# The Korean Twin Registry — Methods, Current Stage, and Interim Results

Joohon Sung[1], Soo-Hun Cho[2], Sung-Il Cho[3], David L. Duffy[4], Jin-Hee Kim[1], Ho Kim[3], Kyung Sook Park[5], and Sue Kyung Park[5]

[1] *Department of Preventive Medicine, Kangwon National University, Korea*
[2] *Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea*
[3] *Department of Epidemiology and Biostatistics, Seoul National University School of Public Health, Seoul, Korea*
[4] *Epidemiology Unit, The Queensland Institute of Medical Research, Brisbane, Australia*
[5] *Department of Biology, Sungshin Women's University, Seoul, Korea*
[6] *Department of Preventive Medicine, Kunkook University, Chungju, Korea*

The Korean Twin Registry is the first nationwide twin study in Korea. We compiled 154,783 twin pairs from existing nationwide data sources, mainly from address and national health insurance data. The coverage of this registry was almost complete for the twins born since 1970, but less complete as age increased, so that there were only 990 pairs who were born before 1930. The twins' health examination ($N$ = 54,390 persons) and questionnaire ($N$ = 44,546 persons) results were incorporated into the registry, yielding 12,894 and 9074 concordantly informative pairs. Morbidity and mortality outcomes have been followed up since 1990, for most diseases. For preliminary analysis of complex diseases, we selected ventricular septal defects (VSD) in young twins, stomach and colorectal cancers in adult twins. We identified 353 VSDs, 284 stomach cancers, and 116 colorectal cancers among twins. The prevalence rates of cancers, but not that of VSD, were lower in twins than those in population. The difference in the cancer prevalence was marked for twins born before 1926, implying some degree of selection. Like-sex (LS) twins showed familial recurrence risks ($\lambda_{LS}$) of 41.2 for VSD and 22.4 for colorectal cancers, and 1.74 for stomach cancers. For opposite-sex (OS) twins, we could estimate $\lambda_{OS}$ of 19.8 for VSD only. These results were compatible with previous studies for VSD and colorectal cancers, but not for stomach cancers. Despite the strength in size, availability of health outcomes, and some lifestyle and basic laboratory data, we need accurate zygosity information to improve the validity of the results.

Twin studies have added to the knowledge of human disease, by discriminating genetic or environmental etiology (Hrubec & Robinette, 1984). Twin research has survived criticisms on potential biases and the lack of population representativeness (Phillips, 1993). Twin studies also have provided rich methodology for genetic dissection and gene mapping of complex human diseases (Martin et al., 1997). Although twin study has these specific strengths, most twin studies have been conducted in western countries or Caucasian populations (Boomsma, 1998). In Korea, data for medical and health research had not been systematically maintained until mid-1980s. As Korean National Health Insurance (KNHI) became obligatory in 1988, related agencies started keeping centralized nationwide databases. In addition, the national cancer registry was launched in 1987. Some recent epidemiologic studies utilized this wealth of data (Song & Sung, 2001; Suh et al., 2001). We had few good genetic studies about Korean population, despite relative abundance in genetic technology. Our primary motivation for a twin registry was to set up the basis for genetic studies, with an emphasis on complex human diseases or traits.

The Korean Twin Registry (KTR) is the first nationwide twin study in Korea. This study is population based, covering most age groups and is complete for twins born since 1979. It is an ongoing study, which started with the lists of twins from nationwide data sources. This approach is similar to those in Northern European countries (Kaprio et al., 1979; Kyvik, 1995; Pederson & Lichtenstein, 2000). We have been gathering information on almost full range of morbidity and mortality, mainly through nationwide medical utilization and death report data since 1990. This registry, with means of detecting most health outcomes, enables us to ascertain cases with higher efficiency and less ascertainment bias. We also incorporated the results of health examination and questionnaire of registered twins, mainly adults. Despite some strength in size, population representativeness, and case ascertainment, we have not yet conducted main surveys on the zygosity and genetic information.

We submit this report to describe the methods and present interim results, as well as the basis of the future results.

