The Occurrence of Chronic Disease and Other Conditions in a Large Population-based Cohort of Native Californian Twins

Myles Cockburn, Ann Hamilton, John Zadnick, Wendy Cozen, and Thomas M. Mack

University of Southern California/Keck School of Medicine, California, USA

We describe the prevalence of chronic diseases and conditions in a large cohort of twins, which has been developed to facilitate studies of the role of genetics and environment in the development of disease. The California Twin Program (CTP) comprises twins born in California between 1908 and 1982. Birth records from all multiple births (256,616 in total) were linked (multiple times between 1990 and 2001) with the California Department of Motor Vehicles (DMV) roster of licensees to obtain address information. The linkages have revealed 161,109 matches and, because of less complete DMV records in some years, were less successful in older females than in all others. To date over 51,000 of these twins have completed a detailed 16-page mailed risk factor questionnaire. Based on estimates of numbers of individuals receiving a questionnaire, our crude response rates are as high as 63.6% (among females currently in their 50s), with an overall crude response rate of 37.9%. Similar to our previous report regarding the first 42,000 twins, the current group who have completed the questionnaire are representative of the population from which they were drawn (in terms of age, sex, race and residential distribution). Self-reported disease frequencies are provided, along with current estimates of future cancer incidence and mortality rates likely to be observed in the group. We outline our plans for cohort expansion, additional studies using the cohort, and future plans for inviting collaboration.

We have previously argued that many of the liabilities of a twin cohort established solely on the basis of chronic disease occurrence can be overcome only by developing a population-based twin cohort established in a manner unbiased by common exposure or outcome (Cockburn et al., 2001). Subsequently, we reported on the establishment of the California Twin Program, a cohort of twins based on the 256,616 twin individuals born in California between 1908 and 1982. Included were all twin subjects with and without pertinent environmental exposures, healthy and diseased, paired and surviving, monozygotic and dizygotic, and like-sex and unlike-sex. We described the formation and enrollment of approximately 41,000 members of the cohort, discussing the degree to which they represented California twins and Californians generally (Cockburn et al., 2001).

In this report we provide both an update on our recent activities, including the enrollment of a further 10,000 twins, and provide the first published data on the prevalence of self-reported chronic diseases (including cancer). While we have conducted linkage of the cohort with the California Cancer Registry, those results are preliminary, so we will elucidate in greater detail prospects for further ascertaining in an objective and verifiable manner incidence of and mortality from cancer in the cohort. We report also on the prevalence of many exposures and occurrences likely to be of interest to twin researchers from a variety of disciplines.

Finally, in order to confirm the potential of this resource as a mechanism for future twin research, we outline the means by which we have conducted, and plan to further, our own twin studies within the cohort, and how we are ensuring that the resource remains available for future collaborative research while it continues to grow.

Materials and Methods

Full details of the source of the "birth" cohort of twins, the recruitment of participants, and determination of the representativeness of resulting "respondent" cohort, are contained elsewhere (Cockburn et al., 2001), and we summarize them only briefly here, for the purpose of updating as needs be.

Cohort Establishment and Recruitment of Participants

Records of live multiple births in the state of California occurring between 1908 and 1982 were obtained from the California Department of Vital Statistics. This set was linked to the records of the California Department of Motor Vehicles (DMV) in 1989, 1998, 1999, 2000 and 2001 using first and last name (linked to first, last and "a.k.a." name of DMV record) and date of birth, and returned a current address and new married name, when available.

We carried out recruitment in 4 "waves", one in each of 1991, 1998, 1999 and 2000–1. Each wave was conducted in a similar manner. After comparing the DMV-linked file to the National Change of Address Index (NCOA) to remove or update addresses (due to typographical errors,

Address for correspondence: Myles Cockburn, 1441 Eastlake Ave MC9175, Los Angeles CA90089-9175, USA. Email: cockburn@usc.edu

https://doi.org/10.1375/twin.5.5.460 Published online by Cambridge University Press

or out-of-date street names or zip codes), we sent letters of invitation to the twins with valid addresses. The letter contained a reply-paid postcard for the twin to update their information, inform us of the location of their twin, and space to inquire about the study. At this stage, and at each subsequent mailing, we removed twins with incorrect addresses (returned by the postal services) from the subsequent questionnaire mailing. We then mailed a 16-page questionnaire with its reply-paid envelope to all those individuals whose addresses were thought to be valid. After a further 5 weeks we sent a reminder postcard again urging twins to take part in the study, and after a further 3 weeks (on average), we sent a second copy of the questionnaire to those twins not yet responding, provided they had not refused and their package had not been returned with an unknown address.

