Genetic Study of the Height and Weight Process During Infancy

Paula van Dommelen,[1] Mathisca C. M. de Gunst,[1] Aad W. van der Vaart,[1] and Dorret I. Boomsma[2]

[1]Department of Mathematics, Vrije Universiteit, Amsterdam, the Netherlands
[2]Department of Biological Psychology, Vrije Universiteit, Amsterdam, the Netherlands

Longitudinal height and weight data from 4649 Dutch twin pairs between birth and 2.5 years of age were analyzed. The data were first summarized into parameters of a polynomial of degree 4 by a mixed-effects procedure. Next, the variation and covariation in the parameters of the growth curve (size at one year of age, growth velocity, deceleration of growth, rate of change in deceleration [i.e., jerk] and rate of change in jerk [i.e., snap]) were decomposed into genetic and nongenetic sources. Additionally, the variation in the estimated size at birth and at 2 years of age interpolated from the polynomial was decomposed into genetic and nongenetic components. Variation in growth was best characterized by a genetic model which included additive genetic, common environmental and specific environmental influences, plus effects of gestational age. The effect of gestational age was largest for size at birth, explaining 39% of the variance. The differences between monozygotic and dizygotic twin correlations were largest for size at 1 and 2 years of age and growth velocity of weight, which suggests that these parameters are more influenced by heritability than size at birth, deceleration and jerk. The percentage of variance explained by additive genetic influences for height at 2 years of age was 52% for females and 58% for males. For weight at 2 years of age, heritability was approximately 58% for both sexes. Variation in snap height for males was also mainly influenced by additive genetic factors, while snap for females was influenced by both additive genetic and common environmental factors. The correlations for the additive genetic and common environmental factors for deceleration and snap are large, indicating that these parameters are almost entirely under control of the same additive genetic and common environmental factors. Female jerk and snap, and also female height at birth and height at 2 years of age, are mostly under control of the same additive genetic factor.

The growth of an individual can be viewed as the phenotypic expression of his or her genotype and the influence of environmental factors. In this paper we estimate the influence of genetic and environmental factors on variation in height and weight during the first 2.5 years of age in a large sample of Dutch twin pairs born between 1986 and 1992 (Boomsma et al., 2002). To this aim a two-stage method was used. Firstly, the longitudinal measurements on height and weights of individuals were reduced to parameters of a polynomial of degree 4. This was done separately for mono- and dizygotic male and female twins. Height and weight at birth and at 2 years of age was estimated by interpolating the polynomial of the individuals. Secondly, a multivariate biometric analysis was performed on the fitted coefficients and the interpolated values by decomposing the variances of the parameter values into genetic and environmental components. Because both gestational age and sex have been shown to be significant predictors for physical features of infants (Ben-Amitai et al., 1990), these variables were included as explanatory variables in the second stage of the study.

Growth curve models can be used to describe growth at particular time-points as well as the process of growth over time. They are well suited to analyze longitudinal data when the times of measurement are irregularly spaced and differ for different individuals, as they describe growth with a limited number of interpretable parameters, such as height and weight at 1 year of age, growth velocity, deceleration of growth, rate of change in deceleration (i.e., jerk) and rate of change in jerk (i.e., snap). The uniform description through such growth parameters makes it possible to compare individuals. A number of growth curves have been suggested in the literature and have been shown to be representative at different periods of life (see Karlberg et al., 1987a, 1987b and Thissen & Bock, 1990).

The present study was conducted to expand on previous research on the genetics of height and weight in young Dutch children, in particular the study of Baker et al., 1992. In Baker et al.'s study, longitudinal data on height of a subset (996 twin pairs) of our

sample (2701 pairs for height and 3477 pairs for weight) were summarized by parameters of the quadratic polynomial growth curve via a multiple regression procedure for each individual. These parameters were then subjected to a multivariate biometrical analysis.

A polynomial of degree 4 has been used as the growth curve, which turned out to be better suited than polynomials of degree 2 or 3 to describe growth of children from birth to 2.5 years of age. To estimate the growth curve parameters a mixed-effects model was used. Mixed-effects models are generally used to describe a relationship between a response variable and covariates in data from individuals that are grouped according to one or more classification factors. The grouping is reflected by the fact that each parameter in the model (size at 1 year of age, growth velocity, deceleration of growth, jerk and snap) is the sum of a fixed and a random component. The fixed components are the same for individuals in the same group, but may vary for the different groups. The random components are different for every individual, but for individuals of the same group originate from the same normal distribution. The individuals in one group are thus treated as a random sample from a population of similar individuals (Lindstrom & Bates, 1990). Mixed-effects models are designed to estimate the average behavior of an individual in the population, as well as the variability among and within individuals (Pinheiro & Bates, 2000). An advantage of using mixed-effects models instead of performing multiple regression is a reduction of the number of free parameters. Because the number of observations (i.e., the number of time instances at which measurements were taken) per individual can be small, the parameter estimates obtained by applying a multiple regression procedure to each individual separately are subject to a relatively large estimation error. By using a mixed-effects model we use data from all individuals to estimate the growth coefficients for each individual by taking into account all other individuals. This will result in parameters that fluctuate less between individuals than if multiple regression had been used. Of course, the procedure will be unreliable if the growth curves cannot be described well by a mixed-effects model.

In the second stage of the analyses, the variation in the growth curve parameters was decomposed into genetic and nongenetic components. Two series of multivariate analyses were carried out: one which simultaneously analyzed size at 1 year of age, growth velocity, deceleration of growth, rate of change in deceleration (i.e., jerk) and rate of change in jerk (i.e., snap). The second series of genetic models analyzed size at birth and at 2 years of age. The genetic model fitting was done separately for the height and the weight data.

