Healthy Twin: A Twin-Family Study of Korea — Protocols and Current Status

Joohon Sung,[1] Sung-Il Cho,[2] Kayoung Lee,[3] Mina Ha,[4] Eun-Young Choi,[5] Ji-Sook Choi,[6] Ho Kim,[2] Jihae Kim,[7] Kyoung Sue Hong,[7] Yeonju Kim,[8] Keun-Young Yoo,[8] Chan Park,[9] and Yun-Mi Song[10]

[1] Department of Preventive Medicine, Kangwon National University College of Medicine, Korea
[2] Department of Epidemiology and Biostatistics. School of Public Health and Institute of Health and Environment, Seoul National University, Korea
[3] Department of Family Medicine, Busan Paik Hospital, Inje University College of Medicine, Korea
[4] Department of Preventive Medicine, Dankook University College of Medicine, Korea
[5] Department of Family Medicine, Dankook University College of Medicine, Korea
[6] Center for Clinical Research, Samsung Biomedical Research Institute, Korea
[7] Department of Psychiatrics, Samsung Medical Center, SungKyunKwan University School of Medicine, Korea
[8] Department of Preventive Medicine, Seoul National University College of Medicine, Korea
[9] National Genomic Research Institute, National Institute of Health, Korea
[10] Department of Family Medicine, Samsung Medical Center, SungKyunKwan University School of Medicine, Korea

'Healthy Twin' is a twin family study extension of the existing Korean Twin-Family Register. Healthy Twin recruits adult like-sex twins over the age of 30 and their adult family members. Healthy Twin protocols are primarily tailored to the study of the quantitative trait loci of complex traits as well as to the role of environment in the etiology of complex diseases. A full-length survey is underway, including questionnaires, health examinations and the collection of biological specimens. So far, 820 individuals (169 twin pairs and their families) have participated in the survey and 1068 individual twins (608 twin pairs) have replied to the mailed zygosity questionnaire as of July 2006. The first phase (2005–2006) of Healthy Twin will recruit 1550 individuals (including about 380 twin pairs), and the second phase a proposed 1500 to 2500 additional participants. We report study protocols and zygosity and the distribution of family size of the study participants.

'Healthy Twin' is an extended twin and family study of the existing Korean Twin-Family Register (KTR; Sung et al., 2002). Healthy Twin is primarily aimed at elucidating genetic variations responsible for common complex diseases/traits. The researchers are also eager to incorporate the unique strength of twin studies in evaluating the etiologic importance of environmental factors and epigenetics. We have focused on the recruitment of twins and their families unselected by health status. This kind of family study design has specific strength in detecting quantitative trait loci (QTL) and genetic variations underlying common diseases (Peltonen & McKusick, 2001). Metabolic syndrome-related phenotypes are the primary target of the study, and a range of traits have been measured as well. A biorepository for DNA extracts, immortalized cell line, sera, plasma, and urine sample of every participant is being preserved for examining diverse research hypotheses.

In Korea, twin studies or human genetic studies were uncommon. Since the success of the Human Genome Project, human genetics have begun to be widely accepted as a priority in medical research. Healthy Twin was originally planned as a family-based module of the large Korean Genomic Cohort Study of adult individuals (KoGeS; Yoo et al., 2005) by the National Genomic Research Institute of the Center for Disease Control, Korea (NGRI). Healthy Twin shares basic protocols with this population-based mother cohort, the KoGeS. While work on the KTR has been ongoing, Healthy Twin is new as it has started to survey participants in person, and is an extended study of the nationwide KTR as it includes the family members of twins as well. We believe this study to be one of the rare non-Caucasian twin studies including family and genomic information. The first year of the study has been success with the recruitment of 704 individuals (including pairs of twins and their families), and the second year of the study is ongoing and to be completed by December 2006 (in total, 1550 individuals including 365 twin pairs). The next phase of Healthy Twin is being planned to survey an additional 1500 to 2500 twin individuals and their family together with the analyses of gathered data. We report the study design, protocols, and interim results of Healthy Twin for future reference.