Estimation of the True Denominator of Twins Able to Receive a Questionnaire and Subsequent Response Rates

We made estimates of the true denominator for our response rates based on a number of assumptions, first eliminating those twins known to have bad addresses (from postal information), and second by reducing the denominator by the proportion whom we could locate but whom we were certain would not have received mailed materials (15% of all respondents with no known incorrect address (Cockburn et al., 2001)). We assessed the variation in response rates, by age, sex and geographical location (county) to determine potential sources of selection bias. Finally, we considered the breakdown of respondents by zygosity, derived from self-report, and compared the pairwise distribution to that of the original birth cohort to assess their representativeness by sex and zygosity.

The Questionnaire

The 16-page questionnaire (available for viewing at http://twins.usc.edu/questionnaire) asked about basic demographic characteristics (age, sex, education, occupation, marital status), perceived zygosity (Kasriel & Eaves, 1976), growth and development, reproductive history, use of medical services, dietary preference, disease experience (including cancer occurrence), and lifestyle choices (smoking, alcohol consumption, exercise, sun exposure).

Self-reported Chronic Disease Prevalence and Estimates of the Future Cancer Burden in the Cohort

We list the numbers of concordant and discordant pairs for self-reported chronic disease outcomes and we report these by zygosity. We took the age distribution of the cohort and applied the 5-year age-specific rates of cancer incidence and mortality obtained from the Surveillance, Epidemiology and End Results program of the United States (http://seer.cancer.gov) data for the past 5 years to calculate estimates of the expected number of cancer incident cases and deaths in this cohort in the next 10 years (2002–2011).

Results

There were 265,616 multiple births registered in the State of California between 1908 and 1982, and 161,109 successful results from linkage of this cohort with the records of the DMV. Female names were less successfully linked, especially those currently over age 53, with the disparity between the sexes increasing with increasing age (Table 1). Twins in their 30s were most often found in DMV linkages, with both younger and older twins less successfully linked. We found a further 5431 twins by other means, the majority by referral from respondent co-twins (this number is smaller than previously reported, because we have subsequently "found" many of those twins in DMV linkages). A small number of others heard about our study and contacted us via website: http://twins.usc.edu, or by telephone. Females were more likely than males to be ascertained in this manner (not shown).

Response and Response Rates

Of 142,434 individuals sent questionnaires (those with usable addresses after the introductory letter mailing), only 136,156 could possibly have received the questionnaire — the remainder were either later found to be deceased or were without a known address. Of these, 51,609 were returned for a crude overall response rate of 37.9%. Above the age of 62 years, males were slightly more likely than

Table 1

Demographic Comparison of Twins Identified in the California Twin Program and the Original Birth Cohort from Whom They Were Drawn, by Age and Sex

		California birth record of multiple births, 1908–1982 (*n* = 265,616)				Twins found by any method (*n* = 166,540)					
		Males		Females		Males			Females		
Birth year	Age (2000)	Number	%distribution	Number	%distribution	Number	%distribution	%found	Number	%distribution	%found
1908–17	83–92	2596	2.0%	2645	2%	677	0.8%	26.1%	116	0.1%	4.4%
1918–27	73–82	4853	3.7%	5121	3.8%	2037	2.3%	42.0%	558	0.7%	10.9%
1928–37	63–72	5956	4.5%	6164	4.6%	3296	3.7%	55.3%	949	1.2%	15.4%
1938–47	53–62	12,381	9.4%	12,801	9.6%	7920	9.0%	64.0%	5143	6.6%	40.2%
1948–57	43–52	24,958	18.9%	25,174	18.9%	18,130	20.5%	72.6%	16,540	21.1%	65.7%
1958–67	33–42	34,021	25.7%	33,575	25.2%	23,176	26.3%	68.1%	21,823	27.9%	65.0%
1968–77	23–32	26,454	20.0%	26,558	19.9%	18,466	20.9%	69.8%	18,465	23.6%	69.5%
1978–82	18–22	21,100	15.9%	21,259	15.9%	14,533	16.5%	68.9%	14,711	18.8%	69.2%
Total		132,319	100%	133,297	100%	88,235	100%		78,305	100%	

https://doi.org/10.1375/twin.5.5.460 Published online by Cambridge University Press

females to respond, but under the age of 63 years, females were far more likely to respond, with almost twice the proportion of females than males responding in the youngest twins, those aged 18 to 23 years, among whom the lowest response rates were seen overall (Table 2). Response rates were similar in Whites, Latinos and Asians, and substantially lower among the small number of African-American twins and those reporting "other" race. These differences did not vary markedly by age (not shown) or sex (Table 2).