Materials and Methods

Longitudinal growth data were obtained from the Netherlands Twin Register (NTR) at the Vrije Universiteit in Amsterdam, the Netherlands. Parents

of twins responded to questionnaire items on twin similarity, gestational age, and height and weight as measured by the Youth Health Services up until the age of approximately 2.5 years. Parents were asked to indicate when height and weight had been measured (Boomsma et al., 1992). Similarity items were used to obtain zygosity of same-sex twin pairs. The agreement between zygosity assigned by the replies to the questions and zygosity determined by DNA markers/blood typing was around 93% (Rietveld et al., 2000).

We started with a total of 4649 twin pairs who were born between 1986 and 1992. After checking the data for suitability of inclusion in the analysis, there were 4137 twin pairs for the height analysis and 4154 pairs for the weight analysis. These twin pairs had a known zygosity and both the youngest and oldest of the twin pair had at least one measurement for height or weight. The median number of measurements was 9 (SD = 2.5) for height and 12 (SD = 2.8) for weight per child. The maximum number of measurements was 20.

The twin pairs were divided into six zygosity groups, MZM (monozygotic, males), DZM (dizygotic, males), MZF (monozygotic, females), DZF (dizygotic, females), DOSmf (dizygotic opposite sex, male born first), and DOSfm (dizygotic opposite sex, female born first).

For each child, each individual growth pattern was summarized into the parameters of a polynomial of degree 4 (for descriptions see below). The measurements on height and weight were analyzed separately. To ensure good parameter estimates each child was required to have at least one measurement before the age of 3 months, at least one between 3 months and 1 year and 3 months, and at least one after the age of 1 year and 3 months. With these requirements we have a total of 472 MZM, 434 DZM, 528 MZF, 412 DZF, 447 DOSmf and 408 DOSfm for the height analysis. For the weight analysis there were 587 MZM, 546 DZM, 663 MZF, 543 DZF, 595 DOSmf and 543 DOSfm. The sample of twin pairs with known gestational age is a reduced dataset which consists of 444 MZM, 415 DZM, 505 MZF, 395 DZF, 434 DOSmf and 394 DOSfm twin pairs for height and 550 MZM, 526 DZM, 636 MZF, 518 DZF, 573 DOSmf and 522 DOSfm twin pairs for weight.

Estimates of the polynomial growth parameters for each individual were obtained by first fitting mixed-effects models with maximum likelihood and next computing the estimated conditional modes of the random effects given the observations. In this step the data from DZM and DOS males, as well as from DZF and DOS females were combined, because no large differences in heights and weights between these groups were noticed (van Dommelen et al., 2004). The estimation procedure was therefore based on four groups of individuals, namely MZM, MZF, DZM and DOS males, and DZF and DOS females.

For the computations, the function nlme in Splus 6.1 was used. Growth at birth and 2 years of age was estimated by interpolating the polynomial of the individuals. The individual sets of growth parameters and interpolated values were then subjected to further analyses.

The height and weight data were analyzed by mixed-effects models. A mixed-effects model assumes each growth parameter to be the sum of a fixed and a random component, where the fixed component is the same for every individual, and the random component is different but has the same normal distribution. Therefore, this model accommodates individual variations through the random effects, but ties the individuals together through the fixed effects and the covariance matrix of the random effects. The fixed effects represent the mean values of the parameters in the subpopulation of individuals. The random effects represent the deviations of the individual coefficients from their subpopulation average. Therefore, random effects contribute to the covariance structure of the data. These effects may introduce correlations between cases. In our situation each of the four groups of MZM, MZF, DZM/DOS males, and DZF/DOS females is viewed as a subpopulation with its own parameter values.

Let n be the number of children, t the age in years and $y_i(t)$ the height (in cm) or weight (in kg) of the i-th child at age t. Then, for $i = 1,…, n$, the dependency of the response variables height and weight on age is given by the following polynomial of degree 4 (centred at age 1):

$$y_i(t) = \alpha_1 = \alpha_2(t-1) + \alpha_3(t-1)^2 + \alpha_4(t-1)^3 + \alpha_5(t-1)^4 + \varepsilon_{it}$$

In this model, α_1 represents the height/weight at 1 year of age, α_2 the instantaneous rate of growth at 1 year (velocity), α_3 the amount of deceleration in the individual's growth curve, α_4 represents the rate of change in deceleration (jerk) and α_5 the rate of change in jerk (snap).

A mixed-effects model assumes each growth parameter to be the sum of a fixed and a random component, which is $\alpha_k = \alpha_{k0} + \alpha_{ki}$, with α_{k0} fixed effects and α_{ki} random effects, $k = 1,…,5$. The measurement errors ε_{it} are assumed to be independent across individuals and to be normally distributed with mean zero and a common variance. For the mixed-effects procedure it is assumed that for different individuals the random effects have the same multivariate normal distribution with mean vector zero, and are independent of the measurement errors.

Growth parameters of children can be estimated with only three observations by a fourth-order polynomial mixed-effects model, as this model ties the individuals together through the fixed effects and the covariance matrix of the random effects. Therefore, the model borrows strength across individuals in estimating individual parameters. Thus with three

observations, estimation with a fourth-order polynomial, the mixed-effects model is less of a problem than with the simpler method that estimates the parameters for each individual separately. It is also possible to estimate a child's growth parameters by a fourth-order polynomial mixed-effects model with only one or two observations, but in this study, each child was required to have at least three measurements, or the growth curve would have been smoothed too much towards the average curve.

Size at birth, β_1, and size at 2 years of age, β_2, was obtained by interpolating the polynomial of degree 4 with its estimated parameters.

Genetic Model Fitting

Pearson correlation coefficients were used to summarize twin resemblance for each of the growth parameters (α_1 through α_5 and β_1 and β_2). Pearson correlations were also used to quantify the relationship between gestational age and the growth parameters and size at birth and at 2 years.

