects on an amygdala-associated brain network. Molecular Pain 2008;4(1):55-71.
4. Hui KKS, Liu J, Marina O, et al. The integrated response of the human cerebro-cerebellar and limbic systems to acupuncture stimulation at ST 36 as evidenced by fMRI. Neuroimage 2005;27(3):479-496.
5. Li G, Huang L, Cheung RTF, Liu SR, Ma QY, Yang ES. Cortical activations upon stimulation of the sensorimotor-implicated acupoints. Magnetic Resonance Imaging 2004;22(5):639-644.
6. Kong J, Kaptchuk TJ, Webb JM, et al. Functional neuroanatomical investigation of vision-related acupuncture point specificity-a multisession fMRI study. Human Brain Mapping 2009;30(1):38-46.
7. Wu MT, Sheen JM, Chuang KH, et al. Neuronal specificity of acupuncture response: a fMRI study with electroacupuncture. Neuroimage 2002;16(4):1028-1037.
8. Napadow V, Makris N, Liu J, Kettner NW, Kwong KK, Hui KKS. Effects of electroacupuncture versus manual acupuncture on the human brain as measured by fMRI. Human Brain Mapping 2005;24(3):193-205.

9. Penny WD, Holmes AP, Friston KJ. Random effects analysis. Human brain function 2004:843–850.

10. Aron AR, Gluck MA, Poldrack RA. Long-term test-retest reliability of functional MRI in a classification learning task. Neuroimage 2006;29(3):1000-1006.

11. Havel P, Braun B, Rau S, et al. Reproducibility of activation in four motor paradigms. Journal of Neurology 2006;253(4):471-476.

12. Marshall I, Simonotto E, Deary IJ, et al. Repeatability of motor and working-memory tasks in healthy older volunteers: assessment at functional MR imaging. Radiology 2004;233(3):868-877.

13. Wagner K, Frings L, Quiske A, et al. The reliability of fMRI activations in the medial temporal lobes in a verbal episodic memory task. Neuroimage 2005;28(1):122-131.

14. Yoo SS, Wei X, Dickey CC, Guttmann CRG, Panych LP. Long-term reproducibility analysis of fMRI using hand motor task. International Journal of Neuroscience 2005;115(1):55-77.

15. Liu JZ, Zhang L, Brown RW, Yue GH. Reproducibility of fMRI at 1.5 T in a strictly controlled motor task. Magnetic Resonance in Medicine 2004;52(4):751-760.

16. Smith SM, Beckmann CF, Ramnani N, et al. Variability in fMRI: a re-examination of inter-session differences. Human brain mapping 2005;24(3).

17. Kong J, Gollub RL, Webb JM, Kong JT, Vangel MG, Kwong K. Test–retest

study of fMRI signal change evoked by electroacupuncture stimulation. Neuroimage 2007;34(3):1171-1181.

18. Cho ZH, Hong IK, Kang CK, et al. Acupuncture-stimulated auditory-cortical activation observed by fMRI: a case of acupoint SJ5 stimulation. 2000. p 327.

19. Hui KKS, Liu J, Makris N, et al. Acupuncture modulates the limbic system and subcortical gray structures of the human brain: evidence from fMRI studies in normal subjects. Human Brain Mapping 2000;9(1):13-25.

20. Li G, Cheung RTF, Ma QY, Yang ES. Visual cortical activations on fMRI upon stimulation of the vision-implicated acupoints. Neuroreport 2003;14(5):669-674.

21. Li G, Liu HL, Cheung RTF, et al. An fMRI study comparing brain activation between word generation and electrical stimulation of language-implicated acupoints. Human brain mapping 2003;18(3):233-238.

22. Seghier ML, Lazeyras F, Pegna AJ, Annoni JM, Khateb A. Group analysis and the subject factor in functional magnetic resonance imaging: analysis of fifty right-handed healthy subjects in a semantic language task. Human Brain Mapping 2008;29(4):461-477.