Comparison of Respondent Demographic Data with Census Data

Due to the deficit of older females linked to the DMV, the female respondents were very much younger on average than the California-born 1990 resident population of females. Latino respondents were slightly over-represented, and accordingly whites and African-Americans were slightly under-represented. This difference was more striking among females (Table 3). Respondents were less likely to have completed more than 12 years of education than the comparative population of California-born 1990 residents, more so in males than in females (Table 3). However, as previously reported, respondents were substantially better educated than the 1% sample of all US residents (for less than 12 years education, 12 years education and greater than 12 years education respectively, males: 38.9%, 23.0%, 38.0%; females: 37.3%, 26.7%, 36.0%), and better educated than the California-born US residents (for less than 12 years education, 12 years education and greater than 12 years education respectively, males: 41.3%, 18.2%, 40.5%; females: 39.0%, 19.9%, 41.1%). Finally, all Californian counties each with more than 5% of the population residing in them had nearly identical distributions of respondents and California-born 1990 resident population (not shown).

Double- and Single-respondent Twin Pairs, Zygosity and Sex of Respondents Compared to Expectation

A substantially larger proportion of females than males belonged to double-respondent pairs in each age group. Female-female pairs were slightly more prevalent among respondents than in the birth cohort of California twins (Table 4). The distribution of pairs by gender is similar to that of the birth cohort (Table 4), but the proportion of MZ males and females in the respondent cohort was lower than their estimated proportion among live births, and subsequently the proportion of DZ twins, including like and unlike sex, was higher than their proportion estimated among live births (Table 5). In total we received responses from at least one member of 36,965 pairs of twins, with both members responding in 40% of pairs (Table 5).

Prevalence of Chronic Diseases

Table 6 provides the prevalence of pairs discordant and concordant for self-reported chronic diseases, including cancer, nervous system disorders, endocrine-related diseases, autoimmune conditions, infectious diseases, allergic conditions, gastrointestinal conditions, cardiovascular problems, and other interesting outcomes such as alcohol and drug dependence. We provide numbers of discordant pairs as the data most likely to be studied with various twin methods — the corresponding number of concordant pairs are in parentheses. While there are likely to be some biases in self-reported disease prevalence that require further validation, this table provides basic information on the numbers of pairs available for studying disease outcomes. The total number of respondent pairs indicated at the top

Table 2

Response Rates Among All Twins Identified in the California Twin Program (*n* = 51,609 Respondents) by Age, Sex and Race.

		Male	Response rates		Female	Response rates	
			1	2		1	2
Birth year	Age (2000)	Responses	%	%	Responses	%	%
1908–17	83–92	282	49.3%	56.0%	48	50.0%	51.6%
1918–27	73–82	1027	57.2%	60.4%	267	56.0%	58.6%
1928–37	63–72	1688	59.5%	61.9%	448	55.4%	57.4%
1938–47	53–62	3508	52.1%	53.6%	2856	61.9%	63.6%
1948–57	43–52	6601	42.8%	44.4%	8545	57.6%	59.3%
1958–67	33–42	5062	27.2%	28.8%	7635	42.4%	44.4%
1968–77	23–32	2907	18.6%	19.8%	5182	32.4%	34.3%
1978–82	18–22	1913	14.9%	15.5%	3640	27.7%	28.8%
Race							
White/Latino		21,345	34.6%	39.3%	25,695	46.5%	50.9%
Black		670	9.9%	12.5%	1308	18.9%	22.4%
Asian		425	32.8%	36.5%	526	46.3%	50.4%
Other		548	11.9%	17.2%	1092	23.7%	36.9%
Overall response and response rates		22,988	30.9%	32.4%	28,621	42.1%	43.9%

Notes:1 – Response rate as % of twins to whom we sent a questionnaire

2 – Response rate as % of twins likely to have received the questionnaire (removed those known to be deceased and Post Office returns)

