

---

# Ascertainment of a Mid-Western US Female Adolescent Twin Cohort for Alcohol Studies: Assessment of Sample Representativeness Using Birth Record Data

Andrew C. Heath, William Howells, Kathleen K. Bucholz, Anne L. Glowinski, Elliot C. Nelson and Pamela A. F. Madden

*Department of Psychiatry, Washington University School of Medicine, St Louis, USA*

Female twin pairs were identified from birth records, and their families invited to participate in a prospective study of the determinants of alcohol problems in women. We investigated sampling biases arising because of failure to locate families, or non-cooperation of families. Out of 2644 families with a live-born pair (born between July 1975 and December 1986) who survived beyond infancy, contact was established and a brief screening interview completed with 90% ( $N = 2380$ ). Fewer than 6% of located families declined to participate in the initial screening interview. Predictors of failure to locate a family or to obtain a screening interview were identified from information recorded in birth records, and from neighborhood characteristics identified from 1990 US Census block group data for the family residence when the twins were born. African-American families were under-represented in the final sample, but this effect was barely significant when other variables were controlled for. Under-represented were families where the mother was 19 or younger at the birth of the twins, where the mother herself was born out-of-state, or where information about biological father was not reported in the birth record. Non-participating families on average came from neighborhoods with a higher proportion of residents living in poverty, and with a higher proportion of African-American residents. Sampling biases were however small. The unusual cooperativeness in research of families with twins persists.

---

Results of twin and adoption studies, as well as evidence from studies in Asian populations of single gene effects on risk, provide compelling support for the hypothesis that genetic factors play a critical role in the etiology of alcoholism (reviewed in Heath et al., 1997; Jacob et al., 2001). Fundamental questions remain however about the mediators of genetic influences on alcoholism risk (how genetic influences arise) and about the moderators (protective or vulnerability factors) that interact with genetic predisposition. The twin study design is particularly powerful for addressing such effects (Jacob et al., 2001; Heath et al., in press), particularly when implemented prospectively, thereby avoiding the many potential biases associated with retrospective recall. Because of the early onset of alcohol dependence symptoms, with median onset age 20 or younger (e.g. Nelson et al., 1996), a prospective

study of the determinants of alcoholism risk is necessarily a pediatric study.

A variety of recruitment methods have been used in studies of juvenile twins, including ascertainment through school systems or through media appeals, and systematic ascertainment through birth or hospital records. Traditionally, twin researchers and other behavioral geneticists have been negligent in their inattention to the representativeness of the samples that are thus recruited. Identification of variables that predict reduced probability that a family will participate in a study enables sampling weights to be generated, to re-weight the observed data back to the population of families with twins (Heath et al., 1998; Heath et al., 1999a). When ascertainment is conducted from birth records, birth record data, together with neighborhood characteristics identified from public records such as census data (Meyer et al., 1996), can be used to check for major sampling biases. Here we use birth record and census data to examine potential sampling biases in a mid-western US twin cohort identified from birth records.

## Methods

Computerised state birth records from a mid-western US state, for the period 1 July 1975 to 30 June 1987, were scanned to identify all female like-sex twin pair births. Because of interruption of research funding, the cohorts of twins born 1 January to 30 June 1985, and 1 January to 30 June 1987, were incompletely ascertained, and therefore are excluded from totals and analyses presented here. In cases where an adoption occurs, biological parent information is sealed at the time of adoption under state law, so adoptee birth records were excluded. Infant deaths were recorded in the computerised birth record, so families with

---

*Address for correspondence: A. C. Heath, Missouri Alcoholism Research Center, Department of Psychiatry, Washington University School of Medicine, St Louis, Missouri, USA. Email: andrew@matlock.wustl.edu*

pairs where one or both twins were known to be deceased were also excluded.