23. Friston KJ, Holmes AP, Price CJ, Büchel C, Worsley KJ. Multisubject fMRI studies and conjunction analyses. Neuroimage 1999;10(4):385-396.

24. Petersson KM. Statistical limitations in functional neuroimaging II. Signal detection and statistical inference. Philosophical Transactions of the Royal Society B: Biological Sciences 1999;354(1387):1261-1281.

25. Brammer MJ, Bullmore ET, Simmons A, et al. Generic brain activation mapping in functional magnetic resonance imaging: a nonparametric approach. Magnetic Resonance Imaging 1997;15(7):763-770.

26. Kherif F, Poline JB, Mériaux S, Benali H, Flandin G, Brett M. Group analysis in functional neuroimaging: selecting subjects using similarity measures. Neuroimage 2003;20(4):2197-2208.

27. Beijing S. Nanjing Colleges of Traditional Chinese Medicine. Essentials of Chinese Acupuncture. Beijing, Foreign Language Press; 1980.

28. Kong J, Gollub R, Huang T, et al. Acupuncture de qi, from qualitative history to quantitative measurement. The Journal of Alternative and Complementary Medicine 2007;13(10):1059-1070.

29. Cook RD, Weisberg S. Residuals and influence in regression: Chapman and Hall New York; 1982.

30. Xue K, Lu X. Essentials of Chinese Acupuncture. Beijing, Foreign Languages Press; 1993.

31. Yang JS, Cao YM. Aupuncture Point. Shanghai: Shanghai Science and Technology Press 181p 1987.

32. Hsieh JC, Tu CH, Chen FP, et al. Activation of the hypothalamus characterizes the acupuncture stimulation at the analgesic point in human: a positron emission tomography study. Neuroscience letters 2001;307(2):105-108.

33. Pariente J, White P, Frackowiak RSJ, Lewith G. Expectancy and belief

modulate the neuronal substrates of pain treated by acupuncture. Neuroimage 2005;25(4):1161-1167.

34. Zhang WT, Jin Z, Cui GH, et al. Relations between brain network activation and analgesic effect induced by low vs. high frequency electrical acupoint stimulation in different subjects: a functional magnetic resonance imaging study. Brain research 2003;982(2):168-178.

35. Nagai M, Kishi K, Kato S. Insular cortex and neuropsychiatric disorders: A review of recent literature. European Psychiatry 2007;22(6):387-394.

36. Shelley BP, Trimble MR. The insular Lobe of Reil-its Anatamico-Functional, behavioural and Neuropsychiatric attributes in humans-a review. World Journal of Biological Psychiatry 2004;5(4):176-200.

37. Davis KD, Kwan CL, Crawley AP, Mikulis DJ. Functional MRI study of thalamic and cortical activations evoked by cutaneous heat, cold, and tactile stimuli. Journal of Neurophysiology 1998;80(3):1533-1546.

38. Biella G, Sotgiu ML, Pellegata G, Paulesu E, Castiglioni I, Fazio F. Acupuncture produces central activations in pain regions. Neuroimage 2001;14(1):60-66.

39. Ingvar M, Hsieh JC. The image of pain. Textbook of Pain 1999:215-234.

FIGURE LEGENDS

Figure 1:

Experimental paradigm: Panel A indicates that acupuncture needle manipulation was performed at acupoint ST36 (Zusanli, arrow pointing to red dot). An acupuncture needle was inserted at ST36, and after resting for 1 min, the needle was manipulated for 1.5 minutes. The needle then remained inserted at the acupoint for another 12.5 minutes. Panel B presents the design paradigm for the checkerboard task. Visual stimulus was presented for 1 min, with 1 min rest between stimuli, preceded by a 1 min baseline, followed by 1 min rest. Total scanning time lasted 5 min.

Figure 2:

Panel A shows 2D multidimensional scaling (MDS) representations of distances between subjects in the checkerboard task. Panel B shows 2D MDS representations of distances between subjects during acupuncture stimulation (ACUP) without a possible outlier. The mean distances of both tasks are displayed in panel C respectively. The mean distance of ACUP is significant larger than in the ACUP stimulation ($P < 0.05$).