## Methods and Materials

### Data Sources

The lists of twins born in Korea were collected through matching several nationwide data sources longitudinally: address data of the Ministry of Internal Affairs, membership data of Medical Aid and Korea National Health Insurance

*Address for correspondence: Joohon Sung, Department of Preventive Medicine, Kangwon National University, 192–1 Hyoja2-Dong, Chuncheon, Kangwon-Do, Korea, 200–701. Email: sungjohn@kangwon.ac.kr*



https://doi.org/10.1375/twin.5.5.394 Published online by Cambridge University Press

(KNHI). The birth report data lacked 13-digit individual identification number (ID) and even the exact birthday, so that it could not be used for twin identification. This 13-digit ID system, given to all Koreans on birth report, consists of the date of birth (first 6 digits, year-month-day), gender (7th digit), area of birth (8th–11th digits), and two digits for checking duplication and check sum. This 13-digit ID system guarantees exact data matching among various sources. The KNHI now covers more than 95% of 45 million Koreans; Medical Aid covers the rest. The KNHI and Medical Aid membership log carries information on the exact familial relationships between the insured and their dependants, classified into more than a hundred different kinds. Although some families could split in two or even more parts depending on the number of employees in the family, longitudinal KNHI data provided rich sources of twin lists.

#### Identifying Younger Twins

Initially, twin list was compiled from nationwide address data, by confirming same birthday, family name, address code, and householder in 1996. The first twin list was restricted to those born since 1979. We have added new twins, who were confirmed from KNHI and medical aid membership data, to this initial list. We selected and confirmed twin pairs who share a birthday (allowing 1 day's difference), the same relationship with the same insured, and family name. We also identified some higher order multiples, but excluded them.

#### Identifying Adult Twins

The adult twins, who usually do not share addresses, were identified based on an algorithm in Figure 1. When we selected possible twin pairs who share birthday, birth area, and family name, there were about tenfold more candidates than expected. To select genuine twin pairs out of these candidates, we considered several factors: family name, first name and insurance number. In Korea, some family names are very common; big five family names comprise about 55% of all Koreans. The probability of sharing a family name by chance is 8%. The big five family names contributed most of this chance; all other family names contributed less than 0.6%. Note that women do not follow husbands' family names in Korea, thus adult opposite-sex twins could be identified based on their current family names.

We also considered whether the possible twin pairs shared part of their first names. A first name in Korea usually consists of two Chinese characters (or two syllables of pure Korean). Traditionally, one of the two first name characters is already settled by the patriarchs, and shared by the same generation of the clan, which is a part of paternally related families sharing a family name. For example, suppose a name "Dae-Jung Kim", where "Kim" is the family name and "Jung" is the prefixed name, his brother may have a name "Min-Jung Kim", and his

