ANNALS OF HUMAN BIOLOGY, 1978, VOL. 5, NO. 1, 1–24

A new family of mathematical models describing the human growth curve

M. A. PREECE
Department of Growth and Development, Institute of Child Health, University of London

and M. J. BAINES
Department of Mathematics, University of Reading

[Received 1 September 1977; revised 17 October 1977]

Summary. A new family of mathematical functions to fit longitudinal growth data is described. All members derive from the differential equation $dh/dt = s(t) \, . \, (h_1 - h)$ where h_1 is adult size and $s(t)$ is a function of time. The form of $s(t)$ is given by one of many functions, all solutions of differential equations, thus generating a family of different models.

Three versions were compared. All were superior to previously described models. Model 1, in which $s(t)$ was defined by $ds/dt = (s_1 - s)(s - s_0)$ was especially accurate and robust, containing only five parameters to describe growth in stature from age two to maturity.

Derived "biological" parameters such as Peak Height Velocity were very consistent between these three members of the family but, in some cases, differed significantly from previous estimates.

1. Introduction

In any of its varied aspects, the nature of growth makes it a suitable candidate for mathematical methods of description and analysis; this is especially true of the individual growth curve. The problem to be studied here is the fitting of relatively simple mathematical models to growth curves of individual children.

There are two main reasons for this type of study:

(i) By reducing large amounts of growth data to a small number of parameters for each individual, considerable parsimony may be achieved. It is then possible to compare individual with individual and, often more importantly, population with population, using the parameters derived from fitting the equation(s). This may be helpful when data are collected at irregular intervals or when, for example, measurements on every subject are not available at every age.

(ii) Insight may be gained into functional relationships within the growing process. The parameters provide numbers that may also be used to examine relationships between the physical aspects of growth and other observations, e.g. biochemical concentrations, maturity indices, etc.

There have been many attempts at representation of the growth curve by mathematical functions. The complexity of its shape has inevitably made the approaches varied and, until recently, attention has been concentrated on parts of the curve rather than the whole (Count, 1942, 1943; Jenss and Bayley, 1937; Deming, 1957; Marubini, Resele and Barghini, 1971; Marubini, Resele, Tanner and Whitehouse, 1972; Tanner, Whitehouse, Marubini and Resele, 1976). Sometimes two or three functions have been combined to cover the whole curve.

Linear models, principally high-order polynomials (Joossens and Brems-Heynes, 1975) have been used, and also smoothing spline functions (Stützle, Gasser and Largo, 1976), but these suffer from the problem of large numbers of parameters with no clear meaning. Attempts at developing specific non-linear functions to cover all ages have recently been made by linear summation of two or three logistic functions (Bock, Wainer, Peterson, Thissen, Murray and Roche, 1973; Bock and Thissen, 1976). The former of these suffered with considerably poor fits in particular parts of the growth curve; the latter was better but contained a rather large number of parameters.

The aim of the present study was to develop a new function or family of functions that could describe the whole growth curve. We required that the fit should be better than previous models but, as far as possible, retain relatively few parameters. Five or six parameters felt to be a maximum if adequate representation could be achieved within this constraint.

2. Methods and subjects

Mathematical models

A detailed description of the mathematical derivation of the new models is given in the Appendix. They share the parent differential equation:

$$\frac{dh}{dt} = s(t) \,.\, (h_1 - h),$$

where h is height at time t, h_1 is final (or adult) height and $s(t)$ is a function of time which differs between the models.

The models take the form:

Model 1

$$h = h_1 - \frac{2(h_1 - h_\theta)}{\exp[s_0(t-\theta)] + \exp[s_1(t-\theta)]},$$

where s_0 and s_1 are rate constants, θ is a time constant and h_θ is height at $t = \theta$.

Model 2

$$h = h_1 - \frac{(h_1 - h_\theta)}{\{\frac{1}{2}\exp[\gamma s_0(t-\theta)] + \frac{1}{2}\exp[s_1'(t-\theta)]\}^{1/\gamma}}$$

where γ is a dimensionless constant and $s_1' = \gamma s_1$.

Model 3

$$h = h_1 - \frac{4(h_1 - h_\theta)}{\{\exp[p_0(t-\theta)] + \exp[p_1(t-\theta)]\}\{1 + \exp[q_1(t-\theta)]\}}$$

where p_0, p_1 and q_1 are rate constants similar to s_0 and s_1 in Model 1.

For comparison with these three related functions, the model described by Bock *et al.* (1973) was also studied. This model, referred to as *Model 4*, takes the form:

$$h = \frac{a_1}{1 + \exp\left[-b_1(t - c_1)\right]} + \frac{f - a_1}{1 + \exp\left[-b_2(t - c_2)\right]}$$

where a_1 is the upper limit of the first component, f is adult height (a fixed constant, not estimated during fitting) and b_1 and b_2 are rate constants. The parameters c_1 and c_2 are time parameters relating to the first and second components respectively.

In all, therefore, four models were fitted to the individual height/age data of 58 children.