Variation between individuals in the growth parameters was analyzed as a function of additive genetic influences, common and specific environment, and gestational age. For each pair of twins, the gestational age, the set of growth parameters and interpolated values were collected in a vector, and multivariate modeling was carried out on the variance–covariance matrices of these vectors with the computer package Mx 1.52 (Neale et al., 1999). The unknown parameters of the multivariate model, which are denoted by the vector θ, were estimated by maximum likelihood under the assumption that the observational vectors, that is, for each pair of twins the vector containing gestational age and the growth parameters or the interpolated values, are sampled independently from a multivariate normal distribution, with the form of the covariance matrix depending on the zygosity group. This corresponds to minimizing with respect to θ for the six twin groups simultaneously a distance function between the covariance matrix $\Sigma_i(\theta)$ of the form particular to the group ($i = 1, 2,…, 6$) and the sample covariance matrix S_i of the observations in the group. In particular, the function

$$F_{ML} = \sum_{i=1}^{6} N_i \left(\log \det \Sigma_i(\theta) - \log \det S_i + \text{trace}(S_i \Sigma_i^{-1}(\theta)) - p\right)$$

was minimized where N_i denotes the number of twin pairs in zygosity group i, and p equals the dimension of the observational vectors (in the first series of analyses $p = 11$, i.e., 5 growth parameters for the oldest twin, 5 for the youngest twin, and gestational age; in the second series of analyses $p = 5$, i.e., 2 interpolated values for the oldest twin, 2 for the youngest twin and gestational age). The likelihood ratio statistic was used to determine the goodness-of-fit of the different models relative to the model with unrestricted covariance matrices (i.e., when the covariance matrices of the six twin groups are estimated by the

Paula van Dommelen, Mathisca C. M. de Gunst, Aad W. van der Vaart, and Dorret I. Boomsma
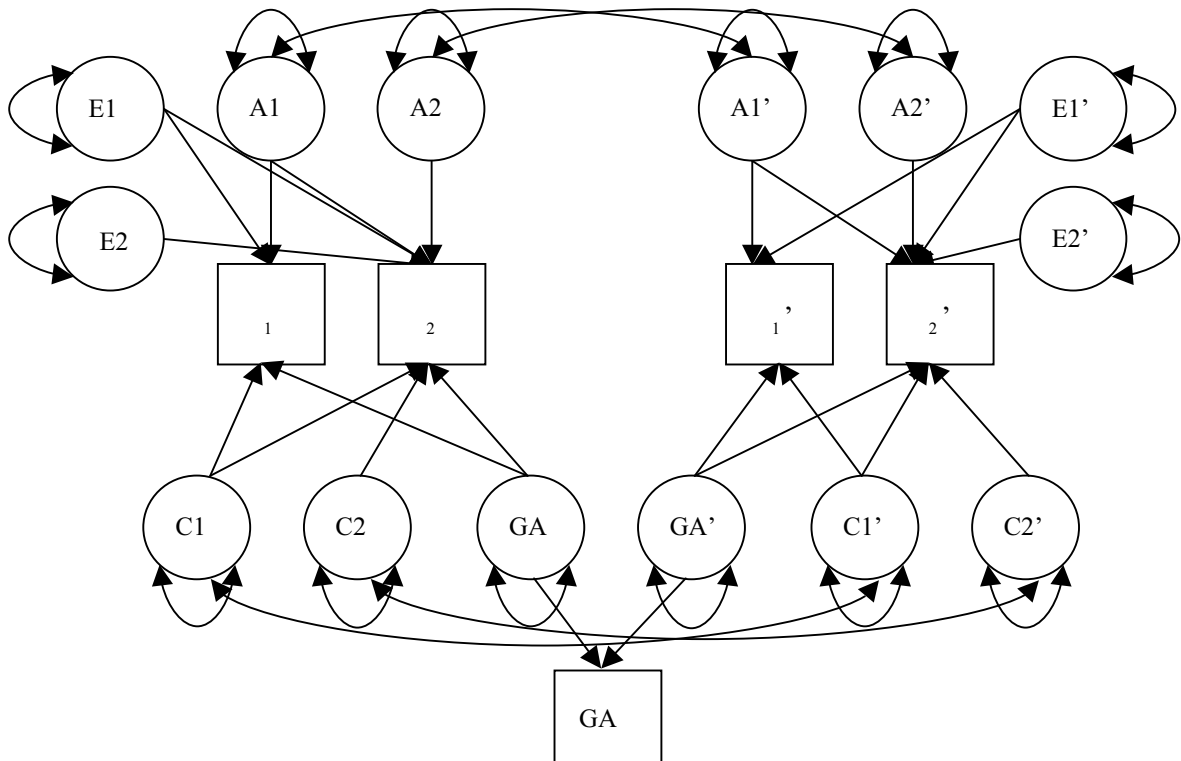
Figure 1
Path diagram of the interpolated values; size at birth (β_1 and β_1') and size at 2 years of age (β_2 and β_2').

sample covariance matrices S_i). As it may be difficult to obtain a good-fitting model with large datasets when the number of observations is large (Hair et al., 1998), the Normed Fit Index (NFI), Akaike's information criterion (AIC) and Bayesian Information Criterion (BIC) were also calculated. Values of NFI close to 1 and negative values of AIC or BIC indicate that the model under consideration provides a good fit to the data relative to the corresponding model with unrestricted covariance matrices.

Variation in growth parameters was assumed to be the sum of additive genetic variance, and common and unique environmental variances. Two series of analyses were carried out for both height and weight. In the first series, size at 1 year of age, growth velocity, deceleration of growth, rate of change in deceleration (jerk) and rate of change in jerk (snap) were simultaneously analyzed with a triangular decomposition. In the second set of analyses, size at birth and at 2 years of age was simultaneously analyzed. The estimated growth parameters (α_1 through α_5 and β_1 and β_2) were modeled as linear functions of the latent variables *additive genetic effects* (*A* male; *A'* female), *common environment* (*C* male, *C'* female), *specific environment* (*E* male, *E'* female) and the observed variable *gestational age* (*GA* male, *GA'* female; Boomsma et al., 2002; Neale et al., 1992; Neale et al., 1999).

