Behav Genet (2007) 37:668–677
DOI 10.1007/s10519-007-9163-2

ORIGINAL PAPER

Estimation of Variance Components for Age at Menarche in Twin Families

Carl A. Anderson · David L. Duffy · Nicholas G. Martin · Peter M. Visscher

Received: 7 December 2006 / Accepted: 4 July 2007 / Published online: 7 August 2007

Abstract Age at menarche (AAM), time of first menstrual period, is an important developmental milestone in females. Follow-up data from 1,302 adolescent twins and their sisters were used to partition the normal variation in AAM. The proportion of censoring was 14.1%. Both a standard and a survival analysis method were used. The best fitting model from the survival analysis method was an ACE model, where 57% and 23% of the variance in AAM was explained by additive genetic and environmental effects, respectively. The best fitting model when using a standard variance decomposition method was an AE model, where 82% of the variance was explained by additive genetic effects. The lack of correspondence between the results of the two methods was an artefact of the different ascertainment of AAM reports from siblings and twins. After the removal of the sibling sample, both methods indicated that an ACE model was the most likely. Standard and survival analysis methods estimated the proportion of variance explained by additive effects to be 0.50 and 0.54, and common environmental effects to be 0.31 and 0.29, respectively. We conclude that variation in AAM can be explained by additive genetic and common environmental components.

Keywords Survival analysis · Age at menarche · Variance components · Frailty model · Censoring · Twin families

Edited by Dorret Boomsma

C. A. Anderson · D. L. Duffy · N. G. Martin · P. M. Visscher
Genetic Epidemiology Group, Queensland Institute of Medical Research, Brisbane 4029, Australia

C. A. Anderson
Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, Scotland, UK

C. A. Anderson (✉)
Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK
e-mail: carl.anderson@well.ox.ac.uk

Introduction

The onset of menses is an important event both biologically and socially. Age at menarche (AAM) has been identified as a risk factor for several traits including depression (Kaltiala-Heino et al. 2003), eating disorders (Kaltiala-Heino et al. 2001) and breast cancer (Velie et al. 2006) and it is thought to be an important evolutionary trait (Kirk et al. 2001). AAM is a complex trait which is determined by an array of genetic and environmental factors. Several twin studies have been carried out to partition inter-individual trait variation in age at menarche into genetic and environmental components and their findings are summarised in Table 1. Genetic factors clearly play a role in age at menarche, with monozygotic (MZ) twin correlations in the range of 0.51–0.95, and dizygotic (DZ) twin correlations in the range of 0.17–0.58 corresponding to estimated heritabilities in the range 0.30–0.95. Studies which make use of other family structures support these findings. A recent study of age at menarche, carried out using family data from the Fels Longitudinal Study (Roche 1992), reported a heritability (h^2) of 0.49 (95% CI = 0.24–0.73) (Towne et al. 2005). The study analysed data from 371 white females from extended families and found not only a significant genetic effect on age at menarche, but also a year of birth effect that explained 0.02 of the residual phenotypic variation. A twin study suggested that common

Table 1 Previous twin studies of age at menarche

Study	Number of pairs		r_{MZ}	r_{DZ}	Best fitting model	Heritability (h^2)
	MZ	DZ				
Treloar and Martin (1990)[a]	1,177	711	0.65	0.18	A+D+E	0.61–0.68[b]
Meyer et al. (1991)[a]	1,178	711	0.65	0.18	A+D+E	0.71
Kaprio et al. (1995)	234	189	0.75	0.31	A+E	0.74
Loesch et al. (1995)	44	42	0.95	0.58	–	0.95
Kirk et al. (2001)	1,373	1,310	0.51	0.17	A+D+E	0.50
van den Berg and Boomsma (2007)[c]	1,340	793	0.71	0.30	A+A*C+E	0.70

MZ = Monozygotic twin pair, DZ = Dizygotic twin pair. Best fitting model described in terms of (A) additive genetic, (C) common environment, (D) dominant genetic and/or (E) non-shared environmental effects

[a] The Treloar and Martin (1990) and Meyer et al. (1991) studies used the same cohort of individuals

[b] Heritability varied between several analysed age-cohorts

[c] Total sample included 4,995 individual twins, 1,296 sisters, 2,946 mothers and 635 female spouses of male twins

genes (but different environments) influence the sequence of pubertal events (van den Berg et al. 2006) of which AAM is the most readily scored. A study of extended twin families by the same group proposed that 70% of the true variation in AAM is underpinned by additive genetic factors and a further 1.5% by gene-environment interactions (van den Berg and Boomsma 2007).

Of the six twin studies described in Table 1, two were carried out on samples only containing adolescents and these frequently contain censored observations, which occur when the true time of event, in this case age at menarche, is unknown. For individuals with censored data, it is only known that they had not experienced menarche at the time of last interview. Kaprio et al. (1995) collected age at menarche data from 323 Finnish twin pairs who were within 3 months of their sixteenth birthday at the time of data collection. No censored observations were present in the sample. Loesch et al. (1995) analysed age at menarche data from a small sample of Polish adolescent twin pairs. The twins were examined annually throughout adolescence, up to the age of 18. As a result of the long period of follow-up, no censored observations were present in the sample. The adolescent samples from both Kaprio et al. (1995) and Loesch et al. (1995) are small so it is difficult to compare the results to the larger adult cohorts. Furthermore, the differences seen between the adolescent and adult studies (see Table 1) are potentially due to the different study designs.