Figure 3:

Panel A depicts the mean distances of each subject during acupuncture stimulation (ACUP) in descending order. Subject 5 shows a much larger mean distance than the others. Panel B presents the BOXPLOT of these mean distances, showing that subject 5 must be an outlier at this statistical benchmark. Panel C shows the random effect analyses before and after removal of the outlier ($P < 0.001$, uncorrected).

Figure 4:

The figure shows the Pearson correlation coefficient between the VAS index and the mean distance and regression plots. The panel in top right corner shows the VAS index of each subject.

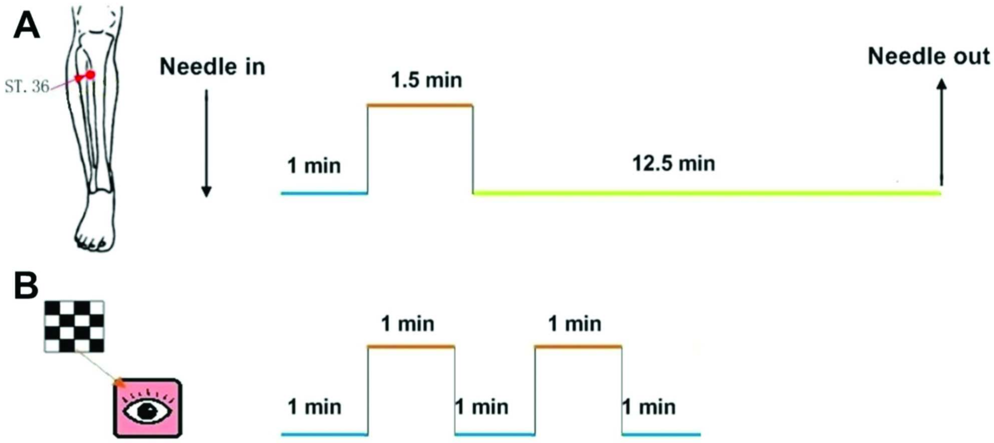

Figure 1. Experimental paradigm: Panel A indicates that acupuncture needle manipulation was performed at acupoint ST36 (Zusanli, arrow pointing to red dot). An acupuncture needle was inserted at ST36, and after resting for 1 min, the needle was manipulated for 1.5 minutes. The needle then remained inserted at the acupoint for another 12.5 minutes. Panel B presents the design paradigm for the checkerboard task. Visual stimulus was presented for 1 min, with 1 min rest between stimuli, preceded by a 1 min baseline, followed by 1 min rest. Total scanning time lasted 5 min.
254x120mm (600 x 600 DPI)