Families were targeted for initial assessment when twins were aged 19, 17, 15, 13, or, from the third year of data-collection onwards, 11 years old. A cohort sequential sampling design was used (see Heath et al., 1999b), with recruitment of cohorts of 19, 17, 15 and 13-year-old twin pairs over a two year period, and continuing recruitment of 13 and 11-year-old twins in the two subsequent years. An initial screening interview was conducted with eligible families, to confirm the identity of the twin pair, and ask a few brief questions about twin pair zygosity. When parents gave permission, a follow-up diagnostic interview was scheduled with at least one parent (wherever possible the mother). The parent diagnostic interview (Hudziak et al., 1998; Neuman et al., 2001) covered prenatal, perinatal and early childhood risk-factors of the twins (e.g. maternal smoking and drinking during pregnancy, twins' birth weight and history of externalising disorders and depression) and critical aspects of parental psychopathology (parent's and co-parent's history of alcohol abuse or dependence, parent's history of depression, and of childhood conduct disorder) and sociodemographic factors (education, religious affiliation and involvement, family income, marital status, twins' rearing history). When adult twins gave permission, or when parents and minor twins (aged 17 or younger) gave permission, a follow-up diagnostic interview was also scheduled with the twins. The twin assessment (Bucholz et al., 2000; Nelson et al., 2000; Glowinski et al., 2001) included a detailed characterisation of experiences with alcohol, drinking patterns, alcohol abuse and dependence symptoms, and assessment of pertinent psychiatric risk-factors (particularly history of major depression, suicidality, childhood conduct disorder, and anxiety disorders) and measures of potential impairment versus achievement (e.g. educational level, school grades). The verbal consent procedures that were used were approved by the Washington University Human Studies Committee. Intake interviews with families were conducted over the period 1995–1999.

Birth record data included parental address at birth, allowing us to merge with US Census block group data (the smallest grouping of households for which Census data are available) to characterise the neighborhoods where twins were born. Data from the 1990 US Census were used. Variables examined were neighborhood percentages for (a) urban versus rural non-farm versus rural farm residences, (b) age-structure of the neighborhood (proportion of individuals under age 19, proportion of individuals aged 65 or older), (c) two parent households with children, (d) single parent (mother) with minor child(ren) households, (e) single resident households, (f) high density households (5 or more occupants), (g) proportion of households below the poverty line, (h) proportion of households with income about \$US 75,000, (i) median household income, (j) proportion of residents with at least some college, and (k) neighborhood ethnic composition (proportion of African-American residents, proportion of Hispanic residents). For these variables we created 3-category summary variables using data for all Missouri households, subdividing into the lowest quartile, middle

50%, and highest quartile, and compared families who did and did not complete the initial screening interview. A small number of families, for whom we had incomplete address information that did not permit assignment to a particular census-defined block group, were excluded from these analyses ( $N = 21$  families).

For some birth years, birth record data also included limited information about the biological mother, in particular maternal state of origin (i.e. the state where the mother was born), and maternal age at the birth of the twins (birth years 1978 and later). For a further subset of birth years the same information about the biological father was included, where provided by the biological mother (birth years 1980 and later). Parental age at birth was reclassified as a categorical variable ( $< = 19$ ; 20–25, 26–29, 30–34 and 35+). State of birth was coded as same state versus other state (including families that immigrated from foreign countries in this latter group).

Simple descriptive statistics were computed with a chi-square statistic used to test the significance of the overall associations between family characteristics (including neighborhood characteristics) at the time of the twins' birth, and whether or not we succeeded in completing an initial screening interview with the family. Classified as non-participants were both cases where we did not succeed in locating the family, and cases where we located the family but did not succeed in obtaining cooperation for an interview. The distinction between these two categories was not always clear, since without completing the screening interview we could not confirm that we had successfully contacted the correct family. All other families who completed the initial screening interview were classified as participants. Multiple logistic regression was used to identify the best joint predictors of non-participation of a family, with all predictors showing a significant univariate association with family non-participation included in the model. Since we were concerned about potential collinearity of some census predictors, many of which would be correlated with neighborhood poverty, we finally fitted a backwards selection stepwise logistic model, in which only variables that were significant at  $p < 0.05$  were retained in the model.