Aims

The ultimate purpose of Healthy Twin is to elucidate genetic and environmental factors underlying complex
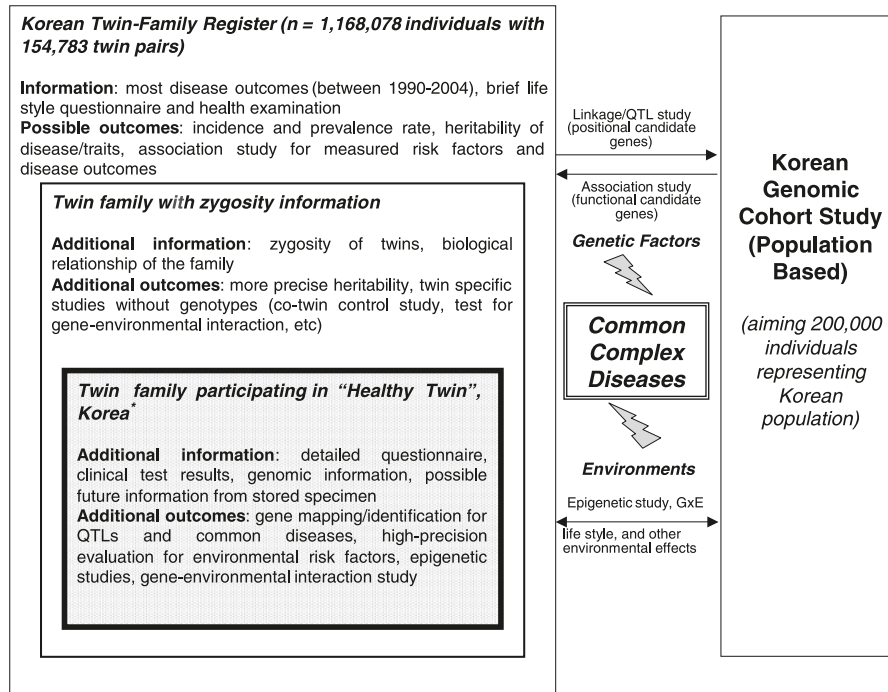
Figure 1

Schematic explanation about the relationship between the Korean Twin Register and the new 'Healthy Twin', and strategic plans for the study.

Note: * current goal: 3000 individuals (including 750 twin pairs) in 3 years.

human diseases and traits. However, in the current phase, the grant from NGRI is confined to survey and data collection itself. The development and validation of research protocols for twin and family study is another goal of the first phase. Metabolic syndrome, obesity and osteoporosis are traits of primary interest. Some of the phenotypic information and a part of the biorepository will be open to the public, and NGRI plans to run a committee for the assessment of researchers wishing to use this.

Methods

Subjects

The study recruits adult (30 years or older) like-sex twins and their first degree family members. Either the twin pair only, or the twin pair with their family (≥ 4 persons in a family) was recruited. For example, a family unit may consist of like-sex twins and two or more of their parents or siblings (parents, a parent and a sibling, and/or two siblings all qualify), or twins and their offspring together with the twins' spouses. For families of more than six participants, a small financial incentive (about $100 USD) was provided, to encourage large twin families to participate.

Healthy Twin was advertised in a nationwide newspaper and also by posters in about 300 hospitals and health-related governmental agencies. Among the twins of the existing KTR, those whose contact information was provided by the participating hospitals were contacted first. Around 12,500 individual twins whose

addresses were available were contacted via mail and introduced to the Healthy Twin health examination and research program, and asked to participate. In addition, a brief zygosity questionnaire consisting of six questions was included. Trained research co-ordinators explained the health examination in detail on the phone and mailed a full-length questionnaire to those who were willing to participate in Healthy Twin, and made an arrangements for the study. Considering the ease of access to high speed internet in Korea, a research website (www.twinkorea.org) was set up where twins and their family members are able to complete a short zygosity questionnaire and apply for participation in Healthy Twin. The relationship between the existing KTR and the new Healthy Twin is schematized in figure 1. As shown in Figure 1, all the participating twins and families share KTR information.

Sample Size

The target sample size of the first phase of this study (2005–2006) was set at 1550 individuals of adult twins (30 years or older) and their first degree family members (20 years or older). If identical twins make up a third of all participants, as was shown from the first year survey, the effective sample size for linkage analysis becomes five sixths of the total, 1290 (half of the identical twins and all others). This figure reaches an approximate sample size to detect quantitative trait loci (QTL) with moderate heritability of between .1 and .15 (Williams & Blangero, 1999). The next phase of Health Twin is pending, but was originally proposed to recruit as many as 3000 to

Table 1

Information on the Questionnaire, Clinical Tests, Physical Measures and Biological Specimens Collected for Healthy Twin

