

The Cuban Twin Registry: Initial Findings and Perspectives

Beatriz Marcheco-Teruel,¹ Marcia Cobas-Ruiz,² Niviola Cabrera-Cruz,² Araceli Lantigua-Cruz,¹ Elsa García-Castillo,¹ Roberto Lardoezt-Ferrer,¹ Zoe Robaina-Jiménez,¹ Evelyn Fuentes-Smith,¹ Francisco Morales-Calatayud,³ María Teresa Lemus-Valdés,⁴ Miriam Portuondo-Sao,² Lenier Comas-Pérez,⁵ Juan M. Pérez-Crispi,⁵ Thais Díaz-De Villal Villa,⁶ Emelia Icart-Perera,⁷ Aida Jordán-Hernández,⁸ Agustín Lage-Castellanos,⁶ Sergio Rabell-Piera,² Juan de Jesús Llibre-Rodríguez,³ Pedro Valdés-Sosa,⁶ and Mitchell Valdés-Sosa⁶

¹National Centre of Medical Genetics, Havana, Cuba

²Ministry of Public Health, Havana, Cuba

³Medical University of Havana, Havana, Cuba

⁴Municipal Center of Medical Genetics, Plaza's Municipality, Havana, Cuba

⁵Municipal Center of Medical Genetics, Caibarien's Municipality, Villa Clara, Cuba

⁶Cuban Centre for Neuroscience, Havana, Cuba

⁷Medical and Psychopedagogic Center, 'La Castellana', Havana, Cuba

⁸Medical and Psychopedagogic Center, 'Benjamin Moreno', Havana, Cuba

The Cuban Twin Registry is a nation-wide, prospective, population-based twin registry comprising all zygosity types and ages. It was initiated in 2004 to study genetic and environmental contributions to complex diseases with high morbidity and mortality in the Cuban population. The database contains extensive information from 55,400 twin pairs enrolled in the period 2004–2006. Additionally, 2,600 new multiple births have been included from 2007 to date. In the past 4 years, more than 130 studies have been carried out using the registry with a classical genetic epidemiological approach in which concordance rates for monozygotic and dizygotic twins and heritability of various disease traits were estimated. This article summarizes the history, registry's methodology, recent research findings, and future directions of work.

■ **Keywords:** Cuban Twin Registry, twin studies, twin registries, genetic registry

For decades, twin studies have made important contributions to understanding the role of genetic and environmental factors on health. They are essential for developing treatment and prevention of disease in humans. Unfortunately, only a few developing countries have reported twin registries. Even if they exist, they rarely include, in a unique database, the majority of twins in the country. In Latin America, only two countries, Chile and Cuba, have been reported to have twin registries (van Dongen et al., 2012).

We here describe in detail the Cuban Registry. This registry targets the population of Cuba, the largest island in the Caribbean Sea, with a population of 11,244,543 inhabitants. Of these, about 75.2% of the population lives in urban areas and about 17.9% is older than 60 years of age. The widest section of the country's population pyramid is in ages between 40 and 50. Life expectancy at birth is 77.97 years. Females live longer than males by 4.02 years on average; however, there are approximately equal numbers of females

and males in the country (Oficina Nacional de Estadísticas de Salud, Ministerio de Salud Pública, 2012).

As a national project, the Cuban Twin Registry (CTR) was initiated in June 2004 and concluded its first stage in April 2006. Information was obtained on 55,400 pairs of twins, including all zygosity variants and all ages. During the subsequent 4 years, the registry has become a useful tool for performing genetic epidemiological studies in complex diseases that present high morbidity and mortality in the Cuban population.

RECEIVED 30 August 2012; ACCEPTED 1 November 2012. First published online 11 December 2012.

