A Structural Equation Modeling Approach for the Estimation of Genetic and Environmental Effects from Twin fMRI Data

Yu Yong Choi, Jong-In Song, Jang Soo Chun, Kun Ho Lee, and Woo Keun Song

***Abstract*—Structural equation modeling (SEM) is a statistical technique widely used in quantitative genetics to measure genetic and environmental variances of human traits. Using SEM, the proportions of genetic and environmental influences can be separated from the phenotypic variance. However, the SEM softwares like Mx or LISREL were not designed for a big data analysis. They can hardly be applied for brain images that comprise hundreds of thousands of voxels. Here, to introduce SEM in the field of neuroimaging, we developed a simple code in MATLAB for multiple computations of Mx. Our method could estimate genetic and environmental variances of neural activations at 153,594 voxels of the whole brain, to be converted to brain images.**

***Index Terms*—fMRI, genetics, structural equation modeling, twin.**

I. Introduction

In the field of human behavioral genetics, twin research is an indispensable tool to understand nature versus nurture. Twin studies use structural equation modeling (SEM) as a standard analysis method to reveal the relative importance of genetic and environmental influences on human traits and behaviors. SEM is a statistical technique that is based on decomposition of the phenotypic variance into the genetic and environmental components. The genetic component is called heritability, which is a fundamental notion in quantitative genetics that summarizes how much of the variation in a trait among individuals is attributable to differences in genotype [1]. Heritability can also be translated as the standardized genetic variance, which means the proportion of total phenotypic variance due to genetic influence.

In SEM, the standardized genetic variance (a^2) is defined as

$$a^2 = V_G / V_P \quad (1)$$

where V_G is the genetic variance and V_P is the phenotypic variance. The standardized environmental variance (e^2) is defined as

$$e^2 = V_E / V_P \quad (2)$$

where V_E is the environmental variance.

The estimation of genetic and environmental variances using SEM is a complex process. Variance decomposition for estimation of the genetic variance requires various sophisticated statistical software tools such as a matrix algebra interpreter and a numerical optimizer (e.g. Mx, or LISREL). The optimizer is used to minimize the fitting function that denotes a discrepancy measure between the expected model and the observed data. The iterative process of model fitting continues until the fitting function appears to reach the minimum.

The SEM fit is so time-consuming that it is not easy to apply to large-scale data like brain images that consist of numerous voxels. Moreover, despite the increasing computational power of the modern computer, the SEM software packages did not provide sequential as well as parallel processing facilities that are necessary to manipulate numerous data together.

We made a simple software tool for neuroimaging. Using our software tool, researchers can apply the SEM technique for neuroimaging analysis to distinguish genetic and environmental influences on the brain function as well as the structure. Additionally, because our tool is coded in MATLAB (http://www.mathworks.com/), the parallel computing toolbox can be used to enhance the performance. The voxels of a brain image are independent of one another in terms of the fitting process. Thus, even a multicore desktop can shorten the total processing time of all voxels of the whole brain.

II. Method

A. Twin Subjects

To measure genetic and environmental variances of the brain activations, we recruited twin volunteers. The study protocol was approved by the relevant institutional review boards (Seoul National University, Catholic University of Korea), and written informed consent was obtained from participants. A total of 26 healthy twin volunteers aged 20.3 ± 1.7 (mean ± SD), consisting of 16 MZ and 10 DZ male twins, were recruited from the community with advertisements. The MZ and DZ pairs were matched for sex.

Manuscript received October 10, 2012; revised February 3, 2013. This work was supported by grants from the Bio Imaging Research Center (BIC) at Gwangju Institute of Science and Technology (GIST), and by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0025102).

Y. Y. Choi, J. I. Song, J. S. Chun, and W. K. Song are with BioImaging Research Center, Gwangju Institute of Science and Technology, Gwangju 500-712 Korea (e-mail: yuyongchoi@gist.ac.kr, jisong@gist.ac.kr, jschun@gist.ac.kr, wksong@gist.ac.kr).