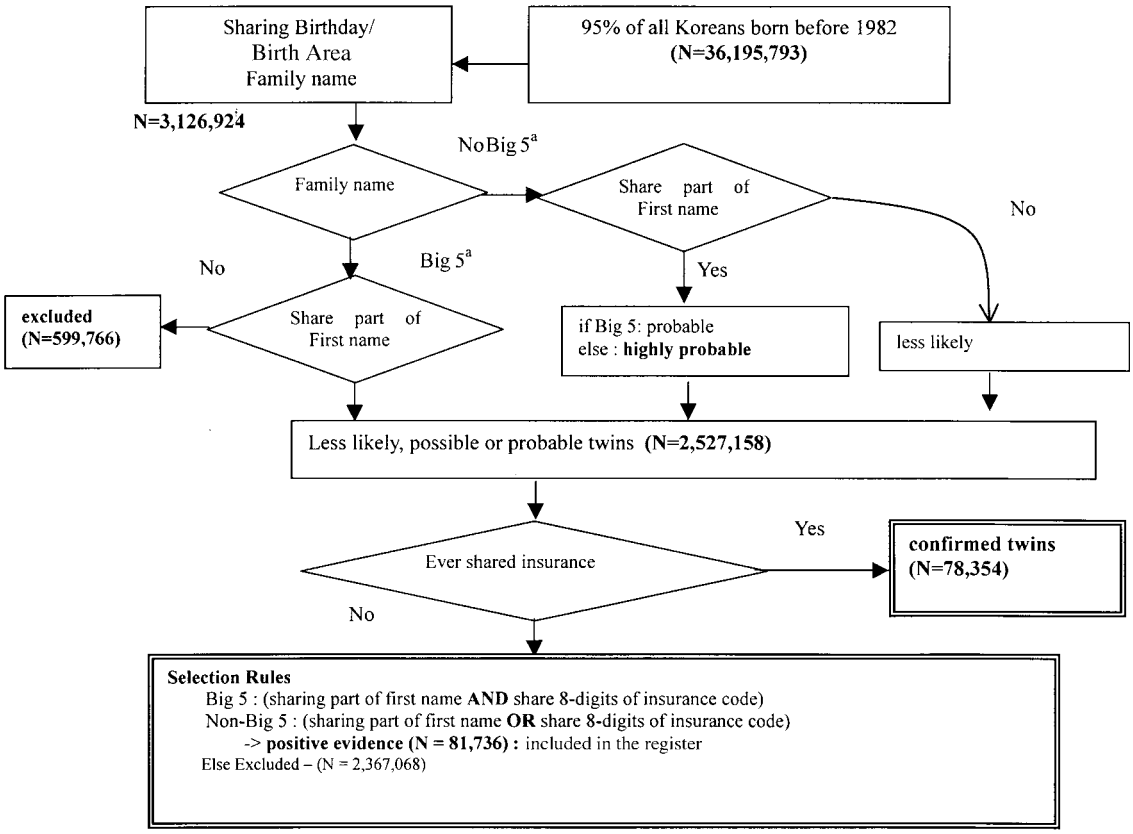


**Figure 1**

Algorithm for selecting adult twins in Korea.

Note: [a]Big 5 family names – Kim, Lee, Park, Choi, and Chung which comprise 55% of all Koreans



https://doi.org/10.1375/twin.5.5.394 Published online by Cambridge University Press

paternally related cousin may have a name "Gong-Jung Kim", while there would be no clue what would be the name of maternally related cousin.

And we considered whether the possible twin pairs have ever shared an insurance code in longitudinal KNHI data. Ever sharing the same insurance number means their being a family, therefore a twin; sharing first 8 digits out of 11-digit insurance number means their sharing detailed address, which was highly suggestive of their being a twin. We only included twins, who were confirmed or who had positive evidence of being a twin (Figure 1).

#### Case Ascertainment

We followed up the vital status of twins by matching their list with National Death Report Data between 1992 and 1999. KNHI also has kept medical claim data since 1990. By matching medical claim data of KNHI, we gathered almost full range of morbidity outcomes of the twins since 1990. Although the accuracy of KNHI medical claim data was not fully proved, we could sieve longitudinal data to get less the reliable part out. We selected three diseases to test the power to ascertain cases: ventricular septal defect (VSD) for younger twins, stomach cancer and colorectal cancer for adult twins. For VSD, we only selected cases detected more than twice, and ever admitted to general hospital, excluding the cases of outpatient visits or diagnosed in private clinics only. For cancers, National Cancer Registry data since 1987 were also available. We confirmed cancer cases, which were found in both medical claim and cancer registry data. We also included cancer cases, which were detected more than twice in general hospital, with a history of admission. This way of case ascertainment may reduce possible ascertainment bias, which was one of critical barriers for many registries. The international classification of diseases and injuries 9th and 10th revision (ICD-9, ICD-10) codes were: ICD-9 745.4 and ICD-10 Q21.0 for VSD; ICD-9 151 and ICD-10 C16 for stomach cancer; and ICD-9 153-154, ICD-10 C18-C21 for colorectal cancer.

#### Questionnaire and Health Examination

We gathered the biannual health examination and concurrent questionnaire data, which is also provided by KNHI free of charge. KNHI provides two kinds of health examination service: periodic examination for all employees or adults older than 40, and cancer screening tests. We elaborated detailed items of these examinations in Table 1. We have not been able to gather information for young twins yet.

#### Coverage of the KTR

The completeness of the KTR was estimated through a comparison with the estimated total twin birth from birth report data. The individual birth report data were available only since 1982 in Korea, and the information on multiple/single birth information was not accurate. Accordingly, we estimated the total count of twin births based on parental information, date of birth, and multiple/single birth information. For twins born before 1982, we considered that the twin birth rates were the same with the average twin rate between 1982 and 1984, applying this rate to total live births of the year.