## Results

### Cooperative Families

A total of 2852 female like-sex twin pairs born between 1 July 1975 and 31 December 1986 (excluding the incompletely ascertained cohort born 1 January 1985–30 June 1985) were identified from birth records. After excluding 208 pairs identified in birth records as having at least once deceased twin (7.3% of all pairs), this left a total of 2644 families with a live-born twin pair that we attempted to trace and invite to participate in a research study. Numbers of available twin pairs, and level of family cooperation, as a function of twins' age when they were targeted for assessment, are summarised in Table 1. Overall, we were able to locate 95.6% of eligible families identified from birth records, with the success rate in locating families varying little as a function of the age of the twins at follow-up; and were able to complete an initial screening interview with

a parent (or occasionally other adult family member) from 90% of targeted families ( $N = 2380$ ). Only 5.6% of eligible families were located but did not wish to complete the initial screening interview. A parent diagnostic interview was successfully conducted with at least one parent from 2061 families, representing 78% of targeted families, and 87% of families who completed the initial screening stage. For those families where twins were aged 13 or older, and therefore targeted for interview assessment, interviews were successfully completed with at least one twin from 77% of targeted families, and 84% of families who completed the initial screening stage. Thus, while parental consent was required before twins aged 17 or younger could be interviewed, parents in some cases gave permission for their twins to be interviewed but did not themselves complete an interview, so that the success rate for obtaining a twin interview was approximately the same as the success rate for obtaining a parent interview. In families where at least one twin agreed to be interviewed, there was a 0.945 probability that the second twin also would complete an interview, leading to an overall total of 1633 female like-sex twin pairs where both twins completed a diagnostic interview (72.5% of targeted pairs, and 80% of pairs where the family completed the initial screening interview).

We were less successful in the recruitment of African-American families, and in obtaining their cooperation, particularly at the initial screening interview stage. There were a total of 479 African-American like-sex pairs born during this period, including 50 pairs where at least one twin died during infancy (10.4% of all pairs), leaving 429 pairs where both twins survived beyond infancy. We succeeded in tracing and completing an initial screening interview with 347 of these families (80.9%), and were unable to locate a further 8.2% of families. Parent interviews were completed with at least one parent from 313 families (73.0% of families identified from birth records, 90.2% of families who completed the initial screening interview). A twin interview was completed with at least one twin from 263 families, out of 362 families with twins aged 13 or older (72.7% of targeted families, 88.3% of families who completed the initial screening stage). We were able to complete interviews with both twins from 231 pairs (0.878 conditional probability that the second twin would complete an interview, given that one twin had been interviewed).

**Census-derived Predictors of Non-participation**

Table 2 summarises differences, in terms of neighborhood characteristics (based on 1990 US Census block group data), between participant and non-participant families. For some neighborhood characteristics there was no evidence of an ascertainment bias. This was true for the proportions of young people and older people living in the neighborhood; single-person households and high density households (with 5 or more residents), and urban versus rural non-farm versus farm households. Other neighborhood characteristics, however, did discriminate significantly between participating and non-participating households in univariate analyses. Non-participating families on average came from neighborhoods with a higher percentage of

**Table 1**

Success Rates in Obtaining Cooperation from Families Identified from Birth Records.

a) Denotes percentage of those families with live twin pairs, (b) Denotes percentage of those families who completed the initial screening interview, and (c) Denotes probability that the 2nd twin in a family has completed an interview, given that the first twin has been interviewed.

Target Intake Age:	Birth Records			Screening Interview			Twin Interview		Parent Interview	
	Total Pairs	Pairs w/Deceased Twin	Live Pairs	Screening Interview Completed (a)	Family Not Located (a)	Families w/Twin Interviewed (a/b)	Complete Pairs Assessed (a/b/c)	Families w/Parent Interview (a/b)		
19	551	63	488	434 (88.9%)	24 (4.9%)	373 (76.4%/85.9%)	343(70.3%/79.0%/92.0%)	370 (75.8%/85.3%)		
17	515	51	464	414 (89.2%)	20 (4.3%)	367 (79.1%/88.7%)	348(75.0%/84.1%/94.8%)	371 (80.0%/89.6%)		
15	577	55	522	462 (88.5%)	19 (3.6%)	391 (74.9%/84.6%)	363(69.5%/78.6%/92.8%)	394 (75.5%/85.3%)		
13	930	52	878	737 (83.9%)	43 (4.9%)	616 (70.1%/83.6%)	579(66.0%/78.6%/94.0%)	641 (73.0%/87.0%)		
11	436	34	402	333 (82.4%)	11 (2.7%)	—	—	285 (70.9%/85.6%)		
Total	3009	255	2754	2380 (86.4%)	117 (4.3%)	1728 (73.5%/84.4%)	1633(69.4%/79.8%/94.5%)	2061 (74.8%/86.6%)		