ADDRESS FOR CORRESPONDENCE: B. Marcheco-Teruel, Centro Nacional de Genética Médica, Ave 31, Esq a 146 Nro. 3102, Reparto Cubanacan, Playa, La Habana, Cuba. E-mail: beatriz@infomed.sld.cu

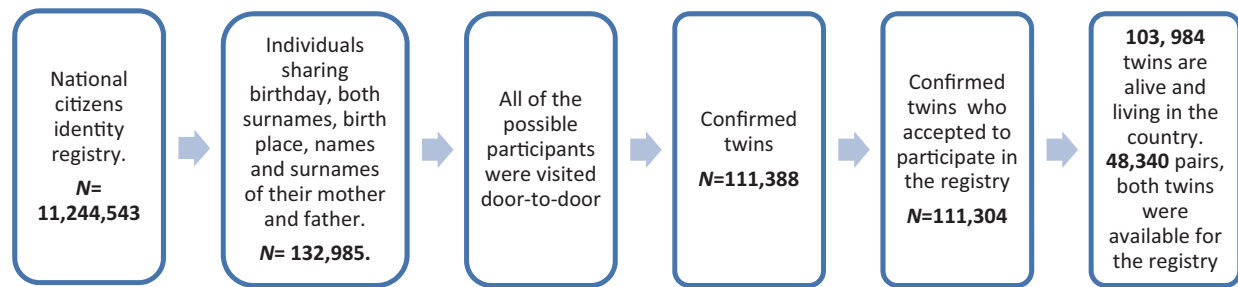


FIGURE 1
(Colour online) Algorithm for selecting twins in Cuba (CTR).

General Aim

The CTR is a resource designed to facilitate research aimed at elucidating the role of genetic and environmental factors in the etiology of complex multi-factorial diseases, taking advantage of the possibilities of classical twin studies, and at the same time allowing assessment of optimal strategies for diagnosis and prevention of such diseases in the Cuban population.

The Initial Database of the CTR

With the support of government authorities, the national citizen identity registry provided a list of potential twins based on the following criteria: individuals with the same surnames; born at the same day, month, and year and in the same place; with identical names and surnames of their mother and father (Figure 1). The information also included their home address and national identification number. The citizen's identity card in Cuba includes all of the information mentioned above.

In Cuba, women do not change their surnames after marriage. At birth, each individual is registered with two surnames: the first is the father's and the second is the mother's, except in a very small number of persons with unrecognized paternity, in which both surnames are almost always derived from the mother. The two surnames usually do not change in the individual's lifetime. This practice allows one to establish with relative ease, genealogical relationships among people (e.g., to draw a pedigree), or as in our case, to define an initial universe, which could include most of the twins in the country. The data fields described for all subjects meeting the required criteria was organized into the initial database of the CTR comprising all candidates for the registry.

Recruitment of the Twins

From the database of candidates, lists were drawn up for each province and municipality containing the names and surnames, national identification number, date of birth, and personal address ($N = 132,985$). This allowed door-to-door visits by trained personnel who were primarily medical

students, doctors, nurses, psychologists, and social workers, who were supported by the health authorities and the local government of the territories.

At the initial interview each candidate received an explanation of the goals of the registry and was invited to participate in the project. In the case of children under 18 years and persons unable to give an informed consent, this was obtained from their legal representatives. The percentage of acceptance for participation in the study was over 99%.

Ethical Issues

All issues related to the CTR, as well as any research project involving it, must be approved by the Scientific Council and the Ethics Committee of the National Centre of Medical Genetics, which is the national reference center for the Cuban program of diagnosis, management, and prevention of genetic diseases and birth defects. The national project to create the registry was authorized by the Ministry of Public Health.

In all cases, after agreeing to join the registry, either the twins or their representatives signed an informed consent form. Furthermore, the staff that conducted house visits and interviews had signed a confidentiality agreement with a commitment to not reveal any information relating to the twins and/or their families obtained during the project.

Data Collection

Each twin or his/her legal representative was invited to an interview in order to answer an extensive baseline questionnaire containing 171 questions organized in 17 sections that related to four broad areas:

1. Twin-related information, including (a) data relevant for zygosity determination and for establishing if the pair had been reared together or not; (b) data concerning conception, obstetric history, childbirth, family's social characteristics, other multiple births in the family; and (c) data about presence of genetic diseases and birth defects.