#### Estimating Disease Concordance and Recurrence Risk

We first compared the prevalence in general population and the prevalence of twins by the age group and gender. Lacking zygosity information, we used the like-sex (LS) and opposite-sex (OS) structure for summarizing health outcome of twins. We did not use Weinberg's rule to estimate MZ-DZ concordance structure. We directly calculated population level prevalence with the same case definition and follow-up period as was used in the case ascertainment of twins. We needed this new calculation, because only incidence rates of diseases were available. We calculated case-wise concordance, and recurrence risk of diseases between LS ($\lambda_{LS}$) and OS ($\lambda_{OS}$). The $\lambda_{LS}$ and $\lambda_{OS}$ were calculated from case-wise concordance rates of LS and OS, divided by the prevalence in population (Risch, 1990). We restricted the age window, considering the disease incidence and follow-up period. We only included twins born

**Table 1**

Tests and Items in Health Examinations and Questionnaires Covered by Korea National Health Insurance

| Examination/Questionnaires | Items included |
|---|---|
| Shared tests in all health examinations | Past medical history, ABO blood type, height/weight, visual acuity, audiogram, systolic/diastolic blood pressure, hemoglobin, blood glucose (fasting), total serum cholesterol, serum liver enzymes (AST/ALT), hepatitis B virus antigen/antibody, urinalysis, chest X-Ray (indirect), Pap. smear (for women), dental caries and oral diseases, physical examination results. |
| Shared questions in all kinds of questionnaires | Subjective health status, family history, regularity of exercise, usage of supplementary medicine, regularity of diet, habit of eating salty food, meat consumption (not specifying the kinds), preference to vegetables, taking coffee or tea, alcohol drinking habit, smoking habit, recent weight change of more than 3 kg in a year, symptoms check list, past medical history. |
| Specific tests in cancer screening test | Upper GI study, Upper GI endoscopy, stool occult blood, colon study, sigmoidoscopy, colonoscopy, ultrasonography of abdomen, palpation examination, for breast, mammography, biopsy, summary results of tests taken. |
| Specific questions in cancer screening test | Body weight at 18, 25, 40s, and 50s, test items taken, motivation of taking examination, past medical history related to cancers, history of transfusion, food frequency tables, female reproductive history. |



https://doi.org/10.1375/twin.5.5.394 Published online by Cambridge University Press

since 1979 for analyzing VSD, and twins born before 1965 for cancers.

### Ethical Considerations

KNHI Corporation and National Statistical Office of Korea approved the use of data for study purposes. In Korea, although there is the law on protection of privacy in general, detailed guidelines of using public data for research purposes are being prepared. This study may need refinements of the ethical issues, upon the settlement of the guideline.

## Results

### Twin List

There were 154,783 twin pairs in the registry: 66,433 male LS pairs, 59,003 female LS pairs, and 29,347 OS pairs. The number of listed twins by the years of birth was summarized in Table 2. The KTR captured more than 95% of twins born after 1975. However, the list was less representative for those twins born before 1970. For some periods (1976–1985), the estimated coverage exceeded 100%. Lacking reliable figures of live births before 1970, we could not estimate twin birth rates and overall coverage. Judging from the LS to OS twin ratio, MZ twins rate had been about twice as high as DZ rate until 1980s. This finding was compatible with the results of Japan or Taiwan (Chen et al., 1987; Imaizumi, 1987).

### Health Examination and Questionnaire

There were 11,256 LS (7,360 male and 3,896 female pairs) and 1,638 OS pairs where both co-twins took the health examination, and 7,994 LS (5502 male and 2492 female pairs), and 1080 OS pairs that both filled in the questionnaire since 1986 (Table 3). We could gather slightly fewer data of oral hygiene examination results. However, there were only 32 and 14 pairs concordant for taking cancer screening tests and questionnaire. Since health examination service is only provided to employees or adults, it turned out that about 40% of adult twins took any kinds of health examination and questionnaire.

### Health Outcomes Follow-up

Health outcomes since 1990 were scanned for the registered twins (Table 4). The range of the identified health outcomes covered almost all diseases entities, except injuries, most infectious diseases, and minor health conditions. We could detect 353 VSD cases, 284 stomach cancer, and 116 colorectal cancer cases.

### Prevalence, Concordance, and Recurrence Risks

Prevalence rates (per 100,000 persons) in general population were 180, for VSD, 1,080 for stomach cancer, and 380 for colorectal cancer, respectively (Table 5). The age-standardized prevalence rates of twins were 222, 931, 310 for VSD, stomach cancer, and colorectal cancer. The prevalence in twins was slightly lower than that in general population. Stomach cancer and colorectal cancer showed difference in prevalence between two sexes, while VSD did not. The concordance of each disease for LS and OS was summarized in Table 5. The recurrence risks of diseases between LS were 41.2 (VSD), 22.2 (colorectal cancer), and 1.74 (stomach cancer). Recurrence risks between OS were calculated only for VSD (19.8), because there was no concordant OS pair for cancers.