It is important to use age at menarche data collected from adolescent samples when carrying out genetic analysis of age at menarche because reports of AAM in adult samples have been shown to be inaccurate. In a sample of 60 women with a known age at menarche, Damon et al. (1969) reported a correlation between actual and recalled age at menarche of 0.78 when recall was requested approximately 19 years post-menarche. When recall was attempted 39 years post-menarche the correlation coefficient between actual and recalled age at menarche was 0.60 (Damon and Bajema 1974). Koo and Rohan (1997) showed that even over a period of 3 years the accuracy of age at menarche recall decreases. They reported that after an interval of 1–2 years only 59% of females could recall the exact year and month of their menarche, while 77% of women were accurate to within one month. If, in an attempt to remove the recall bias of adult samples, one ascertains early adolescent samples then a large proportion of the individuals will be censored for age at menarche. A study design which both uses adolescent females and has sufficient follow-up to remove censoring is therefore required for accurate studies into the genetics of age at menarche. To date, the only prospective and longitudinal twin study that has been carried out to investigate the genetics of age at menarche is the Loesch et al. (1995) study, and this only had a sample size of 44 MZ and 42 DZ twin pairs. In large adolescent samples of age at menarche data, censored observations will occur, and survival analysis methods are needed to correctly account for these in the statistical model.

Survival analysis methodology

The semi-parametric Cox proportional hazards method remains the method of choice for most survival analyses. However, the Cox model relies on the survival times of individuals being independent, and this is not always the case. Individuals can be grouped in such a way that their survival times become correlated. For example, individuals could belong to the same family. If this were the case, and there was a common environmental effect on survival time, then one would expect the survival times of family members to be correlated. If the Cox model is applied to non-

independent data then the model parameters are overestimated (Wei et al. 1989). Special methods are needed to analyse data when survival times are correlated.

Frailty models have been derived to model non-independent survival data. A frailty is an unobserved random effect which acts multiplicatively on the baseline hazard. A shared frailty is a random effect which is the same for all members of a group, for example a family effect (Xue and Brookmeyer 1996). With this model, families with a large frailty will experience the event earlier than families with a small frailty. The model therefore allows for the presence of both 'frail' and 'robust' families (Klein and Moeschberger 1999). However, for genetic studies, fitting only a single family effect is unappealing because people within a family are related to differing degrees. To model the genetic relationship of individuals within a family a correlated frailty method must be adopted. These methods fit a per-individual random effect which is correlated according to a relationship matrix. While most random effects methods assume that the frailties follow a gamma distribution, a Gaussian random effects model is most easily generalised to arbitrary covariance matrices. Ripatti and Palmgren (2000) proposed a mixed effects Cox model which includes both fixed and random effects and is given by

$$\lambda(t) = \lambda_0(t)e^{(X\beta+Zb)} \quad (1)$$

where $\lambda_0(t)$ is the baseline hazard, $\mathbf{X}$ is a fixed effect matrix, $\mathbf{Z}$ is a random effect matrix and β and b are the corresponding parameter vectors.

Like the traditional Cox model, the mixed effects Cox model is semi-parametric and does not require the distribution of the baseline hazard to be specified. In addition, the model retains the proportional hazards framework as it is assumed that the conditional individual-specific hazards are proportional over time. The model can be applied to estimate variance components in outbred populations. If an identity by descent (IBD) matrix and a relationship matrix are included as random effects matrices, the model can also be used for mapping QTL in outbred populations. Using this statistical model, Zhao (2005) found that marker D4S1645 contributed significantly to the variance in alcohol dependence in 143 Genetic Analysis Workshop 14 families drawn from the Collaborative Study on the Genetics of Alcoholism.

In the present study, a large sample of adolescent MZ and DZ pairs and their siblings is used to partition the normal variation in age at menarche into genetic and environmental components. The adolescent twins were first seen close to their 12th birthday and were followed up at ages 14 and 16. The age at menarche data is therefore accurate due to recall close to menarche. The sample contains only a small proportion of censored individuals due to the extended period of follow-up. The present study is the largest prospective and longitudinal twin study carried out to investigate variation in age at menarche. The secondary aim of the study is to compare the use of standard and survival methodologies for the analysis of twin data with a small proportion ($\sim$15%) of censored data. It has previously been shown that with a small number of approximations, standard methodology can be successfully applied to map QTL underlying censored traits (Anderson et al. 2006).

Methods

Adolescent twin families

Adolescent twins and their families were recruited for an ongoing study of melanoma risk factors at Queensland Institute of Medical Research, Australia. Twins were interviewed at ages 12 and 14. Every effort was made to interview twins close to their 12th and 14th birthdays, though this was not always possible. In 31 cases the 'age 12' interview was conducted when the individual was aged 13. In a further 6 cases the 'age 12' interview was held when the individual was aged 14. These 6 individuals (and two others) had their 'age 14' interview while aged 15, all other individuals were aged 14. Non-twin siblings were asked to attend the interview if they were more than 10 years old and had not previously attended (i.e., the siblings attend the clinic for interview once only). As part of the clinical protocol, described by Zhu et al. (1999), female adolescents were asked during interview to provide the date of their first menstrual period. Date of first menstrual period and date of birth were used to calculate the age at menarche for each individual (in months). Two age at menarche measures were potentially available for the twins who attended both the 'age 12' and 'age 14' interviews. For individuals with repeat observations the correlation between the age at menarche reported at age 12 and that given at age 14 was calculated (using only those individuals aged 12 and 14 at their first and second interviews, respectively). The present analysis uses age at menarche data collected between May 1992 and February 2006.