2. Specific health status of the twins.
3. Presence of several types of diseases as well as their related risk factors. Diseases studied were cancer, diabetes, hypertension, ischemic heart disease, bronchial asthma, epilepsy, mental retardation, Parkinson's, dementia, schizophrenia, bipolar disorder, alcohol abuse or dependence, and stroke. These are diseases with a late onset; it was felt that their inclusion at the creation of the registry will be of considerable prospective value.
4. For school age twins an additional questionnaire was used to explore problems of inattention and hyperactivity and specific learning disorders.

All persons who agreed to participate in the study answered all the questions. Whenever possible, photographs of the pair, especially those of monozygotic (MZ) twin pairs, were taken and the images stored in the database.

When twins lived in two different municipalities and/or provinces at the time of the interview, the first co-twin contacted was asked to provide the data regarding location of the other twin. The latter was later interviewed at his/her home. The information from the questionnaires was uploaded to a database system specifically designed by the University of Informatics of Havana for this study. Access interfaces to facilitate structured query language (SQL) data queries were also developed for the CTR.

Zygoty Determination

Zygoty was determined with a questionnaire of 24 items regarding physical similarities and the frequency of identity confusion by family members, relatives, and other people. To establish a classification algorithm, the zygoty was independently established using blood group tests and human leukocyte antigen (HLA) antigens identification for a subset of 188 same-sex twin pairs. This method has been reported to have a 90–98% correct classification rate (Cohen et al., 1975; Goldsmith, 1991; Nichols & Bilbro, 1966). The classification weight of each item's answer was computed as a chi-squared statistic between MZ and dizygotic (DZ) twins, as this statistic measures departures from a random distribution between groups. Those items having a low chi-squared value were considered to have poor classification value.

The final score of each pair is the sum of all the chi-square statistic of the corresponding questionnaire. Cut-off points were established on this score using receiver operating characteristic curve (ROC) analysis. Using this algorithm we correctly classified 86.21% of MZ pairs and 100% of DZ pairs in the test subset, whereas the global rate of correct classification was 92.5%.

The Demographic Characteristics of the CTR

A total of 132,985 individuals were visited at their homes in the period 2004–2006, all over the country. Of these

TABLE 1
The Age Distribution of Twins in CTR

Age group (years)	%
<1	1.36
1–4	7.62
5–14	20.31
15–19	10.14
20–29	15.27
30–39	16.69
40–49	12.47
50–59	7.99
60–69	4.96
70–79	2.32
80–89	0.76
90–99	0.11
+100	0.01
Total	100.00

individuals, 111,304 cases were confirmed as being born in a multiple birth, whereas only 84 individuals (0.08%) refused to participate in the project. With a population of just over 11 million inhabitants, the rate of twin pairs in the general population of Cuba was found to be 5.03 per 1,000 individuals.

The CTR database contains information on 55,400 pairs of twins of all ages, of which 391 pairs are triplets; 103,984 twins (48,340 pairs) are alive and living in the country, 835 twins live outside the country, and 6,401 co-twins are deceased.

The zygoty analysis revealed that 14,981 pairs (27.04%) of the twins were MZ, and 32,843 pairs (59.28%) were DZ, while the remaining 7,576 (13.67%) pairs were uncertain for zygoty; 7,885 of the MZ pairs were female and 7,096 were male, with a female–male ratio of 1.1:1. In DZ twin pairs of same sex, the ratio is similar; 6,797 were female and 6,343 were male (1.07:1). In the DZ pairs 60% of were of opposite sex. When inquiring about the presence of other multiples in the family, 20,111 pairs (36.3%) reported a positive history, 28.8% were MZ, 57.6% DZ, and 13.6% of uncertain zygoty.

Longevity in Twins

Among twins who live in Cuba, there are 83 aged between 90 and 94 years, 33 aged between 95 and 99 years, and 13 aged over 100 years (Table 1).

Disease Records of the Twins in the CTR

Genetic diseases in twins. Of the registered twins, 2,782 individuals were reported to have genetic disorders (2.7%). Of these, 751 pairs reported that both members were affected by a genetic disease. The most frequently found disease was sickle cell anemia, with 404 affected twins, and 120 pairs were concordant (51 MZ, 69 DZ). The next most frequent genetic disorder was neurofibromatosis, found in 154 twins, and 49 pairs were concordant (15 MZ, 44 DZ).