Figure 1 depicts the path diagram for DOS-fm twin pairs for size at birth and 2 years of age. The vector (β_1, β_2) on the left contains the data for the female twin; the vector (β_1', β_2') on the right contains the corresponding values for the male twin. The vector (ga, β_1, β_2, β_1', β_2') is expressed linearly in the latent factors which are indicated in circles. The latent factors are represented by two additive genetic factors for the first twin and two additive genetic factors for the second twin. In twin pairs of opposite sex, these factors correspond with A and A'. Likewise, there are four common environmental factors and four specific environmental factors. The model also includes a gestational age factor. The female β_1 loads on A_1, C_1, E_1, GA, and the female β_2 on A_1, A_2, C_1, C_2, E_1, E_2, GA.

The structure of the covariance matrices for the six twin groups follows from the model given in Figure 1 and can be conveniently described by writing the latent vectors as matrix products involving standard normal random vectors. Let the *p*-dimensional observational vector for a given twin pair be denoted as (γ, ξ, ξ^*) representing gestational age and the growth parameters or interpolated values of the youngest and oldest of the twin pair. Then it is assumed that

$$\begin{pmatrix}\gamma\\\xi\\\xi^*\end{pmatrix}=\begin{pmatrix}0\\XA\\X^*A^*\end{pmatrix}+\begin{pmatrix}0\\YC\\Y^*C^*\end{pmatrix}+\begin{pmatrix}V\\S\\S^*\end{pmatrix}GA+\begin{pmatrix}0\\ZE\\Z^*E^*\end{pmatrix}$$

Table 1

Estimates of the Fixed Component Along With the Standard Deviation of the Random Component, the Residual Variance and Akaike's Criterion (AIC) of the Polynomial of Degree 4 Mixed-Effects Model for Different Zygosity Groups

	Height parameters (N = 2701 twin pairs)										
Zygosity	α_1		α_2		α_3		α_4		α_5		Residual variance	AIC
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
MZM	75.3	2.57	14.4	1.94	−4.99	2.58	6.52	1.85	−3.20	1.69	.869	33058
MZF	73.8	2.55	14.9	1.86	−4.50	2.78	5.60	1.69	−2.95	2.06	.799	35122
DZM & DOSm	75.3	2.57	14.2	1.99	−4.17	2.91	6.41	2.08	−3.65	2.44	.895	61092
DZF & DOSf	74.1	2.46	14.8	1.85	−4.38	2.56	5.52	1.89	−2.90	2.13	.860	57814
	Weight parameters (N = 3477 twin pairs)										
Zygosity	α_1		α_2		α_3		α_4		α_5		Residual variance	AIC
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
MZM	9.61	0.96	3.64	0.99	−1.17	1.48	1.57	0.85	−0.98	1.12	.211	11407
MZF	9.11	1.01	3.60	0.92	−1.11	1.32	1.38	0.73	−0.83	1.02	.186	9310
DZM & DOSm	9.65	0.97	3.73	1.05	−1.12	1.52	1.43	0.91	−0.95	1.20	.205	21024
DZF & DOSf	9.15	0.94	3.62	0.92	−1.13	1.33	1.31	0.74	−0.75	0.98	.193	16735

Note: α_1 = size at 1 year of age, α_2 = velocity, α_3 = deceleration, α_4 = jerk, α_5 = snap.

where X, X^*, Y, Y^*, Z and Z^* are deterministic lower-triangular ($k \times k$) matrices (where $p = 2 k + 1$ and k equals 5 for the first series of analyses and 2 for the second series), V is a number and S and S^* are deterministic k-vectors; and A, A^*, C, C^*, E, E^*, are k-dimensional standard normal random vectors and GA is a standard normal random variable. The elements of the matrices X, X^*, Y, Y^*, Z, Z^*, V, S and S^* are called *factor loadings*, and are the unknown parameters θ that are estimated from the data. The model includes the possibility that the factor loadings depend on the sex of the individual. Factor loadings are the same for all individuals of the same sex. For instance, for a monozygotic pair of twins we have $X = X^*$, where this matrix may be different for male and female monozygotic twin pairs. Furthermore, the four vectors in the decomposition on the right are assumed to be stochastically independent, all vectors E and E^* are assumed to be stochastically independent, reflecting different specific environments, and $C = C^*$, reflecting identical common environment for the two individuals in a pair of twins. Finally, it is assumed that $A = A^*$ for a monozygotic pair of twins, reflecting identical genetic make-up. The cross-covariance matrix between A and A^* is assumed to be 0.5 times the identity matrix for a dizygotic pair of twins.

The factor loadings or path coefficients which represent the influence of the latent factors on the observations are estimated, together with the unknown variance of gestational age, by maximum likelihood based on the joint distribution of the growth parameters and gestational age as indicated above. Next it is possible to compute for each of the factors the proportion of the variance that it contributes to the total variance of the observational vector. Refer to Neale and Cardon (1992) for more details on the triangular or Cholesky decomposition.

Several submodels of the general model can be formed by setting appropriate sets of factor loadings equal to zero. In the 'Null Model' EE' [GA][GA'] the loadings on both the additive genetic factors and the common environmental factors are assumed to be zero: X = X* = Y = Y* = 0. In this model, any familial resemblance in growth can only arise because there is variation between twin pairs in GA. The additive genetic factors and common environmental factors are added separately in the models AA'EE'[GA][GA'], and CC'EE'[GA][GA'] respectively, whereas AA'CC'EE'[GA][GA'] is the full model with all factors included. Finally, ACE[GA] denotes the model with all types of factors included, but with the factor loadings constrained to be identical for males and females. These submodels, versus the model in which the covariance matrix of the observational vector is an arbitrary positive-definite matrix, can be tested through the likelihood ratio test.