A second sample of adolescent twin families was recruited to an ongoing study of cognitive ability, again at Queensland Institute of Medical Research, Australia. Twins were interviewed at age 16, with siblings asked to attend if they were 10 years of age or older. As part of the clinical protocol, described previously by Wright et al. (2001), female participants were asked by a research nurse to provide the date of their first menstrual period. The same procedure as implemented in the melanoma risk factor

study was used to calculate the age at menarche. A subset of the 16-year-old cohort (324 individuals) was asked, via a telephone interview at a later date, to give their age at menarche. These individuals attended for interview before age at menarche data was introduced to the study protocol. This current analysis uses age at menarche data collected between July 1996 and February 2006.

Where two or more age at menarche estimates were available for an individual, the estimate provided at the first data collection following menarche was used. It was assumed that the recall closest in time to menarche would be the most accurate. The date of interview was recorded for all individuals. If an individual was censored (i.e., the true age at menarche was unknown because the individual had not started menstruating at the time of last interview), the age at last seen (ALS) was used in the analysis. It is typical in survival analysis to include ALS as a censored observation in the analysis. Thus, including ALS in the standard analysis (which fails to account for the censoring) allows us to directly compare the two methods.

For ease of computation, 30 females were removed from the data set because they were either the last born member of a triplet or members of a second twin pair. In total, the data consisted of age at menarche information for 1,351 adolescent twins and their siblings, 226 (16.73%) of whom had a censored age at menarche. Univariate outliers were identified as individuals reporting an age at menarche more than 4 standard deviations outside of the mean (less than 104 or greater than 208 months). In total, 6 individuals were identified as outliers and removed from further study to ensure that only the normal variation (and not extreme variation) in AAM was retained.

Siblings were asked to attend the clinic for interview if they were 10 years of age or older. Twins attended the clinic for interview at ages 12, 14 and/or 16. It is unlikely that a 10–12-year-old sibling would have started menstruating, and therefore they would be censored for age at menarche. In this scenario, the age at last seen (10–12 years) was used as an age at menarche. The minimum censored age at menarche (age at last seen) for a twin is 12 years. Not only can siblings attend the clinic at a younger age than the twins, but they can also attend when older than 16 years of age. The range of the censored age at menarche estimates for the siblings and twins is 116–208 months and 144–194 months, respectively. The presence of siblings older than 16 years will have a lesser effect on the age at menarche variance because it is unlikely that these individuals will be censored. Thus, because age at last seen is only used when an individual has not started menstruating, it will be used rarely for siblings of 16 years and above. However, given the findings of Koo and Rohan (1997) regarding the reduction in age at menarche recall accuracy over a period of 1–2 years, it is likely that older siblings provided a less accurate estimate of age at menarche.

To ensure that all individuals within the study had the same opportunity to experience menarche, the siblings with an age at interview of less than 12 years were removed from the study. A total of 223 non-twin sisters, including 21 sibling pairs, had an age at interview greater than 12 years. The final sample consisted of 1,302 adolescent females, 184 (14.1%) of whom had a censored age at menarche.

Estimation of variance components

Twin pair correlations can be used to decompose inter-individual trait variation into genetic and environmental components. Inferences are based on the genetic similarity of monozygotic versus non-monozygotic (DZ twins and siblings) twin pairs; monozygotic twins share all their genes in common and non-MZ pairs share on average half of their segregating genes in common. When the phenotypic correlation between monozygotic twin pairs is greater than that of non-MZ twin pairs, it is assumed that genetic influences underlie the increased familiality. If the phenotypic correlation between non-MZ pairs is more than half the phenotypic correlation between monozygotic twin pairs, a common environmental effect on the trait is indicated. If the phenotypic correlation between non-MZ pairs is less than half the phenotypic correlation between monozygotic twin pairs, then this indicates genetic dominance or an epistatic (gene-gene interaction) effect on the trait (or that the common environment effect on MZ twins is not equal to that of the non-MZ sibling pairs). When variance components were being estimated the phenotypic mean and phenotypic variance of the whole sample was used (i.e., separate means and variance were not estimated for siblings, MZ twins and DZ twins). Furthermore, it was assumed that the MZ–DZ, MZ-sibling, DZ–DZ, DZ-sibling and sibling-sibling covariances were equal, and a single correlation was estimated for these non-MZ pairs.

The means and variances of the monozygotic twins, dizygotic twins and siblings were calculated. The covariance and correlation was calculated for monozygotic twin pairs, dizygotic twin pairs, sibling pairs and non-MZ pairs using MX (Neale et al. 2002).