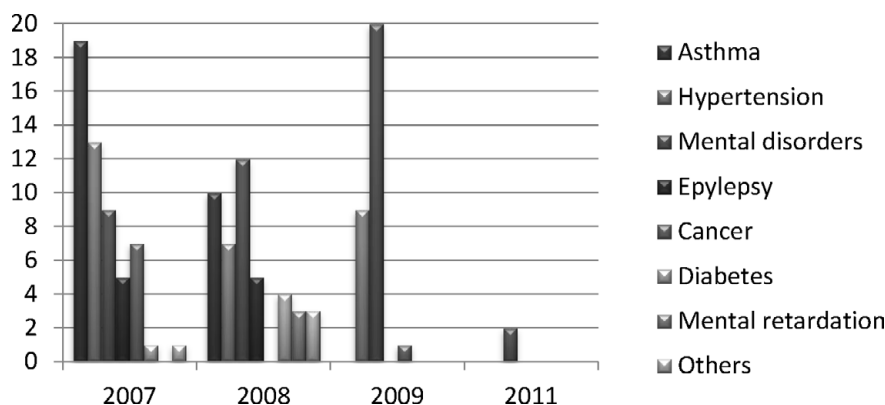


FIGURE 2

Studies conducted with the Cuban Twin Registry by year and disease. 2007–2011.

Down syndrome affected 125 twins, of which 39 pairs were concordant (13 MZ, 26 DZ).

Birth defects. Within the registry, 4,448 twins were found to have a birth defect. Of these, 893 pairs showed that both members had the same defect. The most prevalent problem was congenital heart defects, followed by defects in the feet or face.

Common diseases with complex etiology. Among common diseases, bronchial asthma was the most frequent; 14,476 (13.9%) twins were affected and 4,664 pairs were concordant (2,053 MZ, 2,611 DZ). This prevalence rate was somewhat higher than that (10%) found in the general Cuban population (Oficina Nacional de Estadísticas de Salud, Ministerio de Salud Pública, 2012). In order of frequency, at the time that twins were recorded and specific information about health status of twins was collected, asthma was followed by hypertension with 9,922 twins affected and 2,095 twins were concordant (754 MZ, 1,321 DZ); 2,258 twins reported ischemic heart disease and in 256 pairs, twins were concordant (104 MZ, 152 DZ); 2,278 twins were diagnosed with depression and in 462 pairs twins were concordant (185 MZ, 277 DZ).

Cancer. One thousand one hundred ninety-two twins were reported to suffer from different types of cancer. The most frequent cancer in women was cervical cancer, followed by breast cancer, the colon in both sexes, and the prostate in man. In 243 pairs both members had the same type of cancer.

Neurological disorders. One hundred eight twins were diagnosed with Parkinson's disease, and nine pairs of twins were concordant (4 MZ, 5 DZ); 426 twins were diagnosed with dementia, 43 pairs twins were concordant (13 MZ, 30 DZ); and 631 were reported to have suffered from stroke, of which 45 twin pairs were concordant (18 MZ, 27 DZ).

Summary of Recent Research

Studies of Physical Diseases Based on the CTR

In 2007 the twin registry was started with an aim to study classical genetic epidemiology studies to determine concordance and heritability for various diseases and for extended twin-family designs to investigate familial aggregation of complex disorders. More than 130 studies have been performed since then (Figure 2).

The National Centre of Medical Genetics offers a post-graduate master program in genetic counseling. Every year about 80 students from different parts of the country and abroad take the course. All students are offered the opportunity to carry out research for their thesis using the CTR. One of the goals of this research is to produce results that can be later used in genetic counseling practice. This strategy has the additional benefit that a systematic update of data concerning the twins' health is achieved. Additionally, information about twins' families is obtained to expand the data in the registry.

The results of these investigations have been introduced in genetic counseling practice at medical genetics services located in the primary healthcare level as part of the strategies of the National Health System for Disease Prevention.

