Published by Oxford University Press on behalf of the International Epidemiological Association
International Journal of Epidemiology 2006;**35**:488–498

Letters to the Editor

Do we need more twin studies? The Healthy Twin Study, Korea

From JOOHON SUNG,[1]* SUNG-IL CHO,[2] YUN-MI SONG,[3] KAYOUNG LEE,[4] EUN-YOUNG CHOI,[5] MINA HA,[6] JIHAE KIM,[7] HO KIM,[2] YEONJU KIM,[8] EUN-KYUNG SHIN,[2] YOON-HEE KIM,[2] KEUN-YOUNG YOO,[8] CHAN PARK[9] and KUCHAN KIMM[9]

The success of the Human Genome Project and recent technological progress have made the analysis of individual genetic variation much more feasible. However, genetic variations responsible for common complex human diseases and traits are largely unknown, with a few exceptions of specific subtype of common diseases (e.g. early onset type diabetes or Alzheimer disease).[1] Twin study has evolved from the classical twin study design, in which comparison of concordance in traits or diseases between monozygotic and dizygotic twins provided evidence about the genetic and environmental contributions to the phenotypes/diseases of interest.[2] Current twin studies in Europe, Australia, and other countries are already comprehensive genomic studies rather than a classical twin design, which maximize the presence of twins. Examples of innovation include: linkage study using dizygotic twin pairs; linkage and association study using the family of twins; epigenetic study using monozygotic twins; and tests for gene–environmental interaction using twins. The list is ever growing as creative minds apply themselves.[3,4] Although contemporary twin study has the potential to incorporate multidimensional genetic approach, most existing twin studies were started in the 1970s and 1980s, when genomic information and families of twins were rarely collected. It was not until after the 1990s that the availability of genomic information and the value of recruiting families of twins started to be emphasized. Furthermore, most twin studies have been performed in Caucasian populations, which have different genetic variations as well as different cultures and environments from the other twin collections.

[1] Department of Preventive Medicine, Kangwon National University College of Medicine, Korea.

[2] Department of Epidemiology and Biostatistics, School of Public Health, Seoul National University, Seoul, Korea.

[3] Department of Family Medicine, Samsung Medical Center, SungKyunKwan University School of Medicine, Seoul, Korea.

[4] Department of Family Medicine, Inje University College of Medicine, Korea.

[5] Department of Family Medicine, Dankook University College of Medicine, Korea.

[6] Department of Preventive Medicine, Dankook University College of Medicine, Korea.

[7] Department of Psychiatrics, Samsung Medical Center, SungKyunKwan University School of Medicine, Seoul, Korea.

[8] Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea.

[9] National Genomic Research Institute, National Institute of Health, Korea.

* Corresponding author. Department of Preventive Medicine, Kangwon National University College of Medicine, Hyoja-2-Dong, Chunchon, Kangwon-Do, Korea. E-mail: sungjohn@kangwon.ac.kr

Twin studies in low-income and middle-income countries

In dissecting complexities of common disease, twin study offers great power, by matching all genomic and most of the environmental background. No other human study designs can come close to the twin design in minimizing background noise. Given the huge background signals from multiple genes and environmental factors, and possible interactions between them, the strength of twin research is unique, because it has the natural way of cancelling unwanted signals—by the matching. However, studies of monozygotic twins cannot contribute to gene discovery studies, and dizygotic twins, which would be valuable here, are sometimes are not readily available in conventional twin studies. Family-based studies are suitable for searching genetic variation related to diseases. Classical family studies, ascertained by the presence of specific diseases or traits, as a rule, have greater power on those selected phenotypes. However, selected families have less power for other phenotypes,[5] which make them inappropriate for more general study purposes. Population-based study design has been successful in identifying environmental risk factors of common diseases. However, attempts to identify genetic variations underlying common diseases through population-based association studies have suffered from lack of reliable replication.[5]

Unlike classical twin studies, current twin study design is a hybrid of classical twin, family, and population-based study designs.[2,4] Recruiting the twins is generally the most difficult part of the study. Thus, most twin studies are performed using registers or cohorts, of either volunteers or populations. Most of the twin registers and twin research have been organized in European countries, Australia, and North America.[4,6]

To get a comprehensive view of the current state of twin research in low-income and middle-income countries, we reviewed twin studies from around the world (Table 1). All the published twin studies worldwide since 1968 were searched for using electronic bibliographic databases. We found 84 published papers derived from twin studies from 14 countries low-income and middle-income countries. Only five definitive twin studies derived from registers or using cohort designs were found among these publications. Except for one recent study in Brazil,[7] all the other studies were from Asia and Africa (Table 1). These comprise only 2.7% of the total number of 3069 published papers using twin study designs found since 1950. Currently there are thought to be 40 twin registers/cohorts for various study purposes, but only five are from non-Caucasian populations.[4] Most of the published twin studies were accomplished sporadically and not based on a systematic birth register