## Discussion

### Possible False Twins in Adult Twin Lists

Although all young twins were confirmed, we might include some false adult twins, who were not genuine twins, but met the criteria. By our estimation, the chance probability of false twins was less than 1/10,000 (sharing one of first names and insurance code), yielding as much as 160 false twin pairs out of 3126 thousand of candidates. The estimated false twins may comprise less than 1% of adult twins.

**Table 2**

Summary of the Korean Twin Registry (KTR) by the Year of Birth, and Estimated Coverage of the KTR

| Year | 1901 –1920 | 1921 –1930 | 1931 –1940 | 1941 –1945 | 1946 –1950 | 1951 –1955 | 1956 –1960 | 1961 –1965 | 1966 –1970 | 1971 –1975 | 1976 –1980 | 1981 –1985 | 1986 –1990 | 1991 –1995 | 1996 –1999 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No. of pairs | | | | | | | | | | | | | | | | |
| Total | 226 | 764 | 2137 | 1424 | 1692 | 4199 | 8160 | 10,293 | 15,455 | 19,689 | 20,201 | 18,290 | 15,337 | 19,610 | 17,306 | 154,783 |
| Like Sex (Male) | 68 | 239 | 783 | 432 | 798 | 2074 | 3816 | 4832 | 6748 | 8390 | 8448 | 8103 | 6831 | 8215 | 6656 | 66,433 |
| Like Sex (Female) | 83 | 290 | 739 | 680 | 527 | 1447 | 3133 | 3747 | 5672 | 7422 | 8074 | 7348 | 6072 | 7580 | 6189 | 59,003 |
| Opposite Sex | 75 | 235 | 615 | 312 | 367 | 678 | 1211 | 1714 | 3035 | 3877 | 3679 | 2839 | 2434 | 3815 | 4461 | 29,347 |
| Estimated Twins from Birth report | NA[a] | NA | NA | NA | NA | NA | NA | NA | NA | 23,726 | 20,182 | 17,828 | 16,046 | 20,336 | 18,444 | |
| Coverage (in %) (identified/ birth report) | NA | NA | NA | NA | NA | NA | Less than 50% | Less than 70% | Less than 80% | 83.0[b] | 100.0[b] | 102.5[b] | 95.6 | 96.4 | 93.8 | |

Note: [a] Data not available before 1970

[b] The number of twin births was estimated for the years before 1982, assuming that twin birth rates were constant and same with average rate between 1982–1984.



https://doi.org/10.1375/twin.5.5.394 Published online by Cambridge University Press

**Table 3**

Collected Questionnaire and Health Examination Information in the Korean Twin Registry

| Number of pairs | Health Examination | Questionnaire Screening | Cancer Screening | Questionnaire for Cancer Screening | Oral Hygiene Examination |
|---|---|---|---|---|---|
| Like Sex | | | | | |
| Male | | | | | |
| Both completed | 7360 | 5502 | 25 | 10 | 5167 |
| One completed | 12,555 | 12,006 | 441 | 272 | 11,711 |
| None completed | 46,518 | 48,925 | 65,967 | 66,151 | 49,555 |
| Female | | | | | |
| Both completed | 3896 | 2492 | 2 | 1 | 2452 |
| One completed | 9697 | 8504 | 285 | 212 | 8167 |
| None completed | 45,410 | 48,007 | 58,716 | 58,790 | 48,384 |
| Opposite Sex | | | | | |
| Both completed | 1638 | 1080 | 5 | 3 | 1015 |
| One completed | 6350 | 5888 | 154 | 100 | 5435 |
| None completed | 21,359 | 22,679 | 29,188 | 29,244 | 22,897 |