Growth data

The measurements used to compare the models were all from the Harpenden Growth Study which has been described previously (Tanner *et al.*, 1976). From the total number of children measured, 35 boys and 23 girls were selected from those that had stopped growing in the sense that they had grown less than 1 cm during the last year of measurement, and had at least two years of measurements available before the onset of puberty. All measurements of stature were used for each child; these were six-monthly before puberty, three-monthly during puberty and yearly thereafter. Every measurement was made by Mr. R. H. Whitehouse.

Computation

The fitting of the equations to the data was done by non-linear least squares using the maximum neighbourhood algorithm due to Marquardt (1963). This was

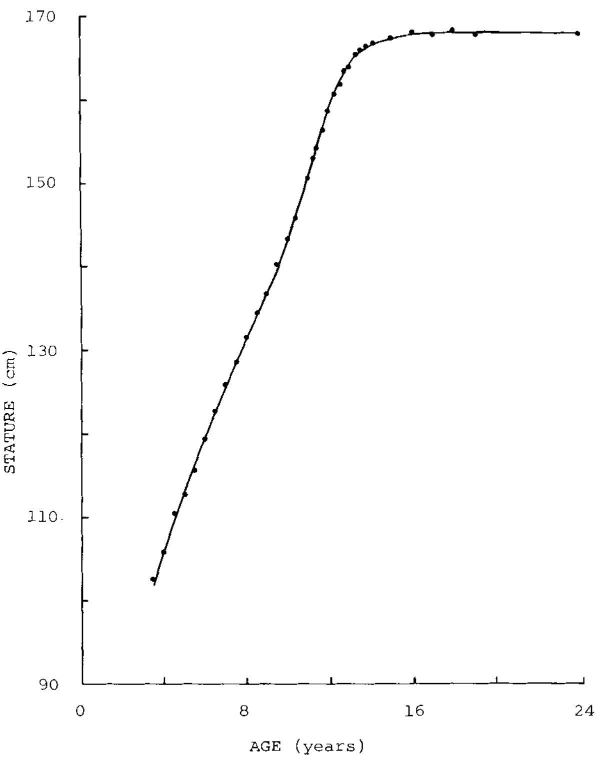


Figure 1. Model 1 fitted to the data of Girl No. 618; Residual Mean Square (RMS)=$0{\cdot}24\ cm^2$.

implemented in Fortran IV as the program NLWOOD (Daniel and Wood, 1971).

Analysis of the parameter estimates and the goodness-of-fit of the various models to the data was complicated by the non-linear nature of the models. Tests analogous to the variance ratios, etc. of linear models are not strictly valid (Draper and Smith, 1966), but "pseudo-*F*" statistics have been calculated to indicate the order of differences seen, although no probability statements have been made. The parameter estimates, themselves, have been considered as unbiased and normal theory tests have been applied as usual.

3. Results

The results are presented in three sections. Firstly, the problems experienced in the actual fitting of the four models are reported. In the second part, a detailed analysis of the validity of the models is given and, finally, in the third section the estimated parameters are studied in terms of their statistical properties and their relationship with biological parameters.

Computing

Models 1, 3 and 4 were fitted without difficulty. Convergence was achieved in an average of 7, 11 and 11 iterations for the three models respectively. There was no evidence of the existence of local minima in any case. A crude test of this was made

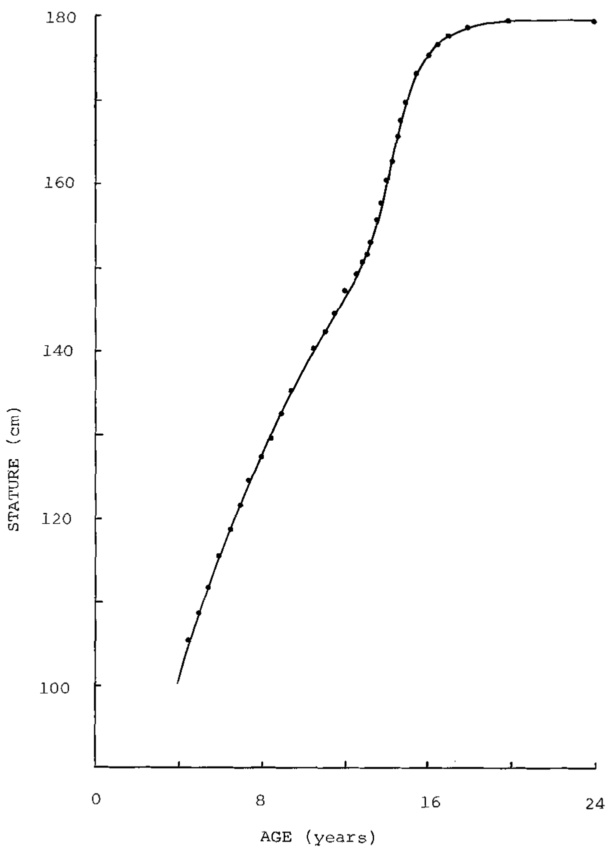


Figure 2. Model 2 fitted to the data of Boy No. 501; RMS = 0·20 cm².

by fitting each individual by each model twice, using widely differing starting values. In all cases, the program converged to identical (within rounding error) estimates of the parameters and test statistics.