'Non-survival analysis' method

Here, the analysis is carried out assuming age at menarche is normally distributed and that the censoring status of each individuals can be ignored. The amount of phenotypic variance explained by additive genetic (A), common

environmental (C), specific environmental (E) and dominance/epistatic effects (D) was estimated through structural equation modelling, using the software package MX (Neale et al. 2002). Censored observations were included in the analysis by giving the age at last seen as an age at menarche. The censored nature of the data was not accounted for in the statistical analysis. To allow a direct comparison to the survival analysis method, described later, the mean and variance of age at menarche was equated across all zygosity groups. An ACE model, or if the non-MZ pair correlation was less than half the MZ correlation, the ADE model, was fitted to the age at menarche data. More simplified models were fitted in turn to test whether A, C (or D), or both parameters could be dropped from the full model. The fit of each sub-model was assessed by the difference in log likelihood between the sub and full models. Twice the difference in log likelihood follows a χ^2 distribution with the degrees of freedom equal to the difference in degrees of freedom between the sub and full models. For variance components, the distribution of likelihood ratio test statistics under the null hypothesis is a 50:50 mixture of a point mass at zero and a chi-squared distribution with one degree of freedom (Self and Liang 1987; Stram and Lee 1994). A chi-squared goodness of fit test was used to directly compare the full model to the reduced models. *P*-values were calculated from a chi-square distribution with 1 degree of freedom and subsequently divided by a factor of two. A *P*-value of less than 0.05 indicates a significant reduction in the fit of the model. If the ACE model provides the best fit to the data, the heritability is reported as the proportion of the variance explained by additive genetic effects. If the ADE model is the best fitting model then the broad-sense heritability is the sum of the proportions of variance explained by additive and dominant genetic effects.

Survival analysis method

To investigate the effect of modelling the censored observations correctly in the biometrical analysis, the general mixed-effects Cox model of Ripatti and Palmgren (2000) was used. The analysis was carried out using the UNIX-based S-PLUS package KINSHIP (Therneau 2003). The package was ported into the R environment (R Development Core Team 2005) for ease of use and free availability. Censored observations were again included in the analysis by inputting the age at last seen as an age at menarche. A status vector was included to distinguish between censored and fully observed age at menarche data, where 0 indicated a censored observation and 1 indicated a fully observed age at menarche. The *makekinship* function within the KINSHIP package creates a symmetrical relationship matrix which is equal to half the nominal relationship matrix used in the non-survival analysis method. The *makekinship* function does not account for identical twins so the relationship matrix was manipulated manually to give the correct genetic relationship. So the relationship matrix remained positive definite, a small constant (0.001) was added to the diagonal of the full matrix.

A common-environment matrix **C** was created by changing all non-zero elements of the relationship matrix to 1. A small constant (0.0001) was again added to the diagonal of the matrix to make it positive definite. A matrix which models dominance effects (**D**) was also created, where the diagonal was fixed to 1.001, MZ pairs had a covariance of 1 and non-MZ pairs had a covariance of 0.25. The variance explained by non-shared environmental effects cannot be estimated by COXME because the error term is not in the linear predictor section of the model. The AC and AD models were fitted to the age at menarche data. Simplified models, A only and C/D only, were then fitted in turn to test if the A or C/D parameters could be dropped from the full model. The fit of each sub-model was evaluated using the difference in integrated likelihoods between the sub and full models. The difference in integrated likelihoods follows a chi-square distribution with degrees of freedom equal to the difference in degrees of freedom between the models. The chi-squared statistic was subsequently divided by a factor of two to ensure the correct distributional properties. A *P*-value of less than 0.05 indicated a significant reduction in the fit of the model.

The interpretation of the variance θ from COXME is not straightforward. As COXME fits a semi-parametric Cox model to the data, a one-to-one transformation of the age at menarche data does not change the variance explained by the model parameters. Thus, a transformation of the time scale to make the hazard constant will not affect the variance component estimates or the likelihoods of the fitted models. A constant hazard indicates that the survival times are distributed exponentially. The exponential distribution is a special case of a Weibull distribution. Therefore, the method first put forward by Yazdi et al. (2002) for interpreting the variance from a parametric Weibull proportional hazards threshold model can be used. Schneider et al. (2005) extended the method to take into account multiple random effects and to make better use of the proportion of censored observations. Using the method of Schneider et al. (2005), the heritability (h^2) of age at menarche from an ACE model is given by

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_c^2 + \frac{1}{1-P_c}} \quad (2)$$

where σ_a^2 is the variance explained by the additive genetic effects (the relationship matrix), σ_c^2 is the variance

explained by the common environment, and P_c is the proportion of censored observations. With this interpretation, the proportion of variance explained by non-shared environmental effects is quantified by $\frac{1}{1-P_c}$ (the error term). Replacing σ_c^2 with σ_d^2 allows one to quantify the proportion of variance in an ADE model explained by dominance (or epistatic) effects.

Results

Of the total 1,302 females with age at menarche data, 117 individuals reported an age at menarche at both the 'age 12' and 'age 14' interviews (and were aged 12 and 14, respectively, at the time of interview). The Spearman rank correlation between the age 12 and age 14 estimates was 0.75 (Fig. 1). Only 1 individual reported an age at menarche at ages 12, 14 and 16. Summary statistics, including means and variances, are given for the MZ twins, DZ twins and siblings in Table 2. The within-pair correlations for MZ/MZ, DZ/DZ, sibling/sibling and twin/sibling pairs are given in Table 3. The MZ correlation is significantly greater than both the DZ and sibling correlation, suggesting a genetic effect on age at menarche. It should be noted that these means, variances and within-pair correlations were calculated using MX, failing to account for the censored nature of data.

