

The USC Adult Twin Cohorts: International Twin Study and California Twin Program

Wendy Cozen,^{1,2,3} Amie E. Hwang,¹ Myles G. Cockburn,^{1,2} Ann S. Hamilton,^{1,2} John Zadnick,¹ and Thomas M. Mack^{1,2,3}

¹Department of Preventive Medicine, USC Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

²Norris Comprehensive Cancer Center, USC Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

³Department of Pathology, USC Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

The study of twin subjects permits the documentation of crude heritability and may promote the identification of specific causal alleles. We believe that at the current time, the chief research advantage of twins as subjects, especially monozygotic twins, is that the commonality of their genetic and cultural identity simplifies the interpretation of biological associations. In order to study genetic and environmental determinants of cancer and chronic diseases, we developed two twin registries, maintained at the University of Southern California: The International Twin Study (ITS) and the California Twin Program (CTP). The ITS is a volunteer registry of twins with cancer and chronic disease consisting of 17,245 twin pairs affected by cancer and chronic disease, respectively, ascertained by advertising in periodicals from 1980–1991. The CTP is a population-based registry of California-born twin pairs ascertained by linking the California birth records to the State Department of Motor Vehicles. Over 51,000 individual California twins representing 36,965 pairs completed and returned 16-page questionnaires. Cancer diagnoses in the California twins are updated by regular linkage to the California Cancer Registry. Over 5,000 cancer patients are represented in the CTP. Twins from both registries have participated extensively in studies of breast cancer, melanoma, lymphoma, multiple sclerosis, systemic lupus erythematosus, diabetes mellitus type 1, mammographic density, smoking, and other traits and conditions.

■ **Keywords:** international twin study, California Twin Program, cancer, chronic disease, etiology, twins

Overview and Mission

Two twin registries were developed at the University of Southern California during the 1980s and 1990s (Cockburn et al., 2001c, 2002, 2006; Mack et al., 2000). One, the International Twin Study (ITS), is a volunteer registry of 17,245 twin pairs affected by cancer or other chronic diseases. The other, the California Twin Program (CTP), is a population-based twin registry of native Californians with information about roughly 75,000 native resident twins. The mission of the registries is to identify twins in order to conduct studies on the etiology of chronic diseases. There is currently no infrastructure support for either registry, and support is maintained via individual hypothesis-based awards. Each registry is further described below.

International Twin Study (ITS)

The ITS is a volunteer-based registry of twin pairs in which at least one member is affected by cancer (12,296 affected pairs) or a chronic disease (4,949 affected pairs), and is created to identify potential subjects for studies of environmental and genetic chronic disease determinants (Table 1) (Mack et al., 2000).

RECEIVED 22 October 2012; ACCEPTED 25 October 2012. First published online 7 December 2012.

ADDRESS FOR CORRESPONDENCE: Wendy Cozen, USC Keck School of Medicine and Norris Comprehensive Cancer Center, University of Southern California, MC-9175, Los Angeles, CA 90089-9175, USA. E-mail: wcozen@usc.edu

TABLE 1
Zygoty and Sex Distribution of Twins Ascertained by the International Twin Study

	Cancer	Multiple sclerosis	Diabetes	CVD ¹	Other chronic diseases	Total
MZ (male)	1,733	104	184	89	193	2,303
MZ (female)	3,810	358	375	165	666	5,374
DZ (male–male)	1,090	113	125	62	106	1,496
DZ (female–female)	2,654	313	232	95	450	3,744
DZ (male–female)	2,487	374	242	134	335	3,572
Unclear sex or zygoty	522	31	84	41	78	756
Total	12,296	1,293	1,242	586	1,828	17,245

Note: ¹Cardiovascular disease.