	Tests in health examination	Remark
Questionnaire (no. of questions/ items)	Demographics (6), past medical history (1 table), family history of disease (1 table), occupation (3), job-related stress (15), medication (1 table), smoking, including environmental tobacco smoking with dependency (17), alcohol habits and dependency (14), physical activity (19, International Physical Activity Questionnaire), lifetime weight change (11), anxiety sensitivity (36), state trait anxiety (40), personality (140, temper and character inventory), food frequency table (food items/groups) (100), eating behavior inventory (33), female reproductive history (20), sexual activity history (10)	Sexual activity questions were only asked for those who had the pap smear test
Twin-specific Questionnaire (no. of questions)	Zygosity screening (5), family members (5), birth and development of twin (20), life-time residence (3), Perinatal history (8)	
Clinical tests	Electrocardiogram, pulmonary function test, chest X-ray, body composition analysis (Impedance), body fat measurement (DEXA), bone density (DEXA, whole body)	
Measures and physical examination	Physical examination, blood pressure, height, weight, sitting height, arm span, head circumference, waist circumference, hip circumference	
Facial and body characteristics*	More than 60 quantitative and 10 qualitative traits of face, hand, and foot	
Cancer screening†	Esophagogastroduodenoscopy (EGD), mammography, cytology of cervix uteri (pap smear)	Participants could select either 'EGD' or 'mammography and pap smear together'
Blood test	Fasting plasma glucose, fasting plasma insulin, 2-hour post loading glucose‡, total cholesterol, high density lipoprotein (HDL) cholesterol, low density lipoprotein (LDL) cholesterol (direct measure), triglyceride, high sensitivity C reactive protein (CRP), total protein, albumin, gamma-glutamyl transferase (GGT), aspartate aminotransferase (AST), alanine amino transferase (ALT), creatinine, calcium, uric acid, complete blood cell count (CBC) with differential count	
Urine tests	Routine urinalysis	
Biological specimens (# of 0.5 ml cryovials)	DNA extract (100 Ìg), immortalized B-Cell line (4), RNA preparation (2) (fixed by RNAlater®), serum (2), plasma (2), 12-hour urine (3 conical tubes), red blood cell (1)	Stored in liquid nitrogen tank, except DNA extract (deep freezer) and urine (freezer)

Note: * performed only in a clinical center located in Seoul

† cancer screening service is provided for those who are aged over 40 and want to take the tests

‡ performed for those who agreed to be tested

4000 individuals of twins and their families, which will work for the gene discovery study of most common diseases as well as QTLs with smaller heritability (< .1).

Measurements

This is a multicenter-based study. Every participant is provided with a health examination at one of the three clinical centers located in different geographical areas (Seoul, Pusan, and Cheonan). Before the beginning of study, efforts to standardize the survey method between the three centers were made in terms of the development of a standard protocol, and the training of research coordinators and research assistants. A wide range of questions are asked in the examination, physical characteristics measured and clinical tests performed (see Table 1 for details). Most test batteries were selected from existing standardized Korean language versions, or newly adapted for this study through translation and back-translation.

Biorepository

The biological specimen collected at the three hospitals are sent to one central laboratory, where blood/urine tests are performed and immortalized cell lines (by EBV transformation) are made for every participant who gives informed written consent. The NGRI checks quality assessment results of this central laboratory every 3 months. Types and the amount of collected biological specimens are listed in Table 1. Two test tubes of blood specimen are centrifuged within 2 hours of sampling and sera are aliquoted and frozen immediately. DNA is extracted within 24 hours. Buffy coat was prepared after RNA fixation (using RNAlater). For women who undertake the cytology test of (pap smear) cervix uteri, a cervical smear specimen is also collected for viral DNA study together with a short sexual activity questionnaire. The funding body, NGRI, stores one set of the biological specimens in a well-equipped biorepository. For the sake of risk reduction, a backup copy of biological specimens is kept by the researchers in a separate biorepository as well.