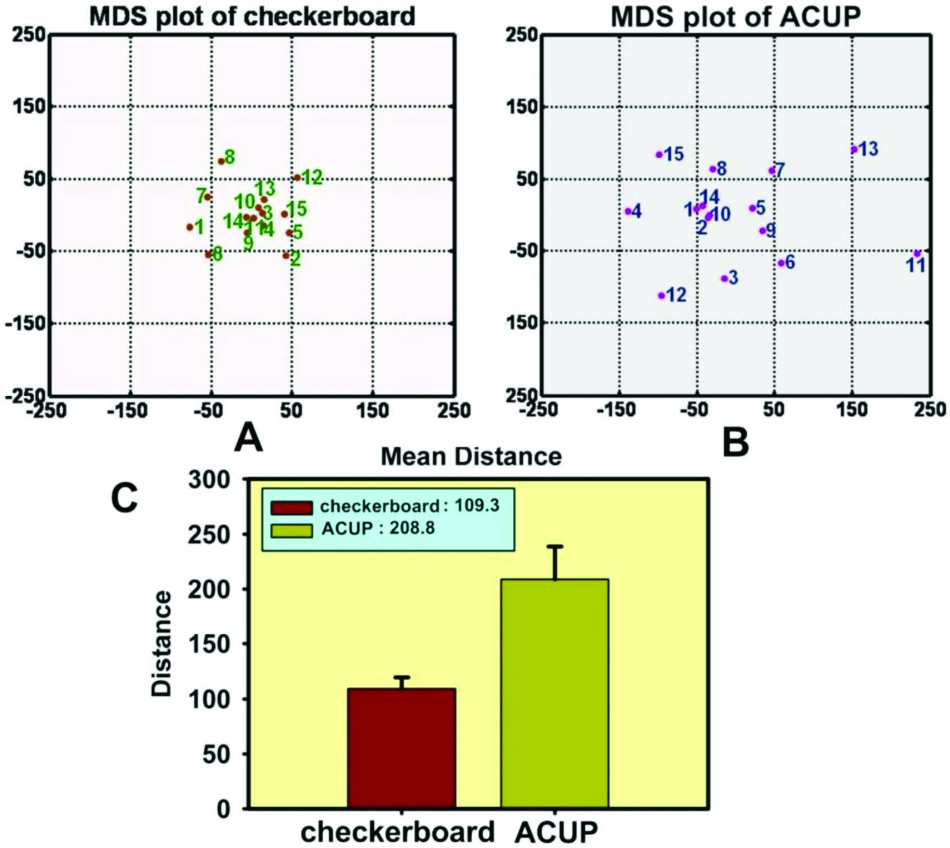

Figure 2. Panel A shows 2D multidimensional scaling (MDS) representations of distances between subjects in the checkerboard task. Panel B shows 2D MDS representations of distances between subjects during acupuncture stimulation (ACUP) without a possible outlier. The mean distances of both tasks are displayed in panel C respectively. The mean distance of ACUP is significant larger than in the ACUP stimulation ($P < 0.05$).
152x133mm (600 x 600 DPI)