From 1980 to 1991, advertisements were placed in over 300 newspapers and magazines across the United States and Canada, with a circulation of roughly 68 million, seeking ‘twins with cancer’ or ‘twins with chronic conditions.’ There was no age criterion; parents of affected twin children responded on behalf of their children. Direct responses from members of affected twins living within the periodicals’ circulation area made up 79% of responses, from those living outside that area 4%; another 17% of the responding twin pairs were referred by friends or family. Once initial contacts were established, a signed permission to obtain medical records, and for cancer cases, pathology reports and diagnostic slides, was requested. Pathology reports and medical records were reviewed for the majority (71.4%) of twin patients, and histological slides were reviewed and classified for 65% of patients, with higher proportions for specific sites. Follow-up with biannual contact was continued until 1993 for the entire registry and longer for study-specific conditions (multiple sclerosis, breast cancer, hematologic neoplasms, melanoma, systemic lupus erythematosus). Respondents were judged by demographic characteristics to be representative of all North American non-Hispanic white twins with cancer or another chronic disease (Mack et al., 2000). Disease concordance patterns by zygoty could be examined and interpreted with possible selection bias in mind. Heritability has been evaluated for type 1 diabetes mellitus (Kumar et al., 1988, 1993), Hodgkin and non-Hodgkin lymphoma (Mack et al., 1995), childhood cancer (Buckley et al., 1996), systemic lupus erythematosus (Deapen et al., 1992), breast cancer (Mack et al., 2002), and multiple sclerosis (Islam et al. 2006) using this cohort. Approximately 6,000 questionnaires have been collected from pairs with systemic lupus erythematosus, diabetes mellitus, multiple sclerosis, melanoma, and breast, lung, bladder, gastrointestinal, and hematologic neoplasms. Biological samples, including blood, saliva, and tumor specimens, have been collected in the context of disease-specific studies, including those of systemic lupus erythematosus ($n = 44$ pairs), Hodgkin lymphoma ($n = 186$ pairs), breast cancer ($n = 280$ pairs), multiple sclerosis ($n = 914$ pairs), bladder cancer ($n = 126$ pairs), and pancreatic cancer ($n = 45$ surviving monozygotic (MZ) twins). Researchers interested in accessing these twins for collaborations should con-

TABLE 2
Response Rates Among Located Individual Twins by Age and Sex, California Twin Program

Birth year	Age in 2000	Male	Response rates		
			Male	Female	
1908–1917	83–92	282	56.0%	48	51.6%
1918–1927	73–82	1,027	60.4%	267	58.6%
1928–1937	63–72	1,688	61.9%	448	57.4%
1938–1947	53–62	3,508	53.6%	2,856	63.6%
1948–1957	43–52	6,601	44.4%	8,545	59.3%
1958–1967	33–42	5,062	28.8%	7,635	44.4%
1968–1977	23–32	2,907	19.8%	5,182	34.3%
1978–1982	18–22	1,913	15.5%	3,640	28.8%
Overall		22,988	32.4%	28,621	43.9%

Note: Reproduced from Cockburn et al. (2002) with editor’s permission.

tact Drs Thomas Mack (tmack@usc.edu) or Ann Hamilton (ahamilt@usc.edu).

California Twin Program (CTP)

The CTP is a population-based registry of twins born in California between 1908 and 1982 (Cockburn et al., 2001c, 2002, 2006). Twins were identified from California birth records, and contact addresses were obtained by linkage to the records of California Department of Motor Vehicles in 1989, 1998, 1999, 2000, and 2001. During this period there were 256,616 multiple births registered, and 161,109 California residents were identified from linkage to the Department of Motor Vehicles (DMV). Recruitment efforts were carried out in four waves between 1991 and 2000, during which letters and invitation were sent to twins with valid addresses. A 16-page screening questionnaire was sent to members of twin pairs with questions on demographic characteristics, zygoty, growth and development, reproductive history, lifestyle factors, dietary preferences, and medical history. Of 115,733 questionnaires mailed, 51,609 (representing 36,965 pairs) were returned, yielding a response rate of 44.6%, similar to or better than those reported among similarly aged persons in other cohort studies. The following tables (Tables 2 and 3) previously published in this journal (Cockburn et al., 2002) describe the cohorts.