K. H. Lee is with Dementia Center, Chosun University, Gwangju 501-759 Korea (e-mail: leekho@chosun.ac.kr).

Blood or hair samples were taken at the date of scanning or cognitive testing. Zygosity was determined by DNA analysis using the 15 highly polymorphic markers.

B. Image Acquisition and Analysis

The functional MRI (fMRI) tasks requiring fluid reasoning ability and the fMRI protocol were as described in our previous paper [2]. The statistical parameters for cortical activation such as *t* scores were acquired using the statistical parametric mapping software (SPM2, http://www.fil.ion.ucl.ac.uk/spm/).

To determine genetic and environmental variances, the conventional, univariate ACE model was adopted [3]. The ACE model decomposes the phenotypic variance into additive genetic (A), shared environmental (C), and non-shared environmental (E) variances. The statistical significance of the genetic variance was derived from chi-square difference between ACE and CE models. Genetic and environmental variances mentioned in this paper mean A and E, respectively.

C. Software Development

We decided to use the methodological heritage of genetic researchers as much as possible. This "Do not re-invent the wheel" approach could save a software developer time and labor, and would enable a researcher to apply easily the methods in genetics to neuroimaging analysis. Moreover, it could produce the reliable results because the methods and tools were verified in the research field.

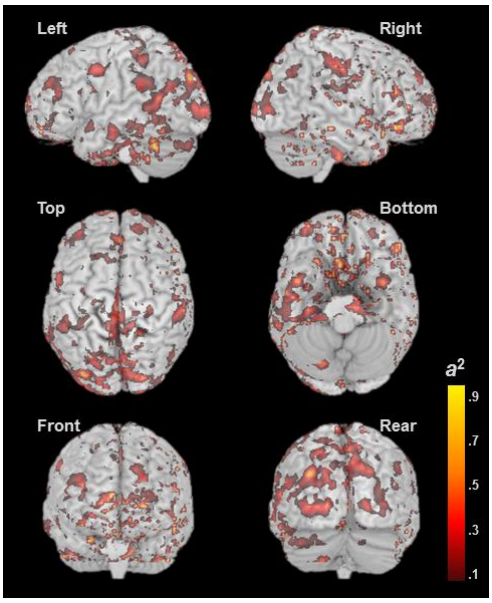

Fig. 1. Genetic variances across the whole brain.

Among the SEM software packages, Mx was chosen for analysis. The Mx software is widely used in human genetics, particularly in twin studies because it facilitates specification of complex models and mixture distributions and provides diverse model fitting functions [3]. To apply the SEM software Mx to neuroimaging analysis, we write a MATLAB code for sequential processing of multiple data. A simple version of the algorithm of our program is below:

Array A[N], E[N];
convert t-score images from SPM2 **to** the text files;
for voxel = 1 to N
  read t-scores of all twins at the voxel;
  write data for Mx;
  execute Mx;
  parse the result_in_text from Mx;
  A[voxel] = the genetic variance from the parsing;
  E[voxel] = the environmental variance from the parsing
end for
write A, E;
convert the A text file **to** the images in SPM2;
convert the E text file **to** the images in SPM2;

D. Software Development

- Programming language: MATLAB 7.7
- SEM software: Mx 1.52b
- fMRI analysis: SPM2
- Brain image viewer: MRICRO 1.39

III. Results

The univariate ACE structural equation model, a standard model for twin analysis, was employed to determine what proportion of variance in a brain-based phenotype (e.g., brain activation) is heritable (a^2), versus the proportions which are due to shared environment or non-shared environment (e^2). The brain images for *t* scores (activation maps) had the dimensions of 53 × 63 × 46. By fitting the ACE model to *t* scores at each voxel of the whole brain, we obtained the standardized genetic and environmental variances of the brain activations at 153,594 voxels. All the standardized genetic and environmental variances were overlaid on the standard template using MRICRO. Finally, we produced 3-dimensional maps of the genetic and environmental variances of the brain activation during fluid reasoning tasks (Fig. 1 and Fig. 2).