Results

The estimates of the growth parameters of the polynomial of degree 4 and the residual variances of the mixed-effects models for the zygosity groups are given in Table 1. The height and weight curves based on the estimated fixed parameters of the mixed model for height and weight for female and male twin pairs are shown in Figure 2.

The Pearson correlations among the growth parameters, the interpolated values and gestational age are shown in Table 2. Size at 1 year of age (α_1) correlates largely with size at 2 years of age (β_2). Also the correlation between deceleration (α_3) with snap (α_5)
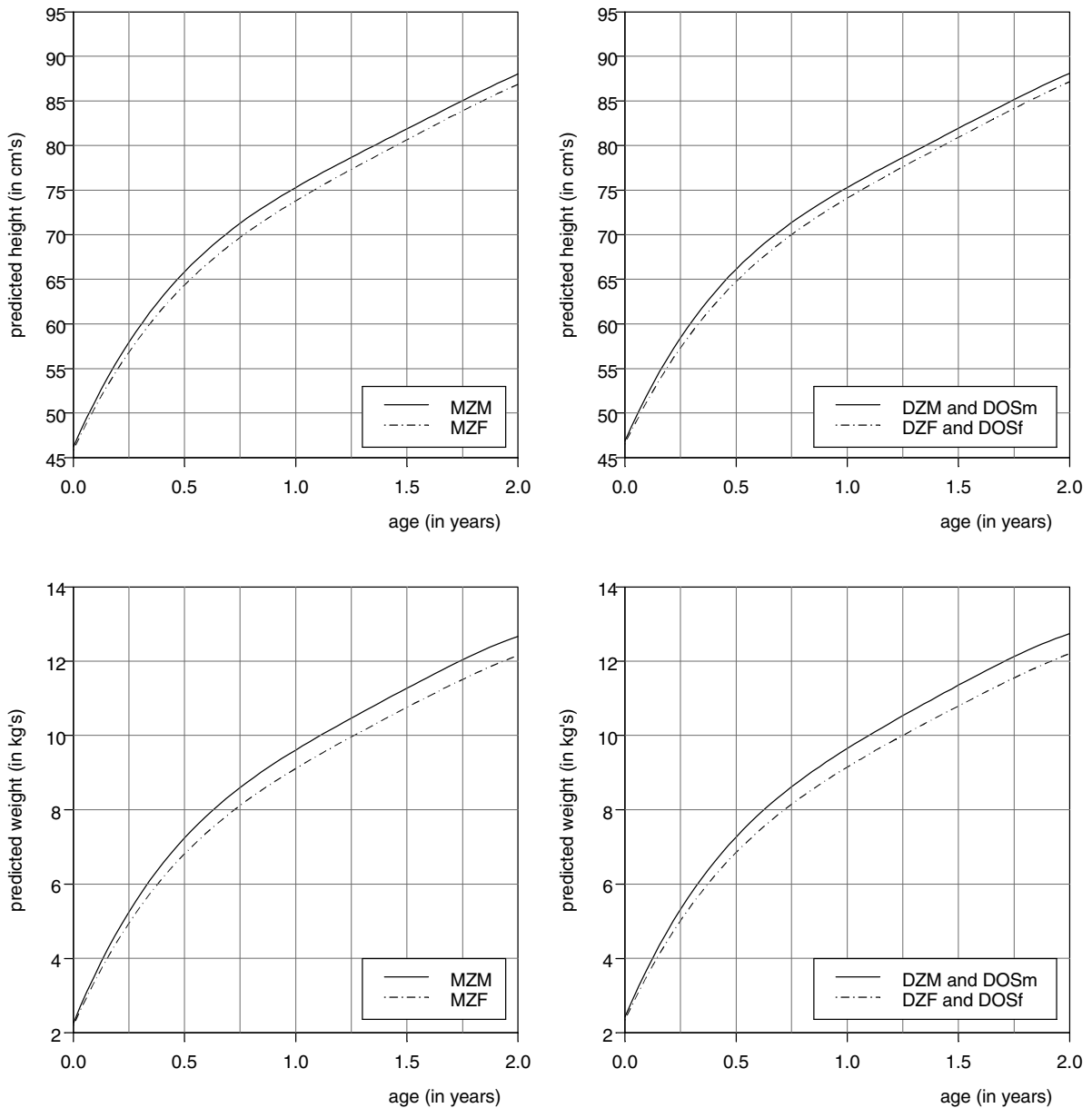
Figure 2

The height and weight growth curves for monozygotic and dizygotic boys and girls based on the estimated fixed parameters of the polynomial of degree 4 mixed-effects model.

is large. The longer the gestation period, the larger the height and weight at birth (β_1). Also, the larger the height at birth or the longer the gestation period, the more slowing over the growth rate (α_2) can be seen. Large growth velocity implies a large and heavy child at the age of 2 and deceleration rate is changing rapidly.

To show the twin resemblance for the growth parameters and the interpolated values, the within-pair correlations corrected for gestational age are shown in Table 3. It can be seen that there is marked twin resemblance for the parameters, with the MZ correlations being significantly larger than the DZ

correlations. This indicates that at least some degree of heritability exists. However, MZ correlations are not twice as high as DZ correlations, which points to an additional influence of the common environment. The differences between monozygotic and dizygotic twins are largest for size at 1 (α_1) and 2 years of age (β_2) and growth velocity of weight (α_2). This means that these parameters are more influenced by heritability than size at birth, deceleration, jerk and snap.