Table 1 Published twin studies in Asia and African countries (between 1967 and 2005)

Continents	Nationality of twins	No of publications	Publication window	Register or cohort based (size of the register)	Systematic collection of genomic information	Main area of interest
Asia	China	6	2000–05	Yes (4500)	No	Common diseases, health behaviour
	India	10	1981–2005	No	For zygosity determination	Various complex diseases
	Iran	1	2004	No	No	Polycyctic ovary
	Israel	9	1992–2005	No	No	Birth and perinatal outcomes
	Japan	18	1969–2004	Yes (12 000)	For zygosity determination	Ageing, lipids, cognition, lifestyle, quality of life
	Korea	5	2002–05	2 registers (154 783/4615)	For zygosity determination	Common complex diseases/traits, cognitive abililty
	Singapore	1	2004	No	No	Twin birth rate
	Sri Lanka	6	2000–03	Yes (20 294)	No	Multidisciplinary (not specifies)
	Taiwan	15	1967–2004	No	No	Complex traits
Africa	Gambia	6	1994–2003	No	No	Tropical infection
	Ghana	3	2001–03	No	No	Twining and birth outcome
	Guinea-Bissau	1	2004	No	No	Vaccination
	Nigeria	2	1995–97	No	No	Twining and birth outcomes
Others	Brazil	1	2005	No	For zygosity determination	Dental health
Total		84		5	10	

or a cohort. All these five registers are being organized in Asia, four of them in Far East countries. All twin registers in Asia lack systematic collection of genomic information that can be linked with clinical/lifestyle information and common disease outcome assessments. This contrasts to the twin registers/cohorts in European countries and Australia, where systematic collection of genotype and phenotype information started in 1990s.[3]

In Korea, we have gathered very large-scale twin lists from the Korea National Health Service study (KNHS).[8] Based on this twin register, and their phenotypic information, the Korean government has now decided to support a new 'Healthy Twin' project which will help fill this information gap. It is based on a very large-scale population-based twin registers (154 783 pairs of twins) that currently has only limited phenotype information and disease outcomes.[8] Recently this register was successfully reconstructed by adding 900 642 family members of these twins. This study now has the potential to offer twin, family, and population-based study designs. Full details of the Korean Healthy Twin Study can be found at IJE Online.

Advantages of twin, family, population designs

Family-based study is generally more powerful than population-based study for detecting genetic variations responsible for a disease/trait. Twin and family designs can add a wide spectrum of study opportunities to conventional family study. Including dizygotic and monozygotic twins in the linkage analysis can improve the power, by contrasting a polygenic background from a shared familial environment.[9] Dizygotic twins can serve as ideal sibling pairs without the noise of age difference.[2] Because age-dependent gene expression exerts a large influence on almost all health conditions, statistical adjustment for age does not clear all the nuisance of the age difference. Dizygotic twin pairs were reported to have superior power to conventional sib-pairs in linkage analysis.[2]

Epigenetic and environmental causes and the Healthy Twin Study

Epigenetic mechanisms have attracted more attention in the aetiological mechanism of cancers, aging, and neuropsychiatric disorders.[10] Monozygotic twins discordant for phenotypes have particular power to detect epigenetic aetiology. Epigenetic variation, such as DNA methylation change, may have weaker signals than the genetic variation itself. Monozygotic twins, who have the same genetic variation, can maximize the chance to detect epigenetic effects. In addition twins and families, representative of general population, can be used to estimate the relative importance of genetic, environmental, and epigenetic effects.

'Family with twin' designs can provide a range of powerful studies on environmental factors, such as the above-mentioned co-twin control study. If genetic variants that interact with environmental factors are found, twin and family designs provide a powerful means of exploring the role of environmental factors with tight control of genetic variation. Given that genetic heterogeneity is an intrinsic aspect of common complex diseases, using nested twin, family, and population designs would provide the ideal means of teasing out the role of environmental factors and genetic variation. Moreover, results on associations and the magnitude of risk can be applied to preventive strategies with less uncertainty.