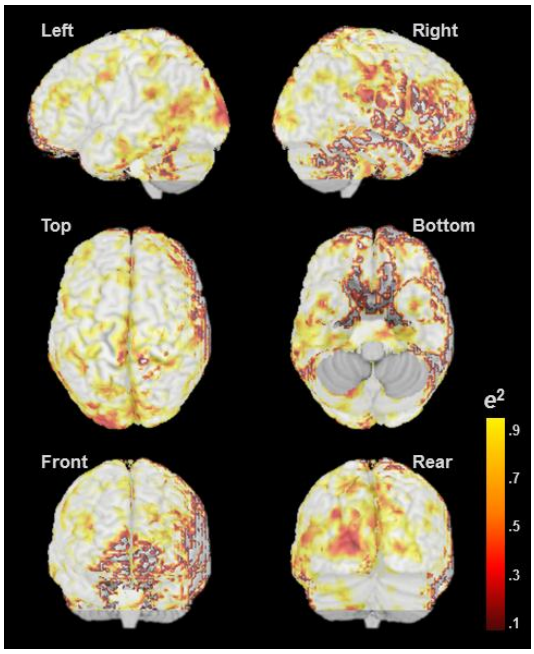

Fig. 2. Environmental variances across the whole brain.

A. Genetic Components of the Variance of Brain Activation

Using our tool, we determined the genetic influence on the

cortical activation during reasoning. As shown in Fig. 1, the genetic variances were scattered all over the cerebral cortex. Compared to other regions, the inferior frontal and superior temporal regions in the right hemisphere showed relatively strong heritability.

B. Environmental Components of the Variance of Brain Activation

The environmental influence on the cortical activation was also determined. As shown in Fig. 2, compared to the genetic variances in Fig. 1, the environmental variances were very high in most regions of the cerebral cortex. The prefrontal, and right temporal regions showed relatively low environmental variances compared to other regions.

IV. Conclusion

We provided a simple software tool to apply genetic analyzing methods for functional neuroimaging studies. This tool is an upgraded version of our previous tool that was applied for anatomical MRI only [4]. Now, this simple tool enables us to perform SEM analysis of functional and anatomical brain images. SEM is a statistical method with many advantages. SEM provides the statistical significance as well as heritability estimates. Moreover, SEM can distinguish shared and random environmental effects. Now, the proposed method facilitates diverse and complex models that were used only in genetic research for neuroimaging analysis including anatomical and functional MRI. In future, we will apply multivariate SEM model for anatomical and functional neuroimaging analysis.

Acknowledgment

The authors thank Dr. Sun Hee Cho, Ms. Min Joo Lee, Prof. Young Bo Kim, Prof. Zang-Hee Cho, and Prof. Ki Hwa Chung for their help in data collection.

References

[1] K. H. Lee, Y. Y. Choi, J. R. Gray, S. H. Cho, J. H. Chae, S. Lee, and K. Kim, "Neural correlates of superior intelligence: stronger recruitment of posterior parietal cortex," *NeuroImage*, vol. 29, pp. 578-86, Jan. 15 2006.

[2] D. S. Falconer and T. F. C. MacKay, *Introduction to Quantitative Genetics*, 4th ed. Harlow, Essex, UK: Longman, 1996.

[3] M. C. Neale, S. M. Boker, G. Xie, and H. H. Maes, *Mx: Statistical Modeling*, 6th ed. VCU Box 900126, Richmond, VA 23298: Department of Psychiatry, 2003.

[4] Y. Y. Choi and K. H. Lee, "Heritability estimation methods of multiple brain measures: A preliminary MRI study in twins," presented at the Medical Image Analysis and Description for Diagnosis Systems, Rome, Italy, 2011.

Yu Yong Choi received the B.S. degree in computer engineering from Seoul National University, Seoul, Korea in 1993, the M.S. degree in computer engineering from Seoul National University, Seoul, Korea in 1995, and the Ph.D. degree in interdisciplinary program in neuroscience from Seoul National University, Seoul, Korea in 2008.