The mean age at menarche within the total sample is 154.9 months, and the variance is 174.3. The sibling variance is not significantly different to that of the MZ or DZ twins. There is a large, though not significant, difference in the correlations between DZ and sibling pairs. The siblings have a much lower correlation than the DZ twins. This is perhaps not unexpected given that the sibling-sibling correlation is calculated using only 21 sister pairs. Assuming a single mean and variance for all zygosity classes, the non-MZ pair correlation for age at menarche is 0.44 (95% CI = 0.35–0.52). Under the above assumptions, the MZ pair, DZ pair and sibling pair correlations are 0.82 (95% CI = 0.78–0.85), 0.57 (95% CI = 0.46–0.66) and 0.23 (95% CI = 0.00–0.55), respectively. The non-MZ correlation is more than half the MZ correlation, therefore suggesting that an ACE model is the correct model to fit to the data.

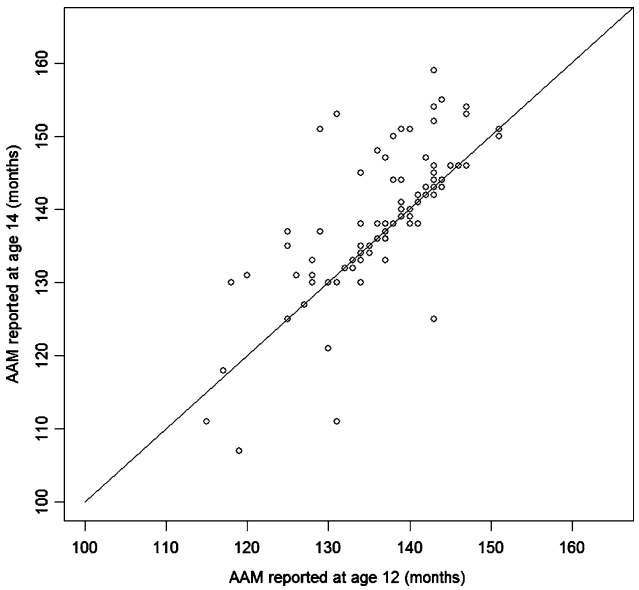

Fig. 1 AAM (in months) reported at age 12 versus age at menarche reported at age 14. From the sample of 1,302 individuals, 117 twins reported an age at menarche at ages of 12 and 14. The solid line denotes the perfect relationship between the two age at menarche recalls

Table 2 Statistical descriptions of the age at menarche data for the MZ and DZ twins and their non-twin sisters

	MZ twins	DZ twins	Non-twin sisters
Number of individuals	446	633	223
Mean (months)	154.6	154.3	156.6
Variance	175.2	169.1	183.1
Standard deviation	13.2	13.0	13.5
Number of censored individuals	62	108	14
Proportion of censored individuals	0.139	0.171	0.063

The DZ sample contains both female-female DZ pairs and female twins from opposite sex twin pairs. Only individuals with an age at last interview of 12 years or more are included

'Non-survival analysis' method

The results of the non-survival biometric analysis, carried out using MX, are given in Table 4. The best fitting model is an AE model. The heritability h^2 of age at menarche within the current adolescent sample, calculated using MX, is 0.82 (95% CI = 0.78–0.85).

Survival analysis method

The results of the biometrical analysis carried out using COXME are given in Table 5. As suggested by the MZ and DZ twin pair correlations, the best fitting model is an ACE model. The approximation of the heritability h^2 of age at menarche within the current adolescent sample, calculated using COXME, is 0.57. COXME does not currently allow the calculation of confidence intervals.

Further analyses were carried out to investigate the different ascertainment of siblings and twins, and its effects on the variance components estimates. The initial analysis,

Table 3 Pairwise statistics for the age at menarche data in MZ pairs, DZ pairs, sibling pairs and twin/non-twin pairs

	MZ/MZ pairs	DZ/DZ pairs	Twin/Sib pairs	Sib/Sib pairs
Number of pairs	223	164	226	21
r (95% CI)	0.82 (0.78–0.85)	0.57 (0.46–0.66)	0.35 (0.21–0.46)	0.23 (0.00–0.55)

All correlations (*r*) and confidence intervals were calculated using MX (assuming one mean and one variance across all zygosity groups). Only individuals with an age at last interview of 12 years or more were included

Table 4 Results from multiple models used to test alternative sources of variation in age at menarche after the removal of siblings less than 12 years old (implemented using a non-survival analysis approach with MX)

Model	V_A (95% CI)	V_C (95% CI)	V_E (95% CI)	A (95% CI)	C (95% CI)	E (95% CI)	−2LL	Δ−2LL	df	*P*
ACE	132.25	10.50	31.47	0.76	0.06	0.18	10,087.23	–	1,297	–
	(102.82–157.25)	(0.00–39.82)	(26.32–38.07)	(0.58–0.85)	(0.00–0.22)	(0.15–0.22)	–	–	–	–
AE*	141.85	–	31.14	0.82	–	0.18	10,087.67	0.45	1,298	0.25
	(127.01–157.75)	–	(26.11–37.45)	(0.78–0.85)	–	(0.15–0.22)	–	–	–	–
CE	–	103.63	71.74	–	0.59	0.41	10,171.87	84.64	1,298	0.00
	–	(88.36–120.78)	(63.68–81.00)	–	(0.53–0.64)	(0.36–0.47)	–	–	–	–
E	–	–	173.98	–	–	1	10,403.22	315.33	1,299	0.00
	–	–	(161.04–187.96)	–	–	(1)	–	–	–	–