**Table 2**

Association Between Family Non-participation in Initial Screening Interview, and Neighborhood Characteristics Defined by 1990 US Census Block Group Data. For All Variables Except Urban-rural Residence, Successive Categories Are the Lowest Quartile, Middle to 75%, and highest quartile Based on the Entire State Population. For Example, in the Entire Population, One-quarter of Neighborhoods Have Fewer than 23.2% Residents Aged 19 or Younger; One-half Have Between 23.2% and 32.6% of Residents in This Age Range; and One-quarter Have More than 32.6% of Residents in This Age Range.

		Non-participants	Participants	Unadjusted		Adjusted	
		(N = 284)	(N = 2339)	OR	95%CI	OR	95%CI
Neighborhood age structure	% 0–19 years						
	0–23.1%	16.4	17.3	1.0	—	—	—
	23.2–32.6%	51.6	52.3	1.02 <sup>NS</sup>	0.71–1.47	—	—
	32.7%+	32.0	30.4	1.09 <sup>NS</sup>	0.73–1.62	—	—
	% 65+ years						
	0–9.4%	32.7	34.4	1.0	—	—	—
	9.5–20.4%	49.8	49.7	1.06 <sup>NS</sup>	0.79–1.44	—	—
	20.5%+	17.4	15.9	1.15 <sup>NS</sup>	0.78–1.71	—	—
Household structure	% 1-person households						
	0–16.6%	27.8	31.4	1.0	—	—	—
	16.7–33.3%	50.9	50.4	1.14 <sup>NS</sup>	0.83–1.58	—	—
	33.4%+	21.4	18.2	1.32 <sup>NS</sup>	0.90–1.95	—	—
	% Families with married couples and related children <18		0				
	0–26.4%	28.8	20.0	1.0	—	1.0	—
	26.5%–43.7%	46.6	47.8	0.68	0.50–0.92	0.97 <sup>NS</sup>	0.69–1.37
	43.8%+	24.6	32.2	0.53	0.36–0.77	0.85 <sup>NS</sup>	0.55–1.32
	% Families with female head household, no husband, child(ren) <18						
	0–2.7%	19.2	20.3	1.0	—	1.0	—
	2.8–12.9%	44.5	55.0	0.85 <sup>NS</sup>	0.60–1.21	0.76 <sup>NS</sup>	0.54–1.08
	13.0%+	36.3	24.8	1.54	1.08–2.21	0.79 <sup>NS</sup>	0.50–1.23
% 5+ person households							
0–5.2%	19.9	18.4	1.0	—	1.0	—	
5.3–13.2%	47.3	51.6	0.85 <sup>NS</sup>	0.81–1.18	0.85 <sup>NS</sup>	0.61–1.18	
13.3%+	32.7	30.1	1.00 <sup>NS</sup>	0.70–1.44	1.00 <sup>NS</sup>	0.70–1.44	
Neighborhood socioeconomic characteristics	% Residents with some college						
	0–23.1%	24.9	19.8	1.0	—	1.0	—
	23.2–47.3%	47.0	48.2	0.76 <sup>NS</sup>	0.56–1.05	1.04 <sup>NS</sup>	0.74–1.45
	47.4%+	28.1	32.1	0.68	0.48–0.98	1.29 <sup>NS</sup>	0.80–2.09
	% Residents unemployed						
	0–2.9%	18.5	28.3	1.0	—	1.0	—
	3.0–9.6%	48.0	49.8	1.48	1.04–2.09	1.33 <sup>NS</sup>	0.93–1.89
	9.7%+	33.5	22.0	2.36	1.64–3.39	1.43 <sup>NS</sup>	0.92–2.24
	% Households below poverty level						
	0–5.3%	19.2	32.8	1.0	—	1.0	—
	5.4–22.7%	48.8	46.9	1.77	1.26–2.49	1.73	1.01–2.97
	22.8%+	32.0	20.3	2.71	1.90–3.88	1.97	1.00–3.86
	Median household income						
	0–\$17,142	27.8	18.0	1.0	—	1.0	—
	\$17,143–\$31,316	45.9	45.3	0.64	0.47–0.87	0.93 <sup>NS</sup>	0.64–1.35
	>\$31,316	26.3	36.7	0.45	0.32–0.64	1.06 <sup>NS</sup>	0.57–1.98
	% Households with income >\$75K						
	0%	33.5	23.3	1.0	—	1.0	—
0.0–5.8%	41.6	43.8	0.66	0.49–0.90	0.81	0.58–1.13	
>5.8%	24.9	32.9	0.53	0.37–0.75	0.80	0.48–1.33	
Urban-rural residences	% Urban	—	63.5	60.0	1.0	—	—
	% Rural, non-farm	—	11.4	11.8	1.09 <sup>NS</sup>	0.71–1.70	—
	% Rural, farm	—	25.2	28.3	1.19 <sup>NS</sup>	0.87–1.62	—
Neighborhood ethnic composition	% African-American						
	0	28.8	36.7	1.0	—	1.0	—
	0.0–7.4%	28.8	35.5	1.03 <sup>NS</sup>	0.74–1.44	1.16 <sup>NS</sup>	0.81–1.64
	>7.4%	42.4	27.7	1.96	1.43–2.69	1.45 <sup>NS</sup>	0.95–2.20
	% Hispanic						
	0	48.8	46.2	1.0	—	—	—
	0.0–1.3%	21.4	27.6	0.73 <sup>NS</sup>	0.53–1.01	—	—
	>1.3%	29.9	26.2	1.07 <sup>NS</sup>	0.78–1.47	—	—
	% Multiple ancestry						
0–25.9%	36.7	24.9	1.0	—	1.0	—	
26.01–40.9%	44.8	45.4	0.66	0.49–0.89	1.03	0.70–1.50	
41.0%+	18.5	29.8	0.42	0.29–0.59	0.78	0.48–1.26	