**Table 4**

Disease Outcomes Available in the Korean Twins Registry

| Types of outcomes and sources | Disease entities | Confer |
|---|---|---|
| Morbidity | | |
| Cancer Registry (since 1987) | All cancers | Pathology attached (60%) |
| Medical Claim (since 1990) | Childhood viral infection, hepatitis, all cancers and benign tumors, anemia, congenital immune deficiency diseases, thyroid diseases, diabetes mellitus, obesity, endocrine disorders, all psychiatric disorders, neurologic disorders, muscular disorders, glaucoma, optic neuritis, hearing loss, hypertension, ischemic heart diseases, arrhythmia, congestive heart failure, cerebro-vascular attacks, chronic obstructive pulmonary diseases, asthma, restrictive lung diseases, pneumoconiosis and toxic lung diseases, peptic ulcers, hernia, inflammatory bowel diseases, chronic liver diseases, atopic dermatitis, psoriasis, rheumatoid arthritis, systemic connective tissue diseases, ankylosing spondylitis, herniated inter-vertebral discs, chronic renal diseases, all congenital malformations. | Accuracy matters for some diseases |
| Mortality | | |
| Death Report (since 1992) | All cause of death including accidents and injuries. | Complete/ Less accurate |

#### Coverage and Selection in Twin List

The KTR has been constructed from population data, not from volunteers or hospital registry. Therefore, selection of subjects according to their health status and self-referral was avoided.

The coverage of the registry decreased for those born before 1950, and this decrease was disproportionately large compared with the population structure. Because we started with the data of Korean population at 1990, we could not capture twins, whose co-twin had expired before 1990. To assess the possible effects from this selection, we compared the cancer prevalence between general population and twins (Table 6). After standardizing by age, observed to expected (O/E) ratios for stomach and colorectal cancers were 87.0% and 93.9%. The O/E ratios of both cancers were the smallest in twins born before 1925. This difference implies evidence of selection for old twins. In adult twins born after 1926, the O/E ratios of stomach and colorectal cancers were 91% and 97%. Judging from the differences in the cancer prevalence between singletons and twins, the effects of selection may be small for the twins born after 1926. Even admitting some degree of selection, unless the degree of selection was differential between MZ and DZ twins (between LS and OS twins in this study) the results from the registry would not be biased.

#### Power to Study Complex Diseases

One of the concerns about the twin research was the difficulty of getting sufficient numbers of twin subjects. This was especially true for rare diseases. In KTR, we could detect 116 cases of colorectal cancer, of which incidence rate was 23 per 100,000 person-years in 1999 (National Cancer Institute of Korea, 2000). Given accurate zygosity information, this size may suit analyses. However, we may need to follow-up longer period for very rare diseases, for example, less than 5 per 100,000 person-years of incidence rate.

#### Genetic Factors of the Diseases

Although being preliminary results, we found little evidence of genetic contribution to stomach cancers. This result was not compatible with the previous study of North European counties (Lichtenstein et al., 2000),



https://doi.org/10.1375/twin.5.5.394 Published online by Cambridge University Press

**Table 5**

Prevalence, Case-wise Concordance, and Recurrence Risk of Like-sex Twins (LS) and Opposite-sex Twins (OS) for Ventricular Septal Defects (VSD), Stomach Cancers, and Colorectal Cancers, in the Korean Twin Registry

| Results in Registry | Disease Status of Twin Pairs | | | Prevalence in Population ($P_P$) and Twins ($P_{TW}$), Case-wise Concordance (CCW), Recurrence Risk ($\lambda$) | | |
|---|---|---|---|---|---|---|
| | VSD | Stomach Cancer | Colorectal Cancer | VSD | Stomach Cancer | Colorectal Cancer |
| Twins | | | | | | |
| Like Sex (All) | | | | | | |
| Both affected | 11 | 2 | 4 | $P_P = 0.0018$ | $P_P = 0.0108$ | $P_P = 0.0038$ |
| One affected | 275 | 208 | 87 | $P_{TW} = 0.0022$ | $P_{TW} = 0.0093$ | $P_{TW} = 0.0031$ |
| None affected | 64,040 | 23,478 | 23,607 | | | |
| Male Like Sex | | | | | | |
| Both affected | 5 | 2 | 2 | | | |
| One affected | 138 | 143 | 59 | | | |
| None affected | 33,381 | 12,897 | 12,991 | $CCW_{LS} = 0.0741$<br>$\lambda_{LS} = 41.2$ | $CCW_{LS} = 0.0188$<br>$\lambda_{LS} = 1.74$ | $CCW_{LS} = 0.0842$<br>$\lambda_{LS} = 22.2$ |
| Female Like Sex | | | | | | |
| Both affected | 6 | 0 | 2 | | | |
| One affected | 137 | 65 | 28 | | | |
| None affected | 30,659 | 10,581 | 10,616 | | | |
| Opposite Sex | | | | | | |
| Both affected | 1 | 0 | 0 | | | |
| One affected | 54 | 72 | 31 | $CCW_{OS} = 0.0357$<br>$\lambda_{OS} = 19.8$ | $CCW_{OS} = NA$[a]<br>$\lambda_{OS} = NA$ | $CCW_{OS} = NA$<br>$\lambda_{OS} = NA$ |
| (Male affected) | (23) | (45) | (20) | | | |
| (Female affected) | (31) | (27) | (11) | | | |
| None affected | 15,035 | 5135 | 5176 | | | |