Units of variance are months^2. Best fitting model marked by *. Alternative models were tested using the χ^2 goodness of fit test to compare the fit of the ACE model to reduced models as potential components of variation were removed. V_A = Variance explained by additive genetic effects, V_C = Variance explained by common environmental effects, V_E = Variance explained by non-shared environmental effects. A = Proportion of variance explained by additive genetic effects (heritability), C = Proportion of variance explained by common environmental effects, E = Proportion of variance explained by non-shared environmental effects. −2LL = −2loglikelihood, df = degrees of freedom

Table 5 Results from multiple models used to test alternative sources of variation in age at menarche after the removal of siblings less than 12 years old (implemented using correlated frailty Cox models in COXME)

Model	V_A	V_C	A	C	E	−2LL	Δ−2LL	df	*P*
ACE*	3.42	1.42	0.57	0.23	0.20	13,534.20	–	2	
AE	5.15	–	0.82	–	0.18	13,539.67	5.47	1	0.010
CE	–	2.01	–	0.63	0.37	13,573.01	38.81	1	0.000

Best fitting model marked by *. Alternative models were tested using the χ^2 goodness of fit test to compare the fit of the ACE model to reduced models as potential components of variation were removed. V_A = Variance explained by additive genetic effects, V_C = Variance explained by common environmental effects, V_E = Variance explained by non-shared environmental effects. A = Proportion of variance explained by additive genetic effects (heritability), C = Proportion of variance explained by common environmental effects, E = Proportion of variance explained by non-shared environmental effects, $1/1-P_C = 1.165$ (see Eq. 2). All proportions of variance calculated using Eq. 2. −2LL = −2 × integrated likelihood, df = degrees of freedom

described above, selected only those twins with an age at interview of 12 years or more. A second analysis was carried out as described previously, but all siblings were removed from the sample prior to the estimation of variance components. A third analysis, which included all siblings regardless of age at last seen, was carried out using the same methods. A summary of the results from all three analyses is given in Table 6.

It was previously hypothesized that the bias introduced to the non-survival analysis estimation of variance components could be reduced by the selection of siblings with an age at last seen of 12 years or more. The results shown in Table 6 prove this hypothesis to be correct. Given that an ACE model is consistently the best fitting model when using a survival analysis method to estimate the variance components underlying variation in age at menarche, it is assumed that this is the best model to describe the variance in the current data. If the siblings with an age at interview of less than 12 years old are removed from the sibling sample, the non-survival analysis variance component estimates become closer to those estimated using the survival analysis method. The siblings are still introducing a small bias in the non-survival analysis, and this is only completely eliminated when the sibling sample is removed altogether.

Table 6 Results from additional analyses carried out to investigate the influence of the siblings on the variance components estimates from COXME and MX

	All siblings		Siblings 12 years +		No siblings	
	MX	COXME	MX	COXME	MX	COXME
Model	ADE	ACE	ACE	ACE	ACE	ACE
A	0.53	0.56	0.76	0.57	0.50	0.54
C	–	0.24	0.06*	0.23	0.31	0.29
D	0.31	–	–	–	–	–
E	0.16	0.20	0.18	0.20	0.19	0.17

All Siblings includes every sibling in the sample, regardless of the age at last interview. Siblings 12 years + summarises the results reported previously, where only those siblings with an age at last interview of 12 years or greater were included in the analysis. 'No siblings' gives the variance components estimated by MX and COXME when only MZ and DZ pairs are included in the analysis. *Variance component is not significantly different from zero

Discussion

In the present study, if the age at menarche for an individual is censored, the age at last seen is used in the analysis. This method is perhaps the simplest method to adopt, and if the censored nature of the observations is accounted for in the statistical model it introduces no bias to the results. However, if the censored observations are not taken into account when analysing the data, as with the MX analysis carried out here, there is a potential for bias to be introduced in the estimation of the variance components. The different ascertainment of the twins and siblings leads to a greater proportion of the siblings being censored (20.3% of all siblings regardless of age at last seen) than the twins (15.8% of the total twin sample). Furthermore, the mean age at last seen differs between the two groups (138.2 months for siblings and 151.5 months for twins). As the proportion of censoring in the sample increases, so does the bias introduced by failing to account for the censoring in the statistical model. The difference between the censored siblings and censored twins with regard to mean age at last seen is also likely to increase the bias. After the removal of the siblings with an age at last seen less than 12 years old, the percentage of censored siblings dropped to 6.3%, with the mean age at last seen within the censored siblings increasing to 169.4 months. There remains a large difference (17.9 months) between the mean age at last seen of the siblings and twins, and this difference is due to the different ascertainment of the two groups. However, after the removal of siblings less than 12 years old, the total proportion of censoring is much smaller. Hence, the amount of bias introduced to the analysis when using a non-survival analysis method is expected to be significantly reduced by this treatment.