households consisting of a single mother with one or more children, and lower percentage of households with two parents plus children; a slightly lower percentage of residents with at least some college education; a higher percentage of unemployed residents; a higher percentage of households below poverty level, and a lower percentage of households with annual income above \$75,000; and a higher percentage of minority (African-American and Hispanic) but a lower proportion of mixed ancestry residents. In a multivariate logistic regression analysis jointly predicting family participation in the screening stage from all of these variables, however, the only neighborhood characteristic that remained predictive of non-participation was living in a neighborhood with a high proportion of the households below the poverty level (highest 25%: OR = 1.97; middle 50%: OR = 1.73). The effect of living in a neighborhood with a high proportion of African-American residents was just short of significance (OR = 1.44, 95%CI 0.95–2.20). When we fitted a backwards selection stepwise model, living in a neighborhood with a high proportion of households below the poverty level remained a significant predictor (highest 25%: OR = 2.15, 95%CI 1.45–3.19; middle 50%: OR = 1.67, 95%CI 1.19–2.35) and living in a neighborhood with a high proportion of African-American residents also reached statistical significance (OR = 1.57, 95%CI 1.16–2.12). The magnitude of the differences between neighborhoods of participant and non-participant families in terms of percent families living below the poverty level (13.3% vs. 17.3%, data not shown) and percent African-American residents (13.9% versus 21.0%) was however not large.