3 – Response rate as % of twins we believe actually received the questionnaire (remove 15% of non–respondents) on the basis of our sub-study: 36.3% in males and 49.5% in females (no age- or race-specific data available)

Table 3

Demographic Comparison of Census-derived Figures for California-born 1990 Resident Population, the California Twin Birth Cohort, and Respondents to our Study, (Percentages Are the Percentage in Each Age/Sex Group) by Age, Race, Education and Occupational Groups

		Male		Female	
Birth year	Age (1990)	Census[1]	response	Census[1]	response
1908–17	73–82	4.6%	2.2%	6.0%	0.4%
1918–27	63–72	10.9%	7.9%	11.9%	2.2%
1928–37	53–62	13.9%	12.9%	14.3%	3.7%
1938–47	43–52	24.4%	26.8%	23.4%	23.6%
1948–57	33–42	46.1%	50.2%	44.5%	70.2%
Race					
White		85.7%	81.0%	85.2%	77.2%
Latino		5.7%	8.6%	5.8%	10.8%
Black		3.7%	2.6%	4.2%	4.0%
American Indian		1.3%	1.4%	1.5%	1.2%
Japanese/Chinese		2.9%	1.5%	2.7%	1.5%
Filipino/Thai		0.5%	0.4%	0.5%	0.5%
Other		0.1%	1.0%	0.1%	1.1%
Missing		0.0%	3.5%	0.0%	3.7%
Education					
Under 12 yrs		13.3%	22.1%	14.3%	17.7%
12 yrs		22.9%	30.8%	28.1%	33.2%
over 12 yrs		63.8%	47.1%	57.6%	49.2%
Occupation					
Managerial, professional and specialty		28.3%	32.9%	25.3%	32.9%
Technical, sales and admin. Support		19.0%	12.9%	33.4%	24.5%
Service		6.8%	4.9%	9.3%	6.4%
Farming, forestry and fishing		3.4%	1.2%	0.8%	0.3%
Precision product, craft and repair		17.3%	10.4%	1.8%	0.3%
Operators, fabricators and laborers		14.0%	13.8%	4.1%	2.5%
Other and unspecified		11.2%	23.9%	25.4%	33.2%

Note: [1] See (Cockburn, Hamilton et al. 2001) for details of census comparisons

Table 4

Paired Gender Comparisons Between Birth Record of California Twins and Respondents, Including Distribution of Zygosity Among Respondents.

	California birth record of multiple births	Estimated proportion of twins live born in Pacific US[1]		Respondents							
Gender of pair	Percent	MZ Percent	DZ Percent	MZ No.	MZ Percent	DZ No.	DZ Percent	unknown No.	unknown Percent	Total	Percent
Male	34.5%	18.0%	15.1%	7544	14.6%	8323	16.1%	508	1.0%	16,374	31.7%
Female	34.9%	18.5%	15.8%	9527	18.5%	9730	18.9%	835	1.6%	20,092	38.9%
Mixed	30.6%	—	32.6%	—	—	14,866	28.8%	—	—	14,866	28.8%
Unknown	—	—	—	—	—	—	—	277	0.5%	277	0.5%
Total	100.0%	36.5%	63.5%	17,071	33.6%	32,919	63.8%	1,619	3.1%	51,609	100.0%

Note: [1] Taken from Mack et al., 2000.

of the table refers to all twin pairs represented by any response – that is, for the "double respondent pairs" discordance is based on the report from both twins, whereas for the "single respondent twins" discordance is based on the response from only one twin, who reported that either they or their co-twin, but not both, had the condition.

Over the next 10 years as the cohort ages we have estimated that over 4000 cancers will occur among twins in the respondent pairs and nearly 7000 cancers will occur among all members of the cohort with addresses obtained (Table 7).