Asthma. A total of 1,487 twin pairs with asthma from 12 provinces were studied: 529 MZ twin pairs and 958 DZ pairs. It was 4.5 times more common to find that both twins of a MZ pair were affected than in the case of a DZ pair. The heritability was 59% (Table 2). The analysis of family history of asthma in both groups of twins showed that asthma was 3.1 times more frequent in first-degree relatives of MZ twins than DZ twins: odds ratio (OR) = 3.1 95% confidence interval (CI; 1.78: 5.42) and 1.5 times more frequent among second-degree relatives: OR = 1.5 95% CI (1.06: 2.09). Regarding risk factors triggering the disease, climate change and exposure to dust were the most frequent. The allergic symptoms most frequently found to be associated with asthma were rhinitis and atopic dermatitis. The non-genetic

TABLE 2
MZ and DZ Twin Concordance and Heritability Estimates for Complex Diseases

Trait	Number of twin pairs	Probandwise concordance%		Heritability		
		MZ twins	DZ twins	h ²	C	E
Asthma	1,487	81.3	51.7	0.59	0.22	0.19
Hypertension	1,207	72.2	52.5	0.39	0.33	0.28
Coronary heart disease	256	24.7	8.2	0.33	–	0.75
Schizophrenia	114	64.7	23.4	0.83	–	0.35
Major depression	573	72.1	45.3	0.54	0.18	0.28
Bipolar disorder	118	59.7	28.9	0.62	–	0.4
Alcohol abuse or dependence	309	75.1	46.6	0.57	0.18	0.25

Note: h² = heritability; C = shared environmental effects; E = non-shared environmental effects including measurement error.

risk factors most frequently associated with the disease were: history of recurrent respiratory infections, living in urban areas, and exposure to tobacco smoke.

Hypertension. A total of 1,207 twin pairs from eight provinces were studied: 437 MZ pairs and 770 DZ pairs. The results showed that concordance in MZ twins was 2.35 times larger compared with DZ twins. The heritability was found to be 39%. The non-genetic risk factors that showed a stronger association with the disease were excessive consumption of salt in the diet, exposure to stress, cigarette smoking, and physical inactivity, in that order. The family history of hypertension was 1.6 times more common in MZ than in DZ twins. The most frequent age at disease onset was between 30 and 39 years in both groups, and the time lag in the onset of the disease among siblings was 3.5 years on average in MZ, and 4.7 years in DZ.

Studies of Mental Disorders Based on the CTR

Major depression. A total of 573 twin pairs have been studied over the past 4 years: 211 MZ and 362 pairs of DZ. Concordance in MZ twins was 3.1 times more frequent than for DZ twins. The estimated heritability was 54%. A history of depression among first-degree relatives is 1.4 times more frequent in MZ twins than in DZ. When exploring mental disorders in other family members, a history of alcohol abuse or dependence, anxiety disorders, and bipolar disorder were the most frequent comorbid conditions found among first- and second-degree relatives. With regards to the age of disease onset, in both MZ and DZ groups the frequency of occurrence was high in the second and third decades of life. Strikingly, the frequency of the onset of the disease after 40 years was 2.3 times higher in DZ than in MZ twins. A history of loss or separation from a beloved family member was the non-genetic risk factor that showed the strongest association with the onset of depression.

Alcohol abuse and dependence. A total of 309 twin pairs from nine provinces have been studied. For MZ twins it was 3.5 times more common than for DZ to find concordance for the disease. The estimated heritability was 57%. MZ

twins showed a history of alcohol abuse in first-degree relatives that was 1.5 times more frequent than in DZ twins. Depression, anxiety disorders, and obsessive-compulsive disorder were the most frequent comorbid mental illnesses that appeared in first- and second-degree relatives of patients.

Bipolar affective disorder. A total of 118 twin pairs from nine provinces have been studied. The concordance rate was 3.6 times greater for MZ twins than DZ twins. The estimated heritability was 62%. It is also 2.2 times more common for MZ twins to have first- or second-degree relatives also affected than DZ twins. The specific patterns of comorbidity among first- and second-degree relatives varied for each sex. For female first- and second-degree relatives, it is more common to find bipolar affective disorder, depression, and anxiety disorder. The pattern is different for men predominantly alcohol addiction, bipolar disorder, and depression.