Since the census variable categories (lowest 25%, middle 50%, highest 25%) shown in Table 2 are defined by census data for the entire state, the extent to which the observed proportions for these categories deviate from 25%, 50%, 25% provides information about the ways in which neighborhoods where twins are born differ, compared to neighborhoods in the state as a whole. Thus for example mothers of twins on average are living in neighborhoods with a higher than expected proportion of young

people, and lower than expected proportion of older people; and are more likely to be living in neighborhoods with families with children, and hence more likely to be living in high density neighborhoods. However, without data from control families with non-twin children, it is not possible to determine whether such differences are entirely due to differences between households with and without young children.

#### Birth Record Predictors of Non-participation

Table 3 summarises individual characteristics identified from birth records that predicted non-participation in the initial screening interview phase. For those birth years where only maternal information was available from state birth records, young maternal age at the birth of the twins (19 or younger), and maternal birth out of state, predicted increased probability of failing to obtain a screening interview with the family. Where both maternal and paternal information was available, the absence of information about the biological father (presumably in many cases indicating fathers who were not expecting to be involved in the care of their children), and father's birth out-of-state, also indicated increased probability of failing to obtain a screening interview. In the latter analysis only, maternal African-American ethnicity also predicted decreased probability that a family would participate in the initial screening interview.

#### Conclusions

Inferences from twin data will be most credible when the representativeness of the twin sample is carefully documented. In this paper, we have discussed the ascertainment of a large cohort of adolescent female twin pairs from birth records for a mid-western U.S. state. Using birth cohort ascertainment itself introduces certain limits to the generalisability of any findings that can be made. For example, geographically mobile families who moved into the state after the birth of their twins will not be represented in the sample, though geographically mobile families who moved out of state are represented. Nonetheless, given the limitations of other

**Table 3**

Association Between Family's Non-participation in the Initial Screening Interview (Not Located, or Not Found), and Birth Record-derived Information about Ethnicity, Parental State of Origin, and Parental Age at the Birth of the Twins.

			Unadjusted Odds Ratios		Adjusted Odds Ratios	
	Non-participants	Participants	OR	95%CI	OR	95%CI
1978–1986 Birth Cohorts <sup>a</sup>	(N = 249)	(N = 1897)				
African-American ethnicity	27.7	14.1	2.34	1.74–3.16	1.42NS	0.98–2.06
Mother born in state	54.4	68.3	0.56	0.43–0.72	0.50	0.38–0.66
Mother $\leq$ 19 at birth of twins	20.6	10.7	2.17	1.55–3.03	1.57	1.08–2.27
Paternal data missing	33.5	15.2	2.82	2.12–3.74	2.29	1.58–3.32
1980–1986 Birth Cohorts <sup>b</sup>	(N = 219)	(N = 1559)				
African-American ethnicity	28.6	13.0	2.68	1.94–3.72	1.54	1.04–2.28
Mother born in state	54.8	68.2	0.57	0.43–0.75	0.53	0.39–0.72
Father born in state <sup>c</sup>	28.3	49.5	0.40	0.30–0.55	0.66	0.46–0.95
Mother # 19 at birth of twins	20.1	9.6	2.38	1.66–3.42	1.68	1.12–2.54
Paternal data missing	35.6	14.8	3.18	2.36–4.29	2.02	1.32–3.09

Notes: a) birth cohorts with maternal data in birth record

b) birth cohorts with paternal data in birth record

c) father born in state, versus father born out-of-state or father's place of birth unknown.

potential sampling procedures, such as excessive cost of using standard household survey procedures because of the rarity of twin births, inability to determine generalisability when ascertainment via appeals for volunteer families through the media are conducted, and incomplete ascertainment of families through school systems (including potential undersampling of pairs discordant for school attendance or for grade level), we considered birth record ascertainment to represent the best option for ascertaining an adolescent twin pair sample.

Using birth record data, together with census data, we were able to investigate potential sampling biases associated with failure to locate and/or failure to obtain cooperation from families. Birth of the twins to a teenage mother, having a mother or father born out of state, having an absentee biological father (imputed from the absence of information about the father on the birth record), and birth to a mother living in a neighborhood with higher than average rates of poverty and of African-American residents were all associated with an increased probability of failure to recruit a family into a study. African-American ethnicity also predicted decreased probability of recruitment. Identification of these predictors of non-recruitment will in the future allow use of sampling weights to correct for the under-representation of such families in our research.