Zygosity Estimation

Six questions to estimate zygosity were asked. Pairs where both twins reported a 'mirror-like resemblance', and schoolteachers or friends 'always or nearly always' had problems distinguishing between them were

Table 2

Age, Sex, and Estimated Zygosity Distribution of Participants in the First Year of Healthy Twin

	Twin and their family members, no. (%)						
Age	Women			Men			Total
	Twin	Family	Subtotal	Twin	Family	Subtotal	
20–29	0	10	10	0	10	10	20 (2.8%)
30–39	140	42	182	60	26	86	268 (38.1%)
40–49	71	64	135	33	26	59	194 (27.6%)
50–59	13	42	55	14	19	33	88 (12.5%)
60–69	1	49	50	2	34	36	86 (12.2%)
≥ 70	0	30	30	0	18	18	48 (6.8%)
Subtotal	225	237	462 (65.6%)	109	133	242 (34.4%)	704 (100%)
Twin							334 (47.4%)
Family							370 (52.6%)

	Zygosity† of twin pairs, no (%)					
	Female twin pairs		Male twin pairs		Total twin pairs	
Zygosity†	Full survey	Mail quest*	Full survey	Mail quest*	Full survey	Mail quest*
Monozygote	74 (65.5%)	224 (65.5%)	40 (71.4%)	235 (68.1%)	114 (67.5%)	459 (66.8%)
Dizygote	20 (17.7%)	83 (24.3%)	9 (16.1%)	95 (27.5%)	29 (17.2%)	178 (25.9%)
Uncertain‡	19 (16.8%)	35 (10.2%)	7 (12.5%)	15 (4.3%)	26 (15.4%)	50 (7.3%)
Subtotal	113	342	56	345	169	687

Note: * 50 pairs who completed a mail questionnaire overlapped with those who completed the full survey

† zygosity estimation was based on questionnaire

‡ in the case of discrepancy between co-twins or ambiguous zygosity of a co-twin

categorized as monozygotic. The pairs who did not report having a 'mirror-like resemblance', and were 'seldom' or 'never' confused by teachers or friends were categorized as dizygotic. All remaining pairs, including those pairs whose zygosity categories conflicted, were regarded as uncertain zygosity.

Genotyping

Although genotyping for full-genome scan or candidate gene study has not been performed, all the blood samples are genotyped for 16 short tandem repeat (STR) markers by NGRI (15 autosomal STR markers + one sex determining marker, using Perkin Elmer's AmpFlSTR Identifier Kit; Cotton et al., 2000) in order to verify the identity of the sample in the future.

Follow-up

Healthy Twin is designed to provide a follow-up health examination for every compliant participant 2 years after initial recruitment. An abbreviated questionnaire is being prepared for repeated survey. We notified each participant with a feedback report on all clinical tests and selected health behaviors by mail. When a participant necessitates further evaluation or intervention regarding newly detected health problems, this is appropriately coordinated and managed. Although the NGRI is covering the cost of the above-mentioned initial health examination and questionnaire survey, any additional medical cost for diagnosing/treating personal medical problems are not supported. The participants will be contacted every 2 years by mail

and/or telephone for a follow-up of health status. Morbidity or mortality can be accessed through nationwide databases, such as cancer registry, vital status report, and the medical utilization record of National Health Insurance as well.

Ethical Considerations

All those who are willing to participate in Healthy Twin are required to give informed written consent to use the information for future research purposes including domestic and international collaborations, storing their biological and genomic specimens, continuing storage longer than 5 years, and genotyping the genomic materials, separately. The institutional review boards from all three hospitals approved the research plans. Any personal information is encrypted, and researchers have differential access levels depending on their roles. A database program for this (based on Oracle version 10e) was provided by the NGRI.

Results

The second year survey is currently underway. To date, 1068 individual twins (608 pairs) have replied to the mailed questionnaire, and 820 individual twins and their families including 386 twins (198 pairs) had completed the full survey as of July 2006. Note that the survey does not obligate simultaneous attendance of a family, which explains the discrepancy between individual and pairwise count. A total of 12,534 were mailed and the response rate to this single mailing was

9.7% (1068 out of 10,989) after excluding returned mail (1545). A total of 704 individuals finished the full survey in the first year. The mean age at examination was 38.9 years for female twins (225 persons), 40.8 for male twins (109 persons), 52 for female family members of twins (237 persons), and 52 for male family members (133 persons). Age, sex, and the family size distribution of the participants, as well as the questionnaire-based zygosity of twins are shown in Table 2. Note that in some twin pairs (one female pair and three male pairs), only one co-twin has participated thus far, which results in the discrepancy between pairwise and individual twin numbers.

Seven hundred and four individuals from 169 families with an average family size of 4.2 as a whole participated, with the average family size increasing to 5.0 if restricted to those twins included with their families. The largest family consisted of 12 family members.