TABLE 3
Zygosity and Sex Distribution of Respondent Pairs in the California Twin Program

Zygosity	Total pairs	%	Pairs with two respondents
MZ (male)	5,381	15	2,196
MZ (female)	6,314	17	3,263
DZ (male–male)	6,270	17	2,112
DZ (female–female)	6,785	18	2,992
DZ (male–female)	11,240	30	3,708
Unclear like-sex	975	3	373
Total	36,965		14,644

Note: Reproduced from Cockburn et al. (2002) with editor's permission.

The respondents represented individuals from 25,725 like-sex and 11,240 unlike-sex twin pairs. The zygosity and sex distribution is shown above (Table 3).

The twins have been shown to be demographically similar to the 1990 California population from which they were drawn in respect to geography, sex, age, race, ethnicity, and social class, except for a slightly lower level of education (Cockburn et al., 2001c). Cancer diagnoses ($n = 3,848$) were initially self-reported, including 508 cases of melanoma, 371 breast cancer, 365 thyroid cancer, 275 endometrial cancer, 236 ovarian cancer, 219 lung cancer, 204 colorectal cancer, 198 prostate cancer, 192 stomach cancer, 160 Hodgkin lymphoma, 141 central nervous system cancer, 124 bladder cancer, 113 testis cancer, and 36 cases of pancreatic cancer. In 2010, we first linked the CTP with the California Cancer Registry, which began state-mandated cancer reporting in 1988, and ascertained about 5,000 cases of cancer in the CTP (with presumably a low level of overlap with the self-reported cases). New cases among these aging twins will accrue prospectively at a number in excess of 500 annually. The web page for the California Twin Program can be found at <http://www.usc.edu>. Interested researchers may contact the director of the program, Dr Thomas Mack, at tmack@usc.edu, or one of the three co-directors: Dr Myles Cockburn at myles@usc.edu; Dr Ann Hamilton at ahamilt@usc.edu; or Dr Wendy Cozen at wcozen@usc.edu.

Research Highlights and Ongoing Projects

There is an extensive bibliography of studies using twins from both registries, with over 50 published papers. Below are highlights of special areas of interest and ongoing projects.

Breast Cancer

We have shown that risk to the unaffected co-twin of a case is 1–2% per year throughout life in the setting of high genetic susceptibility (Peto & Mack, 2000). We subsequently showed that among breast cancer concordant pairs, the twin who had an earlier menarche was 5.4 times as likely to be diagnosed with breast cancer first (Hamilton & Mack, 2003). A study of mammographic density among healthy MZ and dizygotic (DZ) twin pairs showed that both genetic

factors and modifiable environmental factors (e.g., obesity) were important determinants (Ursin et al., 2009). We are currently developing proposals to confirm these findings.

Hodgkin Lymphoma

We were the first to show an extremely high risk of young adult Hodgkin lymphoma in co-twins of MZ cases compared with co-twins of DZ cases (Mack et al., 1995). By using the unaffected MZ co-twin of a case as a surrogate, we defined a susceptibility immunophenotype characterized by high inflammatory and low T-helper-1 (Th1) cytokine levels (Cozen et al., 2004b, 2008). We then showed a strong inverse association between early childhood behaviors linked to infectious exposures and risk of young adult Hodgkin lymphoma (Cozen et al., 2009). A genome-wide association scan (GWAS) of young adult Hodgkin lymphoma revealed significant associations with loci in the HLA class II region of the genome (Cozen et al., 2012). A polymorphism in PR domain containing protein 1 (PRDM1) was found to predict second primary breast cancer among Hodgkin lymphoma patients treated with radiation therapy (Best et al., 2011). Studies comparing Epstein–Barr virus load and gut microbiota in Hodgkin lymphoma-discordant pairs are ongoing.