Note: [a] Not available (none of concordantly affected pairs).

**Table 6**

Prevalence of Stomach and Colorectal Cancers in General Population and Twins

| Year of birth | Total No. | Stomach cancer | | Colorectal cancer | |
|---|---|---|---|---|---|
| | General Population (estimated at 1995) | No. Cases in population (Prevalence per $10^5$)[a] | Expected twin Cases[b] (LS/OS)[b] | No. Cases in population (Prevalence per $10^5$) | Expected twin Cases+ (LS/OS) |
| –1925 | 1,597,000 | 64,371 (4031) | 43.1 (29.5/13.5) | 21,230 (1329) | 14.2 (9.7/4.5) |
| 1926–35 | 2,540,000 | 70,976 (2794) | 74.2 (52.6/21.6) | 23,379 (920) | 24.4 (17.3/7.1) |
| 1936–45 | 3,980,000 | 57,324 (1440) | 76.4 (57.8/19.6) | 20,715 (520) | 28.0 (20.9/7.1) |
| 1946–55 | 5,541,000 | 28,636 (517) | 60.7 (50.1/10.8) | 11,815 (213) | 25.1 (20.7/4.5) |
| 1956–65 | 8,375,000 | 16,429 (196) | 72.3 (60.9/11.5) | 7,212 (86) | 31.8 (26.7/5.0) |
| Total (–1965) | 22,033,000 | 237,735 (1079) | 326 (250/77.0) | 84,352 (383) | 123.5 (95.4/28.2) |
| | Total No. of Twins (LS/OS)[c] | No. Cases in Twin (LS/OS) | Observed/expected (LS/OS) | No. Cases in Twin (LS/OS) | Observed/expected (LS/OS) |
| –1925 | 1070 (732/ 335) | 26 (20/6) | 60.3% (67.8/44.4) | 9 (6/3) | 63.3% (61.7/67.4) |
| 1926–35 | 2654 (1882/ 772) | 79 (55/24) | 106.5% (105 /111) | 21 (12/9) | 86.0% (69.3/127) |
| 1936–45 | 5302 (4014/1364) | 76 (54/22) | 99.5% (93.4/112) | 22 (15/7) | 78.6% (71.8/98.6) |
| 1946–55 | 11,739 (9692/ 2090) | 43 (33/10) | 70.9% (65.9/92.6) | 18 (12/6) | 71.6% (58.1/135) |
| 1956–65 | 36,846 (31,056/ 5860) | 60 (50/10) | 83.0% (82.1/87.0) | 46 (40/6) | 145 % (149/119) |
| Total (–1965) | 57,611 (47,376/10,421) | 284 (212/72) | 87.0% (84.5/93.5) | 116 (85/31) | 93.9% (89.1/110) |

Note: [a] Cumulative Prevalence between 1990–1999, the same case definition was used for singleton and twins
[b] Calculated from (prevalence in general population) x (number of twins) of each strata
[c] LS-like sex twins, OS-opposite sex twins.



https://doi.org/10.1375/twin.5.5.394 Published online by Cambridge University Press

where 28% of heritability was reported. Stomach cancers are most prevalent malignant neoplasm in Korea, having incidence rates of 68 in men, and 32 in women per 100,000 person-years. This may be a true result, implying that excess environmental risks, such as *Helicobacter pylori* infection, caused higher stomach cancer incidence in Korea. This result could derive from the incompleteness of old twins. Considering the results of colorectal cancers, the effect from selection might not have been large, if existed.