Discussion

This respondent cohort is currently the largest available population-based cohort of any kind in California and is

Table 5

Twin Pairs Represented by Current Respondents

Zygosity	Double respondents[1]	Single respondents[1]	Total pairs
MZ (male)	2196	3185	5381
MZ (female)	3263	3051	6314
DZ (male-male)	2112	4158	6270
DZ (female-female)	2992	3793	6785
DZ (male-female)	3708	7532	11,240
Unknown	373	603	976
Total	14,644	22,322	36,965

Note: [1] "double" respondents are pairs where both twins have sent back a questionnaire, "single" respondents are pairs where only one twin has sent back a questionnaire.

representative of native-born resident Californians. We have amassed a comprehensive set of exposure histories, and self-reports of disease that can later be verified by planned linkage to cancer registries and death indices. We previously demonstrated that our methods selected MZ and DZ twins in an unbiased (i.e., representative) fashion and with equal probability, and this also applies to the current 51,609 respondents. The ways in which this respondent cohort varies from the population with respect to age, sex, race, education, occupation, are now known and can be used to adjust the results of studies choosing to use this population simply as a cohort of native Californians. Likewise we can accurately estimate the role of selection bias in studies using these twins as subjects, and determine the extent to which cohort members followed for disease outcomes in future are likely to have been differentially ascertained with respect to zygosity.

These twins, who as paired individuals, share a common genome (MZ twins) or share on average half their genes (DZ twins), also share a majority of childhood exposures potentially pertinent to the etiology of disease. We can identify subsets of twins for further study requiring recontact(such as the collection of DNA samples), identified on the basis of either discordance for exposure(again derived from questionnaire data already at hand) or discordance of disease. One example of great interest is the prevalence and concordance of asthma. There were 6681 reports of asthma (either by a respondent, or by a non-respondent's responding co-twin), with discordance among 1151 MZ pairs (527 double- and 624 single-respondent pairs), 487 concordant MZ pairs (213 double- and 274 single-respondent); and 3272 discordant DZ pairs (1294 double- and 1978 single-respondent pairs), 494 concordant DZ pairs (182 double- and 312 single-respondent pairs). These cases represent the largest sample size anywhere in the world for investigating potential gene-gene and gene-environment interactions in the etiology of asthma. While much of the information regarding current cancer occurrence will both be related to survivorship and limited to retrospective exposure assessment, we expect 5873 deaths (1695 cancer deaths) to occur in the next 10 years in the current respondent cohort who have already provided us with details of exposures.

In addition to being able to contrast risk factors between over 36,000 MZ and DZ twin pairs, we will be able to determine the role of under-ascertainment of respondents by disease status, since we can compare the original birth record to California Cancer Registry and mortality records. We can then comment on the effect of ascertainment on disease concordance estimates and subsequent MZ/DZ comparisons as we have done elsewhere (Mack et al., 2000). We are linking the entire cohort of 256,616 individuals with the California Cancer Registry, a population-based roster of incident cancers occurring from 1988 to the present, and with California mortality data. We will contrast self-report with the result of the linkage, to ascertain survival bias and the accuracy of self- and proxy report of cancer.

We are currently preparing papers using the respondent cohort to investigate population-based risk factors for smoking uptake and cessation (Hamilton et al., 2001), population-based estimates of physical activity levels and characteristics, and risk factors for nevi size and frequency in California. We are conducting classic twin analyses of the risks for mammographic density among female twins, a case-control study of the role of cigarette smoking in the development of cytokines in identical twins discordant for smoking (Cozen et al., 2001), and a study of cytokine levels in twins discordant for Hodgkins disease (Cozen et al., 2001).

Our future aims are to collaborate with investigators from a variety of disciplines to ensure that this resource is utilized to the fullest extent possible. One of the advantages of the cohort design — that it can be used to examine the relationships between multiple exposures and many outcomes — can also provide a dilemma for investigators. For example, while our focus is primarily on cancers and other chronic diseases, twin populations are commonly used to study behavior, educational attainment, or psychosocial outcomes. Twins are also often the subjects of intervention studies aiming to eliminate the possibility that genetic differences account for the differences in intervention versus control outcomes. Therefore, we are compelled to ensure that we on the one hand collaborate with an international community of experts from many fields, and on the other hand, ensure that the way we go about inviting interaction that both meets the highest scientific requirements (e.g., human subjects' approval and analyses worthy of the quality of the data) and also is consistent with the longevity of the dataset. To that end, we have designed a set of protocols for

https://doi.org/10.1375/twin.5.5.460 Published online by Cambridge University Press

Table 6

Numbers of Pairs Discordant (Concordant in Parentheses) for Selected Self-reported Conditions, Given by Response Status (Single or Double Respondent) of the Twin Pair, and by Zygosity.