Melanoma

We demonstrated that recall bias could at least partially explain associations between sunbathing and melanoma, especially when the respondent believed that sun exposure was a cause of melanoma (Cockburn et al., 2001a, 2001b). However, associations between ease of sun burning and tanning did not appear to be subject to bias. We then found evidence that mole prevalence in general was largely constitutional (Cockburn et al., 2001a, 2001b), and the risk of having three or more large nevi in particular was linked to blonde hair, Celtic ancestry, and a propensity to burn rather than tan (all genetically determined characteristics), rather than sun exposure behavior (Cockburn et al., 2007). Studies on the determinants of nevi and melanoma in twins are ongoing.

Multiple Sclerosis

We first demonstrated that twin concordance for multiple sclerosis was higher among MZ compared with DZ twins as expected, but the MZ:DZ concordance ratio was also increased by North European ancestry, younger age at diagnosis, and Northern latitude of birth (Islam et al., 2006), indicating gene–environment interaction. We then showed that sun exposure modified the risk of multiple sclerosis within disease-discordant MZ pairs (Islam et al., 2007). The twins were included in a large GWAS of multiple sclerosis that identified numerous risk loci in immune response genes (International Multiple Sclerosis Genetics Consortium et al., 2011). Studies evaluating gene–environment interaction are ongoing.

Smoking

In a study of smoking initiation and persistence, we found that the strongest influence on initiation was having a twin who smoked, with evidence of modification by closeness (Hamilton et al., 2006). Gender strongly influenced this relationship such that co-twin influence was stronger among female compared with male pairs, presumably on the basis of differential behavior. The only significant determinant of smoking persistence was having a co-twin who continued to smoke. In a second study, we were the first to demonstrate that smoking was directly associated with increased interleukin-5 and interleukin-13 levels (Cozen et al., 2004a).

DNA Methylation

In order to investigate epigenetic mediation of environmental determination of chronic disease in identical twins (Cortessis et al., 2012), we are conducting a study of environmental determinants of DNA methylation in twins and have collected paired blood and saliva samples from each member in over 600 MZ pairs to date. Using an Illumina Infinium platform, we have searched unsuccessfully for differences in the DNA methylation status at 19,000 cytosine loci in whole blood DNA between 25 pairs of MZ twins discordant for multiple sclerosis and between 28 pairs of twins discordant for smoking (unpublished). Preliminary results support differences in DNA methylation patterns by hematopoietic cell subsets separated by flow cytometry prior to DNA extraction.

Other Research

We recently showed a twofold increased risk of infectious mononucleosis in co-twins of MZ compared with DZ cases (Hwang et al., 2012). Twins have participated in studies examining genetic variation in the genes encoding proteins related to catecholamine synthesis and regulation, including tyrosine hydroxylase (Zhang et al., 2004), chromogranin A (Chen et al., 2008), and neuropeptide Y(1) receptor (Wang et al., 2009). Other ongoing projects include a study of the determinants of height differences in MZ pairs, agreement between mothers and twins on childhood experiences, and microbiome studies in twins discordant for traits and disease.

Future Plans

Funding for the resources will be provided by hypothesis-driven projects investigating cancer and chronic disease etiology. We also plan to continue to try to secure infrastructure support through a variety of mechanisms.

Acknowledgments

We acknowledge and thank the twins and their families who take the time to participate in our studies. This work was supported by grants from the National Institutes

of Health (R01CA110836, R03CA110836, R01CA58839, P30AG017265), the California Tobacco-Related Disease Research Program (7RT-0134H, 8RT-0107H, 6RT-0354H), and the Leukemia Lymphoma Society (TR6137-07).