### Acknowledgments

This research was supported by NIAAA grants AA09022 and AA11998 and NIDA grant DA12540.

### References

- Bucholz, K. K., Heath, A. C., & Madden, P. A. F. (2000). Transitions in drinking in adolescent females: Evidence from the Missouri female adolescent twin study. *Alcoholism: Clinical and Experimental Research*, *24*, 914–923.
- Glowinski, A. L., Bucholz, K. K., Nelson, E. C., Fu, Q., Madden, P. A. F., Reich, W., & Heath, A. C. (2001). Suicide attempts in an adolescent female twin sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 1300–1307.
- Heath, A. C., Slutske, W. S., & Madden, P. A. F. (1997). Gender differences in the genetic contribution to alcoholism risk and to alcohol consumption patterns. In R. W. Wilsnack, and S. C. Wilsnack (Eds.), *Gender and Alcohol* (pp. 114–149). Rutgers, NJ: Rutgers University Press.
- Heath, A. C., Madden, P. A. F., & Martin, N. G. (1998). Assessing the effects of cooperation bias and attrition in behavioral genetic research using data-weighting. *Behavior Genetics*, *28*, 415–427.
- Heath, A. C., Madden, P. A. F., & Bucholz, K. K. (1999a). Ascertainment of a twin sample by computerized record matching, with assessment of possible sampling biases. *Behavior Genetics*, *29*, 209–219.
- Heath, A. C., Madden, P. A. F., Grant, J. D., McLaughlin, T. L., Todorov, A. A., & Bucholz, K. K. (1999b). Resiliency factors protecting against teenage alcohol use and smoking: Influences of religion, religious involvement and values, and ethnicity in the Missouri Adolescent Female Twin Study. *Twin Research*, *2*, 145–155.
- Heath, A. C., Todorov, A. A., Nelson, E. C., Madden, P. A. F., Bucholz, K. K., & Martin, N. G. (in press). Gene-environment interaction effects on behavioral variation and risk of complex disorders: The example of alcoholism and other psychiatric disorders. *Twin Research*.
- Hudziak, J., Heath, A. C., Madden, P. A. F., Reich, W., Bucholz, K. K., Slutske, W. S., Bierut, L., Neuman, R., & Todd, R. (1998). The latent class and factor analysis of DSM-IV ADHD: A Twin Study of Female Adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*, 848–857.
- Jacob, T. J., Sher, K., Bucholz, K. K., True, W. T., Sirevaag, E. J., Rohrbach, J., Nelson, E., Neuman, R. J., Todd, R. D., Slutske, W. S., Whitfield, J. B., Kirk, K., Martin, N. G., Madden, P. A. F., & Heath, A. C. (2001). An integrative approach for studying the etiology of alcoholism and other addictions. *Twin Research*, *4*, 103–118.
- Meyer, J. M., Silberg, J. L., Simonoff, E., Kendler, K. S., & Hewitt, J. K. (1996). The Virginia twin-family study of adolescent behavioral development: Assessing sample biases in demographic correlates of psychopathology. *Psychological Medicine*, *22*, 1119–1133.
- Nelson, C. B., Little, R. J. A., Heath, A. C., & Kessler, R. C. (1996). Patterns of DSM-III-R alcohol dependence symptom progression in a general population survey. *Psychological Medicine*, *26*, 449–460.
- Nelson, E. C., Grant, J. D., Bucholz, K. K., Glowinski, A., Madden, P. A. F., Reich, W., & Heath, A. C. (2000). Social phobia in a population-based female adolescent twin sample: Comorbidity and associated suicide-related symptoms. *Psychological Medicine*, *30*, 797–804.
- Neuman, R. J., Heath, A. C., Hudziak, J. J., Reich, W., Bucholz, K. K., Madden, P. A. F., Sun, L., & Todd, R. D. (2001). Latent class analysis of ADHD and comorbid symptoms in a population sample of adolescent female twins. *Journal of Child Psychology and Child Psychiatry*, *42*, 933–94.