Goodness-of-fit tests were performed step-wise from the Null model (EE'[GA][GA']) which contains specific environmental factors and gestational age for

Table 2

Correlations Among Polynomial Parameters, Interpolated Values and Gestational Age (upper = males, lower = females)

	Height									Weight							
	α_1	α_2	α_3	α_4	α_5	β_1	β_2	GA		α_1	α_2	α_3	α_4	α_5	β_1	β_2	GA
α_1	1	.26	−.13	.17	−.01	.50	.91	.21	α_1	1	.39	−.32	.11	.11	.28	.83	.17
α_2	.29	1	−.31	−.46	.41	−.13	.50	−.28	α_2	.46	1	−.08	−.75	.21	.13	.64	−.06
α_3	−.05	−.26	1	−.02	−.71	.42	.05	.41	α_3	−.38	−.13	1	.09	−.92	.23	.09	.21
α_4	.12	−.47	−.04	1	−.65	−.41	.05	−.16	α_4	.10	−.70	.09	1	−.36	−.31	−.05	−.04
α_5	−.17	.32	−.72	−.60	1	.07	−.10	−.13	α_5	.15	.22	−.91	−.37	1	−.11	−.17	−.17
β_1	.53	−.11	.46	−.42	−.06	1	.43	.62	β_1	.30	.08	.23	−.25	−.12	1	.27	.62
β_2	.92	.52	.13	.00	−.25	.46	1	.15	β_2	.85	.69	.01	−.05	−.12	.29	1	.14
GA	.26	−.23	.40	−.16	−.17	.62	.20	1	GA	.18	−.08	.20	−.02	−.16	.64	.15	1

Note: $\alpha_1 - \alpha_5$ see Table 1, β_1 = size at birth, β_2 = size at 2 years of age.

males and females, to the final model (AA'CC'EE'[GA][GA']) which contains all factors of interest in this study. All growth models had the final model as best goodness-of-fit. The results from the Null model to the final model are summarized in Table 4. This table shows that all BIC of the final models are negative, which means that the Cholesky decomposition provides a good fit relative to the model with unrestricted covariance matrix.

The maximum likelihood estimates of the additive genetic, common environmental, gestational age and the specific environmental variance proportions under the best-fitting mixed model for females and males are given in Tables 5a and 5b, respectively. From these tables it can be seen that variation in height and weight at birth is to a large extent determined by gestational age (38% to 40% explained variance). For weight at birth, 21% to 27% of the variance is explained by common environmental factors. Size at 1 year (except for female height) and 2 years of age is mostly influenced by additive genetic factors (55% to 74%, and 52% to 59%, respectively). The largest differences between the sexes are in the height at 1 year of age. Males show a much larger proportion of variance due to additive genetic influences. The weight velocity parameter (57% to 63%) is also mainly determined by additive genetic factors. Height growth velocity, deceleration (except for female height) and rate of change in deceleration are explained by both additive genetic and common environmental factors. Deceleration of female height is mainly determined by common environmental factors (44%), specific environmental factors (28%) and partly by gestational age (16%). The correlations for the additive genetic and common environmental factors for deceleration and snap are large, indicating that these parameters are almost entirely under control of the same additive genetic and common environmental factors. Female jerk and snap, and also female height at birth and height at 2 years of age are mostly under control of the same additive genetic factor.

Discussion

The height and weight process during infancy has been described by summarizing longitudinal data into parameters of a growth model. Several growth models were tried, namely the Jenss–Bayley growth curve (1937), the first component of the Infancy–Childhood–Puberty model, polynomials with a maximum degree of 4, and the polynomial of degree 4 chosen. It had the best fit (i.e., the smallest AICs and residual variances) without being overfitted and has parame-

Table 3

Within-Pair Correlations for Parameters and Interpolated Values Corrected for Gestational Age

	Height							Weight						
Zygosity	α_1	α_2	α_3	α_4	α_5	β_1	β_2	α_1	α_2	α_3	α_4	α_5	β_1	β_2
MZM	.89	.80	.67	.74	.67	.65	.90	.87	.88	.84	.88	.84	.70	.86
DZM	.60	.60	.61	.49	.55	.55	.61	.55	.58	.68	.70	.69	.55	.53
MZF	.89	.79	.72	.69	.69	.65	.90	.84	.89	.83	.86	.84	.70	.87
DZF	.66	.57	.52	.54	.54	.53	.66	.57	.62	.66	.66	.66	.59	.55
DOSmf	.55	.57	.55	.65	.59	.63	.56	.55	.60	.65	.64	.65	.55	.54
DOSfm	.50	.55	.52	.62	.54	.56	.47	.54	.57	.56	.63	.59	.47	.55

Note: $\alpha_1 - \alpha_5$ see Table 1, $\beta_1 - \beta_1$ see Table 2.

Table 4

Goodness-of-Fit Tests of Different Cholesky Decompositions

Measure	Model	χ^2	df	p	AIC	BIC	NFI
Height	EE'[GA][GA']	9715	355	< .001	9005	6925	.00
$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$	AA'EE'[GA][GA']	2072	325	< .001	1422	−482	.81
	CC'EE'[GA][GA']	3036	325	< .001	2386	482	.71
	ACE[GA]	2691	340	< .001	2011	19	.75
	AA'CC'EE'[GA][GA']	1514	295	< .001	924	−804	.87
Weight	EE'[GA][GA']	12448	355	< .001	11738	9569	.00
$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$	AA'EE'[GA][GA']	1410	325	< .001	760	−1225	.91
	CC'EE'[GA][GA']	2251	325	< .001	1601	−384	.84
	ACE[GA]	1215	340	< .001	535	−1542	.93
	AA'CC'EE'[GA][GA']	611	295	< .001	21	−1781	.97
Height	EE'[GA][GA']	3535	79	< .001	3377	2914	.00
β_1, β_2	AA'EE'[GA][GA']	330	73	< .001	184	−244	.93
	CC'EE'[GA][GA']	707	73	< .001	561	133	.82
	ACE[GA]	184	76	< .001	32	−413	.97
	AA'CC'EE'[GA][GA']	123	67	< .001	−11	−404	.98
Weight	EE'[GA][GA']	4302	79	< .001	4144	3661	.00
β_1, β	AA'EE'[GA][GA']	333	73	< .001	187	−259	.94
	CC'EE'[GA][GA']	745	73	< .001	599	153	.84
	ACE[GA]	199	76	< .001	47	−417	.97
	AA'CC'EE'[GA][GA']	135	67	< .001	1	−408	.98