References

- Best, T., Li, D., Skol, A. D., Kirchhoff, T., Jackson, S. A., Yasui, Y., . . . Onel, K. (2011). Variants at 6q21 implicate *PRDM1* in the etiology of therapy-induced second malignancies after Hodgkin lymphoma. *Nature Medicine*, *17*, 941–943.
- Buckley, J. D., Buckley, C. M., Breslow, N. E., Draper, G. J., Roberson, P. K., & Mack, T. M. (1996). Concordance for childhood cancer in twins. *Medical and Pediatric Oncology*, *26*, 223–229.
- Chen, Y., Rao, F., Rodriguez-Flores, J. L., Mahata, M., Fung, M. M., Stridsberg, M., . . . O'Connor, D. T. (2008). Naturally occurring human genetic variation in the 3'-untranslated region of the secretory protein chromogranin A is associated with autonomic blood pressure regulation and hypertension in a sex-dependent fashion. *Journal of the American College of Cardiology*, *52*, 1468–1481.
- Cockburn, M. G., Black, W., McKelvey, W., & Mack, T. M. (2001a). Determinants of melanoma in a case-control study in twins. *Cancer Causes and Control*, *12*, 615–625.
- Cockburn, M., Hamilton, A., Cozen, W., Zadnick, J., & Mack, T. (2002). Chronic diseases in a large population-based cohort. *Twin Research*, *5*, 460–467.
- Cockburn, M. G., Hamilton, A. S., & Mack, T. M. (2001b). Recall bias in self-reported melanoma risk factors. *American Journal of Epidemiology*, *153*, 1021–1026.
- Cockburn, M., Hamilton, A., & Mack, T. (2006). Twins as willing research participants: Successes from studies nested within the California Twin Program. *Twin Research and Human Genetics*, *6*, 927–932.
- Cockburn, M., Hamilton, A., & Mack, T. (2007). The simultaneous assessment of constitutional, behavioral, and environmental factors in the development of large nevi. *Cancer Epidemiology, Biomarkers and Prevention*, *16*, 200–206.
- Cockburn, M. G., Hamilton, A. S., Zadnick, J., Cozen, W., & Mack, T. M. (2001c). Development and representativeness of a large population-based cohort of native Californian twins. *Twin Research*, *4*, 242–250.
- Cortessis, V. K., Thomas, D. C., Levine, A. J., Breton, C. V., Mack, T. M., Siegmund, K. D., . . . Laird, P. W. (2012). Environmental epigenetics: Prospects for studying epigenetic mediation of exposure-response relationships. *Human Genetics*, *131*, 1565–158.
- Cozen, W., Diaz-Sanchez, D., Gauderman, J. W., Zadnick, J., Cockburn, M. G., Gill, P. S., . . . Mack, T. M. (2004a). Th1 and Th2 cytokines and IgE levels in identical twins with varying levels of cigarette consumption. *Journal of Clinical Immunology*, *24*, 617–622.
- Cozen, W., Gill, P. S., Ingles, S. A., Masood, R., Martinez-Maza, O., Cockburn, M. G., . . . Mack, T. M. (2004b). IL-6 phenotype and genotype are associated with a risk of young adult Hodgkin's disease. *Blood*, *103*, 3216–3221.