#### Validity and Completeness of Case Ascertainment

We relied on existing population morbidity and mortality data for case ascertainments. Biases related to self-reporting or the differential case ascertainment between MZ and DZ could be avoided. Accordingly, the validity and completeness of case ascertainment reflects those characteristics of the data sources. Although KNHI medical claim data is relatively complete, the accuracy of the diagnosis matters. Generally, the accuracy of more serious diseases is higher. We had investigated the accuracy of KNHI data for oral cleft and lung cancer. We found 77% cases (89/116) of lung cancers and 96% oral clefts (78/81) were confirmed from medical records. Actual accuracy would be slightly higher, because this did not count probable (not biopsy proven) cases. Accuracy of KNHI data for congenital malformations and cancers was estimated to be high. To increase the accuracy, we refined the cases definition, to exclude medical claims for routine check-up. Cancer registry data usually have pathologic diagnosis. For other diseases other than cancers, we may need to confirm the cases through medical record review or direct contact. The KTR can also provide a basis for this confirmation by scanning probable cases.

## Conclusion

The KTR has strength in size, representativeness, and case ascertainment. This registry is one of the few population-based twin researches in Asian population. We selected some complex human diseases as a pilot study. We identified some VSD and cancer cases for analyzing genetic factors. However, we found some evidence of selection for old twins, which could affect the results in certain conditions. Currently, the lack of zygosity information is major limit of this registry. The KTR is open to domestic as well as international collaboration.

## Acknowledgment

This study was supported by the Ministry of Health and Welfare of Korea, Grant HMP-99-M-09-0007.

## References

Boomsma, D. (1998). Twin registers in Europe. *Twin Research, 1,* 34–51.

Chen, C., -J., Lin, T., -M., Chang, C., & Cheng, Y. -J. (1987). Epidemiological characteristics of twinning rates in Taiwan. *Acta Geneticae Medicae et Gemellologiae, 36,* 335–342.

Hrubec, Z., & Robinette, D. (1984). The study of human twins in medical research. *New England Journal of Medicine, 310*(7), 435–441.

Imaizumi, Y. (1987). The recent trends in multiple births and stillbirth rates in Japan. *Acta Geneticae Medicae et Gemellologiae, 36,* 325–334.

Kaprio, J., Koskenvuo, M., Artimo, M., Sarna, S., & Rantasalo, I. (1979). Section 1. Materials, methods, representativeness and results for variables special to twin studies. *The Finnish Twin Registry: Baseline characteristics.* Helsinki: Kansanterveystieteen julkaisuja M47.

Kyvik, K., O., Green, A., & Beck-Nielsen, H. (1995). The New Danish Twin Register: Establishment and analysis of twinning rates. *International Journal of Epidemiology, 24*(3), 589–596.

Lichtenstein, P., Holm, N., V., Verkasalo, P., K., Iliadou, A., Kaprio, J., Koskenvuo, M., Pukkala, E., Skytthe, A., & Hemminki, K. (2000). Environmental and heritable factors in the causation of cancer — Analyses of cohorts of twins from Sweden, Denmark, and Finland. *New England Journal of Medicine, 343*(2), 78–85.

Martin, N., Boomsma, D., & Machin, G. (1997). A twin-pronged attack on complex traits. *Nature Genetics, 17,* 387–392.

National Cancer Institute of Korea (2000). *Annual reports on cancer incidence 1999* (in Korean). Seoul: Author.

Pedersen, N., & Lichtenstein, P. (2000). The Swedish Twin Registry. A presentation. In B. Smedby, I. Lundberg, & T. I. A. Sørensen (Eds.), *Scientific evaluation of the Swedish Twin Registry.* Stockholm: Swedish Council for Planning and Coordination of Research.

Phillips, D. I. W. (1993). Twin studies in medical research: Can they tell us whether diseases are genetically determined? *Lancet, 341,* 1008–1009.

Risch, N. (1990). Linkage strategies for genetically complex traits. I. Multilocus models. *American Journal of Human Genetics, 46,* 222–228.

Song, Y. M., & Sung, J. (2001). Body mass index and mortality: A twelve-year prospective study in Korea. *Epidemiology, 12,* 173–179.

Suh, I., Jee, S. H., Kim, H. C., Nam, C. M., Kim, I. S., & Appel, L. J. (2001). Low serum cholesterol and haemorrhagic stroke in men: Korea Medical Insurance Corporation Study. *Lancet, 357,* 922–925.