Note: Goodness-of-fit tests of different Cholesky decompositions vary from the null model in which the resemblance of the halves of a twin pair is due to gestational age effect, the model without a common environmental component, the model without an additive genetic component, the model with sex-limitation to the model with additive genetic and common environmental components without sex-limitation.

$\alpha_1 - \alpha_5$ see Table 1, $\beta_1 - \beta_2$ see Table 2.

ters which are straightforwardly interpreted in terms of growth. The second best model is the Jenss–Bayley model (1937), of which the AICs are between 13 to 451 larger for height, and between 2088 to 4504 larger for weight compared to the polynomial of degree 4.

When estimating the growth parameters by a multiple regression procedure instead of a mixed-effects model, several problems were encountered: the normality assumption for the parameters, which is needed for the biometric analyses, is violated, the fluctuation of parameters is strong because of the variability in number of observations per individual, and some individuals had to be removed from the data as the multiple regression fitting procedure cannot handle a small number of observations per individual. Therefore, it was chosen to estimate the growth parameters by a mixed-effects model.

The estimated growth parameters, together with the interpolated values for size at birth, were modelled by several multivariate genetic models. The fit of the multivariate genetic models is reasonably good, as all BIC are less than zero. The results of these models are the proportions of variance and correlation explained by additive genetic, common environmental, gestational age and specific environmental factors. The correlations indicate which parameters are under control of the same additive genetic and common environmental factors.

Statistically significant additive genetic variance was found for variation in height at birth, at 1 and at

2 years of age. In the first year, the additive genetic component for height increased from 0.10–0.15 to 0.44–0.74 and for weight from 0.14–0.24 to 0.55–0.64. Similar results were obtained by Levine et al. (1987) for American twins. Baker et al. (1992) used a subset of our dataset (n = 996 twin pairs out of 2701) and found that the additive genetic component for height varies between 0.25 and 0.45 at 1 year of age.

In common with Baker et al. (1992), it was found that the models with sex-limitation fit the data better than models constraining equality across sexes. It was also concluded that deceleration of height in females is largely determined by common environmental factors. Common environmental factors explain 21% to 27% of the variance for weight at birth. As the common effect of the mother and a more general common environmental effect cannot be separated in this design, the variance is likely to be due to maternal effects. Vandenberg et al. (1965) concluded that genetic factors appear to be of paramount importance for the deceleration of the growth rate based on a polynomial of degree 2 centred at birth. For velocity, it was found that both common environmental factors and additive genetic factors are important, while Baker et al. concluded that additive genetic factors are more important. These differences may be due to the fact that a polynomial of degree 4 instead of degree 2 was used, that the growth parameters were estimated by a mixed-effects model instead of by a multiple regression procedure, and that the data set was larger.

Table 5a

Proportions of Variance (Diagonal) and Correlations (Off-Diagonal) of the Additive Genetic Component (A), Common Environmental Component (C), Gestational Age (GA) and Specific Environmental Component (E) of the ACE Without Sex-Limitation Cholesky Decomposition for the Female Parameters of the Polynomial and the Interpolated Values

Females		A	C	GA	E
Height	α_1	.44	.39	.07	.10
	α_2	.51 .36	.36 .38	.06	.01 .20
	α_3	.42 .63 .12	.10 .02 0.44	.16	.07 .13 .28
	α_4	.17 −.59 −.62 .31	.24 −.45 0.20 0.42	.02	−.03 −.62 −.02 .25
	α_5	.07 −.54 −.77 .95 .33	.29 −.07 −0.69 0.48 0.36	.04	−.04 −.31 −.81 .56 .28
Weight	α_1	.64	.17	.03	.16
	α_2	.64 .57	.10 .31	.01	.35 .11
	α_3	.51 .15 .45	.37 .12 0.35	.04	.30 −.06 .16
	α_4	.07 −.65 .04 .39	.20 −.83 −0.23 0.48	.00	.16 −.68 −.06 .13
	α_5	−.26 −.18 −.91 .23 .37	−.15 −.34 −0.91 0.53 0.43	.03	−.09 −.02 −.92 .30 .17
Height	β_1	.10	.30	.38	.22
	β_2	.92 .52	.17 .34	.04	.43 .10
Weight	β_1	.14	.27	.40	.20
	β_2	.34 .58	.06 .25	.02	.47 .15

Note: $\alpha_1 - \alpha_5$ see Table 1, $\beta_1 - \beta_2$ see Table 2.

Table 5b

Proportions of Variance (Diagonal) and Correlations (Off-Diagonal) of the Additive Genetic Component (A), Common Environmental Component (C), Gestational Age (GA) and Specific Environmental Component (E) of the ACE Without Sex-Limitation Cholesky Decomposition for the Male Parameters of the Polynomial and the Interpolated Values.