- Cozen, W., Gill, P. S., Salam, M. T. R., Nieters, A., Masood, R., Cockburn, M. G., . . . Mack, T. M. (2008). Interleukin-2, interleukin-12 and interferon-(gamma) levels and risk of young adult Hodgkin lymphoma. *Blood*, *111*, 3377–3382.
- Cozen, W., Hamilton, A. S., Zhou, P., Salam, M. T., Deapen, D. M., Nathwani, B. N., . . . Mack, T. M. (2009). A protective role for early oral exposures in the etiology of young adult Hodgkin lymphoma. *Blood*, *114*, 4014–4020.
- Cozen, W., Li, D., Best, T., Van Den Berg, D., Gourraud, P.-A., Cortessis, V. K., . . . Onel, K. (2012). A genome-wide meta-analysis of nodular sclerosing Hodgkin lymphoma identifies risk loci at 6p21.32. *Blood*, *119*, 469–475.
- Deapen, D. M., Horwitz, D. H., Escalante, A., Weinrib, L., Roy-Burman, P., Walker, A., & Mack, T. M. (1992). A revised estimate of twin concordance in SLE. *Arthritis and Rheumatism*, *35*, 311–318.
- Hamilton, A. S., Lessov-Schlaggar, C. N., Cockburn, M. G., Unger, J. B., Cozen, W., & Mack, T. M. (2006). Gender differences in determinants of smoking initiation and persistence in California twins. *Cancer Epidemiology Biomarkers and Prevention*, *15*, 1189–1197.
- Hamilton, A. S., & Mack, T. M. (2003). Puberty and genetic susceptibility to breast cancer. *New England Journal of Medicine*, *348*, 2313–2322.
- Hwang, A. E., Hamilton, A. S., Cockburn, M. G., Ambinder, R., Zadnick, J., Brown, E. E., . . . Cozen, W. (2012). Evidence of genetic susceptibility to infectious mononucleosis: A twin study. *Epidemiology and Infection*, *140*, 2089–2095.
- International Multiple Sclerosis Genetics Consortium, Wellcome Trust Case Control Consortium 2, Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C. C., . . . Compston, A. (2011). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*, *476*, 214–219.
- Islam, T., Gauderman, W. J., Cozen, W., Hamilton, A. S., Burnett, M. E., & Mack, T. M. (2006). Differential twin concordance for multiple sclerosis by latitude: Genetic or environmental determinants? *Annals of Neurology*, *60*, 56–64.
- Islam, T., Gauderman, W. J., Cozen, W., Hamilton, A., & Mack, T. M. (2007). Childhood sun exposure modifies risk of multiple sclerosis among monozygotic twins. *Neurology*, *4*, 381–388.
- Kumar, D., Gemayel, N. S., Deapen, D., Kapadia, D., Yamashita, P. H., Lee, M., . . . Mack, T. M. (1993). North-American twins with IDDM genetic, etiological, and clinical significance of disease concordance according to age, zygosity, and the interval after diagnosis in first twin. *Diabetes*, *42*, 1351–1363.
- Kumar, D., Gemayel, N. S., Gill, S. K., Bray, G. A., Roy-Burman, P., Deapen, D., & Mack, T. M. (1988). Type-specific concordance in young diabetic monozygotic twins. *Advances in Experimental Medicine and Biology*, *245*, 259–267.
- Mack, T. M., Cozen, W., Shibata, D. K., Weiss, L. M., Nathwani, B. N., Hernandez, A. M., . . . Rappaport, E. B. (1995). Concordance in identical twins suggests genetic susceptibility to young adult Hodgkin's disease. *New England Journal of Medicine*, *332*, 413–418.
- Mack, T. M., Deapen, D., & Hamilton, A. S. (2000). Representativeness of a roster of North American twins with chronic disease. *Twin Research*, *3*, 33–40.
- Mack, T. M., Hamilton, A. S., Press, M., Diep, A., & Rappaport, E. B. (2002). Heritable breast cancer in twins. *British Journal of Cancer*, *87*, 294–300.
- Peto, J., & Mack, T. M. (2000). High constant incidence in twins and other relatives of women with breast cancer. *Nature Genetics*, *26*, 411–414.
- Ursin, G., Lillie, E. O., Lee, E., Cockburn, M. G., Schork, N. J., Cozen, W., . . . Mack, T. M. (2009). The relative importance of genetics and environment on mammographic density. *Cancer Epidemiology Biomarkers and Prevention*, *18*, 102–112.
- Wang, L., Rao, F., Zhang, K., Mahata, M., Rodriguez-Flores, J. L., Fung, M. M., . . . O'Connor, D. T. (2009). Neuropeptide Y(1) receptor NPY1R discovery of naturally occurring human genetic variants governing gene expression in cells as well as pleiotropic effects on autonomic activity and blood pressure in vivo. *Journal of the American College of Cardiology*, *54*, 944–954.
- Zhang, L., Rao, F., Wessel, J., Kennedy, B. P., Rana, B. K., Taupenot, L., . . . O'Connor, D. T. (2004). Functional allelic heterogeneity and pleiotropy of a repeat polymorphism in tyrosine hydroxylase: Prediction of catecholamines and response to stress in twins. *Physiological Genomics*, *19*, 277–291.