Discussion

Families ascertained from twins are free from ascertainment bias since twinning is known to be associated with only a few rare health conditions such as cerebral palsy (Newton et al., 2003). Our recruitment process relied both on volunteers/advertisement and population/register. Volunteers may have different risk factors for health characteristics, especially in the studies of psychiatrics (Amori & Lenox, 1989) or sexuality (Strassberg & Lowe, 1995), but the cardiovascular risk factors of volunteers were reported to be similar to those of the general population (Remington et al., 1978). More female and identical twin pairs were recruited than male and fraternal twin pairs in this study. This monozygotic excess was expected as recruitment was restricted to like-sex twins, excluding opposite-sex twins. When we doubled the number of dizygotic like-sex twins to adjust for the absence of opposite-sex twins, the monozygotic–dizygotic twin ratio decreased from 3.9 to 2 for the 704 examinees, and from 2.6 to 1.3 for the mail responders. The female to male twin ratio was also 2 for the examinees which corresponds to the corrected monozygotic–dizygotic twin ratio, and 1 for the mail responders. The general characteristics of volunteer twins (Heath, 2001) were exaggerated in twin-family settings, where adult twins (most of them living apart) were encouraged to participate with their family members, probably as female and identical twin pairs were more likely to contact and organize their families to participate in Healthy Twin. Also, the response rate (9.7%) to a single mail questionnaire was lower for males (30.9%) compared to females (42.1%) in America (Cockburn et al., 2002), 73 to 87% in Japan (Hayakawa et al., 2002), or those in European countries. The response rate may rise slightly if the number of persons who could actual receive the mail were counted. It is thought that this large difference in response rate is partly due to the fact that only twins who wished to participate with their family replied. Further analysis on the predictive factors of response is ongoing.

References

Amori, G., & Lenox, R. H. (1989). Do volunteer subjects bias clinical trials? *Journal of Clinical Psychopharmacology, 9*, 321–327.

Cockburn, M., Hamilton, A., Zadnick, J., Cozen, W., & Mack, T. M. (2002). The occurrence of chronic disease and other conditions in a large population-based cohort of native Californian twins. *Twin Research, 5*, 460–467.

Cotton, E. A., Allsop, R. F., Guest, J. L., Frazier, R. R., Koumi, P., & Callow, I. P. (2000). Validation of the AMPFlSTR SGM plus system for use in forensic casework. *Forensic Science International, 112*, 151–161.

Hayakawa, K., Shimizu, T., Kato, K., Onoi, M., & Kobayashi, Y. (2002). A gerontological cohort study of aged twins: The Osaka University Aged Twin Registry. *Twin Research, 5,* 387–388.

Heath, A. C., Howells, W., Kirk, K. M., Madden, P. A., Bucholz, K. K., Nelson, E. C., Slutske, W. S., Statham, D. J., & Martin, N. G. (2001). Predictors of non-response to a questionnaire survey of a volunteer twin panel: Findings from the Australian 1989 twin cohort. *Twin Research, 4,* 73–80.

Newton, R., Casabonne, D., Johnson, A., & Pharoah, P. (2003). A case-control study of vanishing twin as a risk factor for cerebral palsy. *Twin Research, 6,* 83–84.

Peltonen, L., & McKusick V. A. (2001). Genomics and medicine. Dissecting human disease in the postgenomic era. *Science, 291,* 1224–91229.

Remington, R. D., Taylor, H. L., & Buskirk, E. R. (1978). A method for assessing volunteer bias and its application to a cardiovascular disease prevention programme involving physical activity. *Journal of Epidemiology and Community Health, 32,* 250–255.

Strassberg, D. S., & Lowe, K. (1995). Volunteer bias in sexuality research. *Archives of Sexual Behavior, 24,* 369–382.

Sung, J., Cho, S. H., Cho, S. I., Duffy, D. L., Kim, J. H., Kim, H., Park, K. S., & Park, S. K. (2002). The Korean twin registry-methods, current stage, and interim results. *Twin Research, 5,* 394–400.

Williams, J. T., & Blangero, J. (1999). Power of variance component linkage analysis to detect quantitative trait loci. *Annals of Human Genetics, 63,* 545–563.

Yoo, K. Y., Shin, H. R., Chang, S. H., Choi, B. Y., Hong, Y. C., Kim, D. H., Kang, D., Cho, N. H., Shin, C., & Jin, Y. W. (2005). Genomic epidemiology cohorts in Korea: Present and the future. *Asian Pacific Journal of Cancer Prevention, 6,* 238–243.