From 2006 to 2009, he worked as a researcher for Laxtha Research and Developemnt Instititute, where he developed a technique to measure human intelligence using brain images. From 2009 to 2010, he worked as a researcher and lecturer for Research Institute for Image and Cultural Content in Dongguk University where he studied cognitive enhancement techniuqes through computer games. From 2010 to 2011, he worked as a research professor for Dept. of Computer Engineering in Hanyang University. In 2011, he joined Gwangju Institute of Science and Technology (GIST), Gwangju, Korea, and currently works as a research professor for BioImaging Research Center. His current research interests include early dectection of dementia onset and cognivie enhancement.

Jong-In Song received the B.S. degree in electronics engineering from Seoul National University, Seoul, Korea, in 1980, the M.S. degree in electronics engineering from Korea Advanced Institute of Science and Technology, Daejon, Korea, in 1982, and the Ph.D. degree in electrical and electronics engineering from Columbia University, New York, NY, in 1990.

From 1986 to 1990, he worked as a Graduate Research Assistant for the Center for Telecommunications Research, where he pioneered high-performance GaAs/AlGaAs two-dimensional electron gas (2DEG) charge coupled device research for microwave and infrared-imaging applications. From 1990 to 1994, he joined the Electronics Science and Technology Division in Bellcore, where he worked on the development of microwave transistors including GaInP/GaAs, InAlAs/InGaAs, InP/InGaAs HBTs, and their application to MMICs. In 1994, he joined Gwangju Institute of Science and Technology (GIST), Gwangju, Korea, and is currently a Professor in the Department of Information and Communications. His current research interests include low-power and high-speed devices and cicuits, millimeter-wave over fiber (MMoF) communication systems, and distributed sensor networks.

Jang-Soo Chun received the B.S. degree in biology from Pusan National University, Pusan, Korea, in 1986, the M.S. degree in biology from usan National University, Pusan, Korea, in 1988, and the Ph.D. degree in molecular and cellular biology from Univeristy of Massachusetts, Amherst, Massachusetts in 1992. From 1992 to 1994, he worked as a PostDoc at Harvard Me4dical School where he pioneered biological functions of syndecans in signal transduction. From 1992 to 2000, he worked as an assistant and associate professor at Department of Biology, Kyungpook National University, Dagu, Korea. In 2000, he moved to Gwangju Institute of Science and Technology (GIST), Gwangju, Korea, and is currently a Professor in the Department of Life Sciences. His current research interests include molecular mechanisms of cartilage degeneration.

Kun Ho Lee received the B.S. degree (1989) at Department of Genetic Engineering, Korea University, Seoul, Korea, the M.S. dgree (1992) and Ph.D. degree (1998) at Department of Molecular Biology, Seoul National University, Seoul, Korea.

From 1998 to 1999, he worked as a Post-doctoral fellow at the Center for Cell Differentiation and then moved to Department of Pharmachology, School of madicine, University of Pittsbergh to extend his research as a Research Associate for one and half year. After comeback to Seoul, he focused on neuroimaging study at the Korea Institute of Brain Science for three years. From 2004 to 2009, he joined the Brain Research Center at Seoul National University, where he worked on the development of neurometric model of human intelligence based on MR imaging. In 2009, he moved to Chosun University, Gwangju, Korea, and work at Dementia Center as the Research Director and is also affilated at College of Natural Sciences. His current research interests include brain image analysis and development of prediction model of neurodegenerative diseases based on MR imaging.

Woo Keun Song received the B.S. in Dept. of Molecular Biology from Seoul National University, Seoul, Korea in 1982, and the M.S. degree in Dept. of Molecular Biology from Seoul National University, Seoul, Korea in 1984, and the Ph.D. degree in Dept of Microbiology in University of Illinois at Urbana-Champaign, Illinois, USA, 1992.

From 1992 to1993, he worked as a postdocteral fellow in University of Illinois, where he have developed cellular imaging technology. At 1994, he joined Gwangju Institute of Science and Technology (GIST), Gwangju, Korea, and currently works as a director for BioImaging Research Center and professor in Dept. of Life Science. His current research interest is to develop cellular imaging technology to visualize cell dynamics.