Recent studies on age-related macular degeneration have demonstrated that both family-based and population-based studies may be able to hunt disease genes, when the aetiology of diseases is rather homogeneous and the phenotypes can be

determined with high specificity.[11–13] However, unfortunately, most common complex diseases have shown evidence of aetiologic complexity. Given that multiple gene involvement, genetic heterogeneity/phenocopy, and gene–gene and gene–environmental interactions are the rule rather than the exception, it would be a natural approach to attempt to have as many dimensions of evidence as possible. Combined linkage, association, and functional study findings can much better convince researchers of the authenticity of discoveries. However, in the usual practice of building coherence, each piece of evidence comes from a different study of a different population, raising concerns about latent heterogeneity and uncertainty. The Healthy Twin study, if successfully accomplished, will certainly be able to provide multidimensional evidence. In addition, with combined analysis and collaboration with the existing register, population-based genomic cohort study, and similar studies from other countries, the study can provide appropriate data fit for the ever growing and demanding study needs.

The Korean Healthy Twin study is open to collaborations—designed and ethically approved to be so. Considering the scarcity of twin research in non-Caucasian populations, and the abundant research opportunities that hybrid twin, family, and population-based study can offer, we all hope the study will be able to serve as valuable assets for all epidemiological society, beyond its geographical and cultural limits.

Acknowledgements

This study was supported by the National Genome Research Institute, Korea, National Institute of Health research contract (budgets 2005-347-2400-2440-215). The views expressed in this paper are those of the authors and not necessarily that of any funding body.

References

[1] Glazier AM, Nadeau JH, Aitman TJ. Finding genes that underlie complex traits. *Science* 2002;**298:**2345–49.

[2] Martin N, Boomsma D, Machin G. A twin-pronged attack on complex traits. *Nat Genet* 1997;**17:**387–92.

[3] Peltonen L. GenomEUtwin: a strategy to identify genetic influences on health and disease. *Twin Res* 2003;**6:**354–60.

[4] Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. *Nat Rev Genet* 2002;**3:**872–82.

[5] Terwilliger JD, Weiss KM. Confounding, ascertainment bias, and the blind quest for a genetic 'fountain of youth'. *Ann Med* 2003;**35:** 532–44.

[6] Busjhan A. Twin research across the globe: What's out there in 2002? Guest Editorial. *Twin Res* 2002;**5:**v–vi.

[7] Bretz WA, Corby PM, Hart TC *et al*. Dental caries and microbial acid production in twins. *Caries Res* 2005;**39:**168–72.

[8] Sung J, Cho SH, Cho SI *et al*. The Korean Twin Registry—methods, current stage, and interim results. *Twin Res* 2002;**5:**394–400.

[9] Posthuma D, Boomsma DI. A note on the statistical power in extended twin designs. *Behav Genet* 2000;**30:**147–58.

[10] Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 2004;**429:**457–63.

[11] Klein RJ, Zeiss C, Chew EY *et al*. Complement factor H polymorphism in age-related macular degeneration. *Science* 2005;**308:**385–89.

[12] Edwards AO, Ritter R III, Abel KJ, Manning A, Panhuysen C, Farrer LA. Complement factor H polymorphism and age-related macular degeneration. *Science* 2005;**308:**421–24.

[13] Haines JL, Hauser MA, Schmidt S *et al*. Complement factor H variant increases the risk of age-related macular degeneration. *Science* 2005; **308:**419–21.

doi:10.1093/ije/dyi294
Advance Access publication 19 January 2006

Genes underlying common complex diseases

From JACOB PEEDICAYIL

I read with interest the article by Yang *et al*.[1] on the number of genes underlying the occurrence of common complex diseases in the population. They rightly stated that most common diseases are the result of complex interactions between multiple genes and environmental factors and that the identification of genes underlying these diseases represents a high public health priority since these diseases greatly contribute to the total public health burden.

The authors estimated the number of disease susceptibility genes needed to account for varying population attributable fractions (the proportions of disease cases in a population that would be prevented if an exposure was eliminated, assuming the exposure to be causal; PAF) of complex diseases. For simplicity of illustration the authors considered *N* independent biallelic disease susceptibility loci and that each genetic allele was assumed to have the same prevalence and risk ratio. They also assumed that there is only one at-risk genotype for each disease susceptibility locus. They used two models of interaction of the underlying genes: a purely additive model and a purely multiplicative model and assumed that these interactive effects are of the same magnitude for all genotypes involved.

I would like to make two comments on this article: First, a complexity involving common diseases unmentioned by the authors is epigenetics, which refers to heritable changes in gene expression that occur without any change in DNA sequence.[2] Epigenetics is known to involve three interacting molecular mechanisms: DNA methylation, modification of histones (DNA packaging proteins), and RNA-mediated gene silencing.[3] These mechanisms are known to be markedly influenced by the environment and are thought to play an important role in the pathogenesis of common diseases.[4] Epigenetic variants of a genetic allele have been referred to as epialleles.[5]

Department of Pharmacology, Christian Medical College, Vellore 632 002, India. E-mail: jpeedi@cmcvellore.ac.in