	Double-respondent pairs				Single-respondent pairs			
Selected conditions	MZ		DZ		MZ		DZ	
Total pairs								
1st questionnaire	2627		4817		2197		6538	
2nd questionnaire	2277		3167		3974		8296	
Both questionnaires	4904		7984		6171		14834	
Cancer								
Colon/Rectum	18	(0)	36	(0)	37	(3)	91	(4)
Lung	18	(0)	28	(1)	38	(5)	110	(4)
Stomach	20	(0)	30	(0)	40	(1)	83	(6)
Breast	46	(4)	107	(2)	33	(3)	149	(5)
Ovary	19	(0)	49	(1)	29	(2)	121	(3)
Uterus	37	(5)	71	(0)	40	(6)	93	(3)
Cervix	89	(11)	145	(2)	105	(13)	252	(8)
Testis	11	(0)	23	(0)	23	(1)	48	(2)
Hodgkins***	12	(0)	52	(0)	27	(0)	65	(2)
Prostate	22	(3)	51	(0)	30	(4)	74	(1)
Leukemia	8	(1)	13	(0)	22	(0)	43	(2)
Brain	8	(0)	17	(0)	35	(0)	69	(4)
Melanoma	91	(2)	164	(1)	70	(3)	155	(3)
Other skin cancer	323	(71)	651	(95)	195	(41)	606	(63)
Thyroid*	40	(5)	62	(5)	53	(13)	133	(13)
Bladder*	14	(0)	24	(0)	21	(5)	35	(4)
Esophagus*	1	(0)	8	(0)	6	(0)	12	(2)
Nervous System								
Glaucoma	45	(14)	125	(6)	53	(9)	152	(9)
Migraine	754	(255)	1497	(231)	684	(239)	1926	(318)
Epilepsy	74	(17)	154	(5)	98	(23)	291	(11)
Multiple Sclerosis	28	(2)	48	(2)	27	(2)	73	(6)
Optic Neuritis	18	(0)	29	(0)	14	(0)	33	(2)
Parkinsons	9	(1)	12	(0)	9	(1)	20	(1)
Myopia*	331	(204)	639	(180)	147	(308)	623	(420)
Macular Degeneration*	5	(0)	13	(0)	7	(1)	32	(2)
Schizophrenia*	5	(4)	18	(1)	24	(4)	84	(4)
Bipolar Disorder*	66	(5)	94	(4)	100	(27)	236	(18)
Endocrine								
Graves Disease	51	(2)	90	(3)	42	(3)	101	(9)
Type 1 Diabetes	53	(10)	101	(7)	72	(13)	235	(5)
Other Diabetes	121	(18)	216	(19)	142	(26)	364	(27)
Hypothyroid*	65	(14)	115	(8)	64	(20)	183	(20)
Hashimoto's Thyroid*	10	(5)	18	(1)	16	(4)	41	(5)
Autoimmune Condition								
Gout	121	(17)	285	(13)	95	(23)	334	(20)
Rheumatoid Arthritis	220	(25)	465	(26)	157	(43)	487	(60)
Osteoarthritis*	25	(3)	59	(0)	25	(8)	70	(6)
Lupus	23	(1)	53	(1)	28	(4)	79	(3)
Scleroderma	9	(0)	12	(0)	10	(0)	17	(2)
Dermatomyositis	10	(0)	43	(0)	15	(2)	37	(6)
Scoliosis*	117	(33)	208	(15)	111	(46)	330	(49)
Infections								
Emphysema	295	(32)	553	(36)	224	(71)	690	(91)
TB	41	(3)	79	(7)	58	(5)	158	(16)
AIDS	16	(2)	38	(1)	43	(1)	132	(3)
Infectious Mononucleosis*	202	(28)	290	(21)	226	(48)	421	(48)
Shingles*	112	(5)	166	(9)	122	(8)	327	(12)
Allergic Conditions								
Asthma	527	(213)	1294	(182)	624	(274)	1978	(312)
Allergic Shock	98	(4)	165	(7)	96	(10)	238	(14)
Animal/Plant	901	(302)	1749	(373)	589	(403)	2075	(545)
Drug allergy	1007	(328)	1952	(367)	692	(233)?	2174	(299)
Hayfever	1325	(919)	2889	(1133)	989	(795)	3619	(1321)

https://doi.org/10.1375/twin.5.5.460 Published online by Cambridge University Press