Males		A	C	GA	E
Height	α_1	.74	.13	.04	.09
	α_2	.54 .44	.05 .32	.07	.08 .17
	α_3	.38 .33 .33	.16 .06 .32	.15	.05 .25 .20
	α_4	.30 −.49 −.19 .48	.18 −.55 .08 .30	.03	−.07 −.60 −.03 .19
	α_5	.00 −.41 −.78 .74 .60	.00 −.29 −.71 .60 .25	.01	−.09 −.49 −.73 .68 .14
Weight	α_1	.55	.31	.03	.12
	α_2	.71 .63	−.15 .26	.01	.28 .11
	α_3	.56 .26 .34	.27 −.08 .48	.05	.08 −.17 .14
	α_4	−.24 −.77 −.17 .43	.56 −.79 −.05 .47	.00	.11 −.74 −.09 .11
	α_5	−.35 −.31 −.90 .39 .32	.00 −.16 −.93 .37 .50	.03	.05 .04 −.94 .30 .15
Height	β_1	.15	.27	.39	.20
	β_2	.49 .58	.43 .29	.02	.48 .10
Weight	β_1	.24	.21	.38	.17
	β_2	.16 .59	.26 .25	.02	.47 .14

Note: $\alpha_1 - \alpha_5$ see Table 1, $\beta_1 - \beta_2$ see Table 2.

When observing the correlations obtained from the multivariate genetic models, it was concluded that deceleration and snap are almost entirely under control of the same additive genetic and common environmental factors. Female jerk and snap, and also female height at birth and height at 2 years of age, are mostly under control of the same additive genetic factor.

Literature surveys show that there are several factors that could explain part of additive genetic or common environmental effects on the growth parameters and the interpolated values (Brooke et al., 1989; Guimaraes et al., 1999; Herngreen et al., 1994). Both genetic and environmental factors are mother's educational level and family income. Smoking, alcohol, caffeine and parity may explain the common environmental effects. Further investigation into this would be of value in the future.

References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52*, 317–332.

Baker, L. A., Reynold, C., & Phelps, E. (1992). Biometrical analysis of individual growth curves. *Behavior Genetics, 22, 253–264.*

Ben-Amitai, D., Livshits, G., Levi, I., Kobyliansky, E., & Merlob, P. (1990). The relative contribution of birth weight and gestational age to physical traits of newborn infants. *Early Human Development, 22,* 131–144.

Boomsma, D. I., Busjahn, A., & Peltonen, L. (2002). The classical twin study and beyond. *Nature Genetics: Reviews, 3,* 872–882.

Boomsma, D. I., Orlebeke, J. F., & van Baal, G. C. M. (1992). The Dutch Twin Register: Growth data on weight and height. *Behavior Genetics, 22,* 247–251.

Boomsma, D. I., Vink, J. M., van Beijsterveldt, T. C., de Geus, E. J., Beem, A. L., Mulder, E. J., Derks, E. M., Riese, H., Willemsen, G. A., Bartels, M., van den Berg, M., Kupper, N. H., Polderman, T. J., Posthuma, D., Rietveld, M. J., Stubbe, J. H., Knol, L. I., Stroet, T., & van Baal, G. C. (2002). Netherlands Twin Register: A focus on longitudinal research. *Twin Research*, 5, 401–406.

Brooke, O. G., Anderson, H. R., Bland, J. M., Peacock, J. L., & Stewart, C. M. (1989). Effects on birth weight of smoking, alcohol, caffeine, socioeconomic factors, and psychosocial stress. *British Medical Journal*, 298, 795–801.

Guimaraes, L. V., Latorre, M. D., & Barros, M. B. (1999). Risk factors in the occurrence of short stature of preschool children, *Cadernos de Saude Publica, 15,* 605–615.

Hair, J. F., Anderson, R. E., Tatham, P. L., & Black, W. C. (1998). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice-Hall

Herngreen, W. P., van Buuren, S., van Wieringen, J. C., Reerink, J. D., Verloove-Vanhorick, S. P., & Ruys, J. H. (1994). Growth in length and weight from birth to 2 years of a representative sample of Netherlands children (born in 1988–89) related to socioeconomic status and other background characteristics, *Human Biology, 21,* 449–463.

Jenss, R. M., & Bayley, N. (1937). A mathematical method for studying the growth of a child. *Human Biology, 9,* 556–563.

Karlberg, J., Engström, I., Karlberg, P., & Fryer, J. G. (1987a). Analysis of linear growth using a mathematical model. I. From birth to three years. *Acta Paediatrica Scandinavica*, 76, 478–488.

Karlberg, J., Fryer, J. G., Engström, I., & Karlberg, P. (1987b). Analysis of linear growth using a mathematical model. II. From 3 to 21 years of age. *Acta Paediatrica Scandinavica*, 337(Suppl.), 12–29.

Levine, R. S., Hennekens, C. H., & Jesse, M. J. (1987). Genetic variance of weight and length in infant twins. *American Journal of Epidemiology, 126,* 929–935.

Lindstrom, M. J., & Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics, 46,* 673–687.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (1999). *Mx: Statistical modeling* (5th ed.). Richmond, VA: Department of Psychiatry, Medical College of Virginia.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for Genetic Studies of Twins and Families.* Dordrecht, the Netherlands: Kluwer Academic.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS, statistics and computing.* New York: Springer-Verlag.

Rietveld, M. J., van Der Valk, J. C., Bongers, I. L., Stroet, T. M., Slagboom, P. E., & Boomsma, D. I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Research, 3,* 134–141.

Thissen, D., & Bock, R. D. (1990). Linear and nonlinear curve fitting. In A. von Eye (Ed.), *Statistical methods in longitudinal research, Volume II: Time series and categorical longitudinal data* (pp. 289–318). Orlando, FL: Academic Press.

Vandenberg, S. G., & Falkner. F. (1965). Hereditary factors in human growth. *Human Biology*, 37, 357–365.

Van Dommelen, P., de Gunst, M. C. M., van der Vaart, A. W., van Buuren, S., & Boomsma, D. I. (2004). Growth charts for height, weight and body-mass index of twins during infancy. *Nederlands Tijdschrift Voor Geneeskunde, 148,* 1345–1350.