When using COXME the siblings have little influence on the estimation of the variance components. The best fitting model when using COXME is an ACE model, regardless of the inclusion criteria placed on the sibling sample. This indicates that the COXME analysis is robust to the different censoring properties of siblings and twins. However, the non-survival analysis method, carried out using MX, is not robust to the inclusion criteria placed on the sibling sample. If all siblings are included in the analysis the best fitting MX model is an ADE model; if all siblings are excluded from the analysis the best fitting MX model is an ACE model. The inclusion of the siblings clearly has a large effect on the variance component estimates when using a non-survival method. The only difference between the survival and non-survival methods is that the survival method correctly models the censored nature of the data in the statistical model, whereas the non-survival analysis does not. If all siblings are removed from the sample, there is agreement between the MX and COXME analyses, both with regard to the best fitting model (ACE) and the proportions of variance explained by these components. This suggests that the non-survival analysis is sensitive to the inclusion criteria of the siblings because of the different ascertainment and censoring seen in the siblings in comparison to the twins.

The correlations (r) presented in Table 3 show the concordance in reported age at menarche to be greater in DZ pairs than in twin-sib pairs. This suggests a twin specific environmental effect on variance in reported age at menarche. However, because the correlations were estimated using MX, the censored nature of the data was not taken into consideration. Therefore, the different correlations reported for the DZ pairs and non-twin pairs could be due to the different censoring properties of the two groups. To investigate this further, a twin effect matrix (T) was fitted using COXME (which accounts for the censoring present in the sample). If siblings are included in a sample, the presence of a twin-specific environmental influence on a trait can be assessed through the inclusion of an additional environmental variance component that is shared only by twin pairs (Koeppen-Schomerus et al. 2003). If a significant difference exists between the correlations of the

DZ pairs and non-twin pairs, the ACT model (additive genetic, common environment and twin specific variance components) will give a significantly better fit than the model where the twin specific environment effect has been dropped. The results of this investigation suggested that the twin specific environmental effect could be dropped from the model without significantly reducing the fit of the model ($P = 0.242$). We conclude that there is not a twin-specific environmental effect on age at menarche. In an attempt to see if the apparent twin-effect suggested by the twin pair and non-twin pair correlations is an artifact of the censored data, we repeated the above COXME analysis but classified all individuals as uncensored (i.e., we ignored the true censoring status of individuals). The best fitting model included a twin-specific environmental effect, dropping the T component significantly reduced the fit of the model ($P = 3.3 \times 10^{-4}$). We therefore conclude that the difference in correlation between the DZ pairs and non-twin pairs shown in Table 3 is an artifact of the censoring properties of these two groups. Furthermore, our assumption that the variances and co-variances can be equated across zygosity groups appears to be correct.

Towne et al. (2005) review previous twin studies of age at menarche data and conclude that the heritability of age at menarche is approximately 0.50. This study provides evidence to support the findings of Towne et al. (2005). The components underlying variation in age at menarche are much more unclear. Of the 5 biometric studies shown in Table 1, two different models (ADE and AE) are given to describe variation in age at menarche. The present study suggests that the variation in age at menarche is influenced chiefly by additive genetic effects, with approximately 25% of the variance due to common environmental effects. Common environmental effects have been previously reported for age at puberty (Eaves et al. 2004; van den Berg et al. 2006) though this is the first study to suggest a common environmental effect explicitly on age at menarche. Given that this is the first study to use a large prospective and longitudinal sample of adolescent females further such samples are required to support the finding of a common environmental effect on AAM. However, this study does suggest that the dominance components reported in previous twin studies of AAM could be an artifact of recall bias. It is interesting to note that the only previous study to partition variance in AAM using follow-up data from adolescent twins also failed to identify a dominance component (Kaprio et al. 1995). However, the failure to detect a dominance effect, or indeed a common environmental effect, could be due to the sample size of the study (234 MZ and 189 DZ pairs).

In summary, a biometric genetic analysis has been carried out on a sample of adolescent twins and siblings. The heritability of age at menarche was estimated to be 0.57 using a mixed effects Cox model. The analysis was carried out using a correlated frailty model to account for the 16.7% of individuals with a censored age at menarche. The analysis was also carried out without statistically accounting for the censored observations in the model, and when all siblings were removed from the analysis a heritability of 0.50 was reported. The best fitting model under both methods of analysis separated the variance in age at menarche into additive genetic (A), common environmental (C) and non-common environmental (E) effects. This study demonstrates that with a small proportion of censoring (15% or less) standard biometric methodology can be successfully applied to twin studies. However, great care must be taken to ensure the censoring properties of all groups of individuals are consistent. The present study highlights the potential effects of incorrectly accounting for censored data. Ideally, survival methodology should be applied to correctly model the censored nature of the data.

Acknowledgements We would like to thank all the twins and their families for participating in the study and Ann Eldridge and Marlene Grace for ascertaining the age at menarche data. Thanks to Sri Shekar for an excellent introduction to MX. CAA was funded by a Medical Research Council (UK) postgraduate research studentship. The adolescent twin studies were supported by the Australian Research Council (A79600334, A79801419, A79906588, DP0212016), the Human Frontier Science Program (RG0154.1998-B), the NHMRC (950998, 981339, 241944), and the National Institutes of Health (CA88363). The work was also supported by an NHMRC Bioinformatics program grant awarded to PMV (389982). We would also like to thank the editor and three anonymous reviewers for constructive comments on an earlier version of the manuscript.