Table 6 continued

Selected conditions	Double-respondent pairs MZ		DZ		Single-respondent pairs MZ		DZ	
Total pairs								
Gastrointestinal Conditions								
Primary Billiary Cirrhosis	7	(0)	23	(1)	18	(3)	62	(2)
Crohn's Disease	17	(1)	41	(3)	17	(1)	64	(2)
Ulcerative Colitis	96	(8)	147	(6)	94	(13)	228	(9)
Peptic Ulcer	356	(26)	637	(32)	302	(37)	742	(55)
Hepatitis	314	(51)	674	(45)	306	(76)	778	(160)
Cardiovascular								
Stroke	43	(1)	101	(3)	54	(6)	157	(6)
Heart Attack	68	(6)	168	(5)	118	(15)	331	(17)
High Blood Pressure	467	(230)	1243	(241)	484	(182)	1674	(288)
Congenital Heart Problem	76	(7)	148	(1)	77	(12)	221	(23)
Rheumatic Heart Disease	47	(6)	122	(2)	48	(13)	154	(9)
Anemia	327	(69)	746	(64)	127	(50)	594	(65)
Heart Arrhythmia*	69	(6)	120	(7)	90	(16)	229	(20)
Mitral Valve Prolapse*	81	(22)	142	(2)	83	(23)	181	(10)
High Cholesterol*	154	(38)	268	(31)	159	(55)	403	(47)
Other Problems								
Alcoholism*	110	(33)	237	(19)	181	(107)	679	(151)
Dyslexia*	73	(21)	154	(16)	98	(51)	316	(50)
Drug Dependency*	133	(31)	247	(40)	220	(117)	768	(180)
ADD*	43	(9)	118	(10)	70	(41)	293	(37)

Note: *only asked in new questionnaire
**only asked in old questionnaire
***includes NHL in old questionnaire

Table 7

Expected Number of Deaths, Cancer Deaths, and Incident Cancers to be Identified During 10 Years of Follow-up Based on Age Specific SEER Incidence Rates and US Mortality Rates

	Individuals in Respondent Pairs*			Individuals in Pairs with Addresses**		
	Males	Females	Total	Males	Females	Total
All Deaths	4210	1663	5873	7625	2491	10116
Cancer Deaths	1124	571	1695	1903	858	2761
Incident Cancers	2451	1603	4054	4303	2580	6883

Note: *includes individuals who responded as well as any co-twins of these individuals who did not respond.
**includes individuals in respondent pairs

assessing potential collaboration, and have set up a working group to assess, in an organized fashion, various requests for data or collaboration.

Briefly, we include both investigators from the CTP, and external advisors, who assess each suggestion on the basis of scientific merit, available funding, and appropriate human subjects approval within which we include an assessment of the likely burden of both the proposed study, and its impact on the long-term use of the dataset. For example, a study requiring blood samples from a small number of individuals would only be acceptable if that group had not been approached for study in the past year, or there were no pending similar requests in the year after the proposed completion date of the study. Any investigators wishing to propose collaborations or obtain further details about the California Twin Program should contact the first author directly, at the address given.

Acknowledgments

Initial project management was conducted by Rich Pinder, and the California Twin Program web site was designed by Nick Fox. Jennifer Nedrud, Misha Birch, Saundra McGee and Travis Alexander conducted daily activities including receiving phone calls from twins and processing the enormous volume of mail required for this study. This study was funded by grants from the California Tobacco-related Disease Research Program (8RT-0107H and 6RT-0354H). We also thank the many twins who participated.

References

Cockburn, M., Hamilton, A., et al.(2001). Development and representativeness of a large population-based cohort of native Californian twins. *Twin Research, 4*(4), 1–9.

Cozen, W., Cockburn, M., et al. (2001). Effect of tobacco smoke on cytokines in twins. *Twin Research, 4*(3), 176.

Cozen, W., Masood, R., et al. (2001). Susceptible HD cytokine phenotypes in twins. *Twin Research, 4*(3), 177.

Hamilton, A., Cockburn, M., et al. (2001). Factors related to smoking behavior and exposure in California twins. *Twin Research, 4*(3), 185.

Kasriel, J., & Eaves, L. (1976). The zygosity of twins: further evidence on the agreement between diagnosis by blood groups and written questionnaires. *Journal of Biosocial Science, 8*(3), 263–266.

Mack, T. M., Deapen, D., et al. (2000). Representativeness of a roster of volunteer North American twins with chronic disease. *Twin Research, 3*(1), 33–42.