Schizophrenia. A total of 114 twin pairs were studied: 46 MZ and 68 DZ. MZ twins showed a concordance for the disease that was six times larger than in DZ twins. The estimated heritability was 83%. First-degree relatives of MZ twins suffered from schizophrenia at a rate twice that of same-degree relatives of DZ. Mental disorders such as schizophrenia, depression, and anxiety disorder, in that order, were the most frequent associated disorders among first- and second-degree relatives in both groups.

Twin Registry Update

About 2,655 newborn twin pairs have been included in the registry during the last 5 years (2007–2011). In Cuba, 65 hospitals nation-wide provide coverage to 97% of the total births in the country and there are around 125,000 births per year (Oficina Nacional de Estadísticas de Salud, Ministerio de Salud Pública, 2012). A national registry for birth defects for Cuba was created in 1985. Since then, at all the hospitals where births occur, a neonatologist is responsible for reporting children with birth defects and multiple births. This has allowed the prospective tracking of all births, generating information for both the national birth defect registry and the national twin registry.

Furthermore, in the first 3 months of life, every child is taken to a consultation with a genetic counselor at the community medical genetic service. The genetic counselor performs a detailed physical examination and explores the family history, looking for the presence of genetic diseases, and confirms and reports information data on multiple births, enabling the possibility of long-term follow-up as well as to draw the pedigree of the family, including information about other multiple births and health status of family members. At the same time, this procedure functions as a quality control mechanism for hospital reports.

Concluding Remarks

The Cuban Registry of Twins is the largest reported in Latin America and the Caribbean, and the only one that has included most of the twins available in the total population, with an acceptance rate of 99.9% among individuals initially contacted.

The data presented in this article is the first report after the analysis of the extensive database available that has an inestimable value for future research on the role of genetic and environmental factors in the origin of diseases with high morbidity and mortality in the Cuban population.

Regardless of the extensive work carried out over the past 5 years, the registry is still in its early stages. The main advantage of registration, which is to facilitate access to the twins for their participation in different types of studies, has been used in epidemiological investigations based on classical studies of concordance and heritability.

In future, we expect to work on elucidating more accurately the genetic risk factors involved in genetic and complex disorders, including DNA analysis and other laboratory studies, as well as the design of new and more specific research related to the role of non-genetic factors and gene-environment interactions. We will continue to expand from using twin pairs to twins' families, including twins' parents,

siblings, and offspring, with the purpose of obtaining more complex family designs for genetic epidemiology and further laboratory research on the gene susceptibility in complex traits.

This resource creates exciting opportunities for the Cuban and the international scientific community, which enables a wide range of research on epidemiological studies of genetic and environmental factors involved in the origin of complex diseases and other traits including human behavior, all of which will contribute significantly to the future advancement of science in our country.

Acknowledgments

We are grateful for the large contribution of the Cuban Centre for Neuroscience to the registry's methodology; the University of Informatics and other National Institutes of the Ministry of Public Health; and the medical students, social workers, health professionals, and local government authorities, as well as the twins and their families, who enthusiastically collaborated with the registry all over the country.

References

- Cohen, D. J., Dibble, E., Grawe, J. M., & Pollin, W. (1975). Reliably separating identical from fraternal twins. *Archives of General Psychiatry*, 32, 1371–1375.
- Goldsmith, H. H. (1991). A zygosity questionnaire for young twins: A research note. *Behavior Genetics*, 2, 257–269.
- Nichols, R. C., & Bilbro, W. C. Jr. (1966). The diagnosis of twin zygosity. *Acta Genetica et Statistica Medica*, 16, 265–275.
- Oficina Nacional de Estadísticas de Salud, Ministerio de Salud Pública. (2012). *Anuario Estadístico de Salud*. La Habana, Cuba: Ministerio de Salud Pública.
- van Dongen, J., Slagboom, E. P., Draisma, H. H. M., Martin, N. G., & Boomsma, D. I. (2012). The continuing value of twin studies in the omics era. *Nature Reviews Genetics*, 13, 640–653.