References

Anderson CA, McRae AF, Visscher PM (2006) A simple linear regression method for quantitative trait loci linkage analysis with censored observations. Genetics 173:1735–1745

Damon A, Bajema CJ (1974) Age at menarche of recall after 39 years. Human Biol 46:381–384

Damon A, Damon ST, Reed RB, Valadian I (1969) Age at menarche of mothers and daughters, with a note on accuracy of recall. Human Biol 41:161–182

Eaves L, Silberg J, Foley D, Bulik C, Maes H, Erkanli A, Angold A, Costello EJ, Worthman C (2004) Genetic and environmental influences on the relative timing of pubertal change. Twin Res 7:471–481

Kaltiala-Heino R, Kosunen E, Rimpela M (2003) Pubertal timing, sexual behaviour and self-reported depression in middle adolescence. J Adolesc 26:531–545

Kaltiala-Heino R, Rimpella M, Rissanen A, Rantanen P (2001) Early puberty and early sexual activity are associated with bulimic-type eating pathology in middle adolescence. J Adolesc Health 28:346–352

Kaprio J, Rimpela A, WinterT, Viken RJ, Rimpela M, Rose RJ (1995) Common genetic influences on BMI and menarche. Human Biol 67:739–753

Kirk KM, Blomberg SP, Duffy DL, Heath AC, Owens IPF, Martin NG (2001) Natural selection and quantitative genetics of life-

history traits in western women: a twin study. Evolution 55: 423–435

Klein JP, Moeschberger ML (1999) Survival analysis: techniques for censored and truncated data. Springer-Verlag, New York, NY, USA

Koeppen-Schomerus G, Spinath FM, Plomin R (2003) Twins and non-twin siblings: different estimates of shared environmental influence in early childhood. Twin Res 6:97–105

Koo MM, Rohan TE (1997) Accuracy of short-term recall of age at menarche. Ann Human Biol 24:61–64

Loesch DZ, Huggins R, Rogucka E, Hoang NH, Hopper JL (1995). Genetic correlates of menarcheal age: a multivariate twin study. Ann Human Biol 22:479–490

Meyer JM, Eaves LJ, Heath AC, Martin NG (1991) Estimating genetic influences on the age-at-menarche: a survival analysis approach. Am J Med Genet 39:148–154

Neale MC, Boker SM, Xie G, Maes HH (2002) MX: Statistical Modeling

R Development Core Team (2005) R: A language and environment for statistical computing

Ripatti S, Palmgren J (2000) Estimation of multivariate frailty models using penalized partial likelihood. Biometrics 56:1016–1022

Roche AF (1992) Growth, maturation and body composition: the Fels Longitudinal Study. Cambridge University Press, Cambridge

Schneider MdP, Strandberg E, Ducrocq V, Roth A (2005) Survival analysis applied to genetic evaluation for female fertility in dairy cattle. J Dairy Sci 88:2253–2259

Self SG, Liang K-Y (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J Am Stat Assoc 82:605–610

Stram DO, Lee JW (1994) Variance components testing in the longitudinal mixed effects model. Biometrics 50:1171–1177

Therneau T (2003) On mixed-effect Cox models, sparse matrices, and modeling data from large pedigrees. Mayo Clinic, Rochester, USA

Towne B, Czerwinski SA, Demerath EW, Blangero J, Roche AF, Siervogel RM (2005) Heritability of age at menarche in girls from the Fels Longitudinal Study. Am J Phys Anthropol 128:210–219

Treloar SA, Martin NG (1990) Age at menarche as a fitness trait: nonadditive genetic variance detected in a large twin sample. Am J Human Genet 47:137–148

van den Berg SM, Boomsma DI (2007) The familial clustering of age at menarche in extended twin families. Behav Genet (Epub Ahead of Print)

van den Berg SM, Setiawan A, Bartels M, Polderman TJ, Vaart AWvd, Boomsma DI (2006) Individual differences in puberty onset in girls: Bayesian estimation of heritabilities and genetic correlations. Behav Genet 36:261–270

Velie EM, Nechuta S, Osuch JR (2006) Lifetime reproductive and anthropometric risk factors for breast cancer in postmenopausal women. Breast Dis 24:17–35

Wei LJ, Lin DY, Weissfeld L (1989) Regression analysis of multivariate incomplete failure time data by modelling marginal distributions. J Am Stat Assoc 84:1065–1073

Wright M, Geus Ed, Ando J, Luciano M, Posthuma D, Ono Y, Hansell N, Baal Cv, Hiraishi K, Hasegawa T, Smith G, Geffen G, Geffen L, Kanba S, Miyake A, Martin N, Boomsma D (2001) Genetics of cognition: outline of a collaborative twin study. Twin Res 4:48–56

Xue X, Brookmeyer R (1996) Bivariate frailty model for the analysis of multivariate survival times. Lifetime Data Anal 2:277–289

Yazdi MH, Visscher PM, Ducrocq V, Thompson R (2002) Heritability, reliability of genetic evaluations and response to selection in proportional hazards models. J Dairy Sci 85:1563–1577

Zhao JH (2005) Mixed effects Cox models of alcohol dependence in extended families. BMC Genet 6(Suppl 1):S127

Zhu G, Duffy DL, Eldridge A, Grace M, Mayne C, OGorman L, Aitken JF, Neale MC, Hayward NK, Green AC, Martin NG (1999) A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: a maximum-likelihood combined linkage and association analysis in twins and their sibs. Am J Human Genet 65:483–492