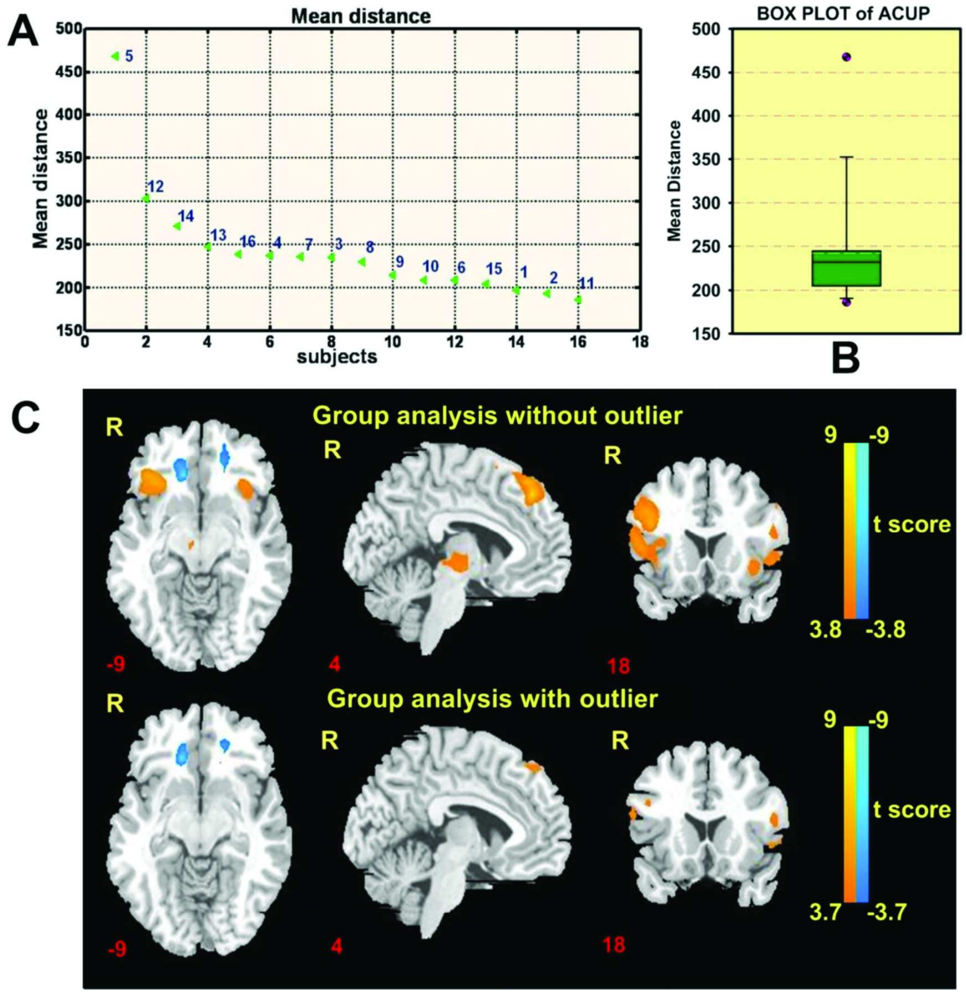

Figure 3. Panel A depicts the mean distances of each subject during acupuncture stimulation (ACUP) in descending order. Subject 5 shows a much larger mean distance than the others. Panel B presents the BOXPLOT of these mean distances, showing that subject 5 must be an outlier at this statistical benchmark. Panel C shows the random effect analyses before and after removal of the outlier (P < 0.001, uncorrected).
152x150mm (600 x 600 DPI)

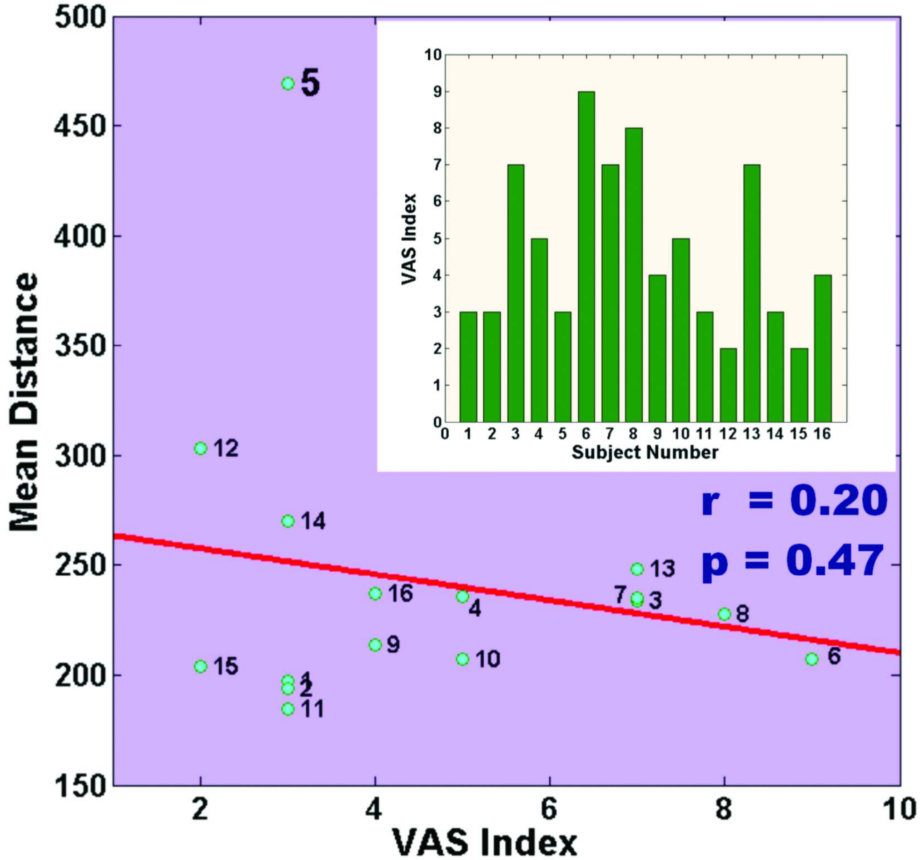

Figure 4. The figure shows the Pearson correlation coefficient between the VAS index and the mean distance and regression plots. The panel in top right corner shows the VAS index of each subject.
152x140mm (600 x 600 DPI)