In the case of Model 2, however, there were difficulties of convergence even with the reparameterization mentioned in Appendix Equation 12a. In 15 out of 35 boys and 6 out of 23 girls, no estimates of the parameters were obtained. In these individuals, convergence was very slow and yielded non-unique estimates. This latter point was indicated by different parameter estimates being obtained, depending on the starting values employed. In all the 21 individuals giving this type of result, there was a very high positive correlation between h_θ and γ and a high negative correlation between θ and h_θ. The range of correlation coefficients obtained (directly from the variance–covariance matrix for each child) was 0·990–1·000 in the former case and −0·985– −0·997 in the latter. This function was therefore extremely ill-conditioned and, compared to the other three models, very difficult to implement. A detailed exploration of the Residual Sum-of-Squares surface in relationship to the various parameters has not been undertaken and, in what follows, results given for Model 2 relate only to those obtained from the 37 satisfactorily fitted growth curves.

Quality of fit and the validity of the models

Figures 1–4 show examples of a fitted growth curve for each model. The individual curve selected was determined by seeking an individual who showed a

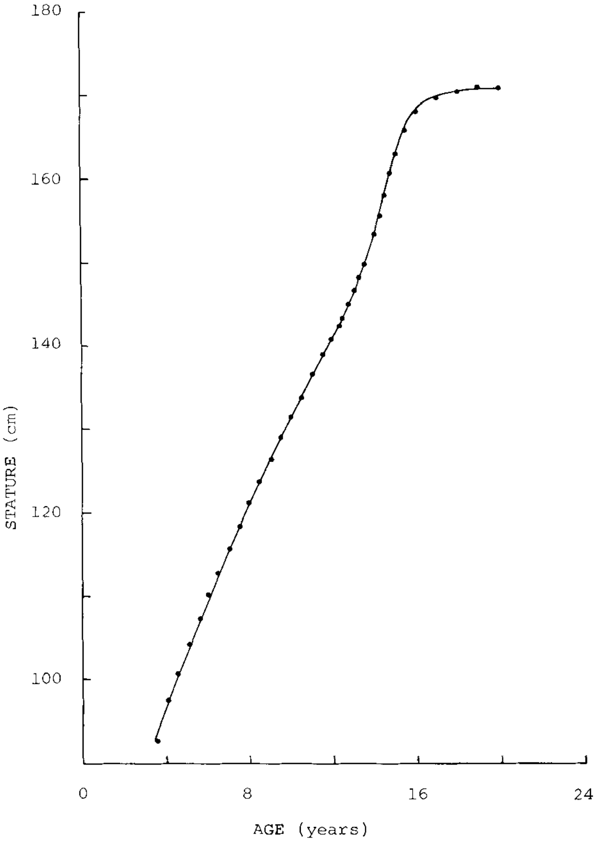


Figure 3. Model 3 fitted to the data of Boy No. 128; RMS=0·15 cm².

result "typical" of those obtained for that model. This was assessed by comparing the Residual Mean Square (RMS) of each child with the pooled within-child RMS over all observations and choosing the child nearest to that value. These curves, therefore, are neither the best nor the worst that were obtained. Clearly, all four models imitate the growth curve moderately well. Model 4, by eye, is the least satisfactory, but there is little to choose between the other three.

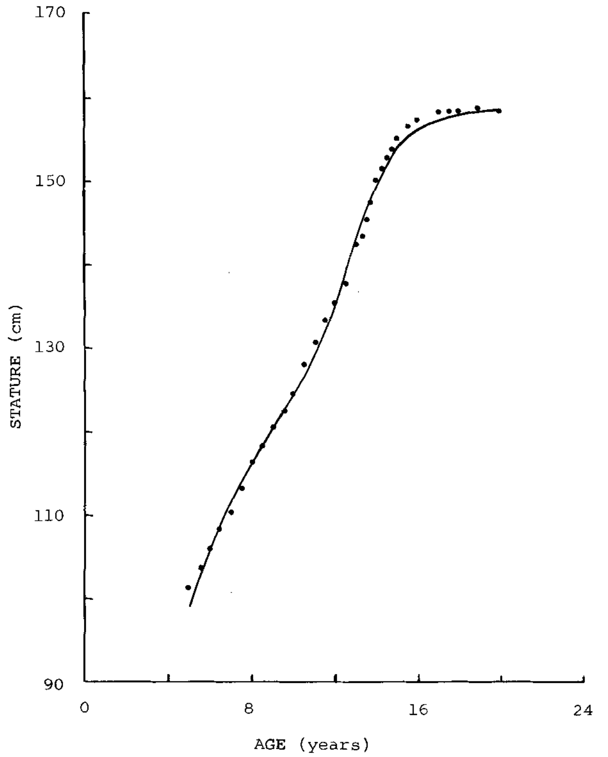


Figure 4. Model 4 fitted to the data of Girl No. 114; RMS = 1·08 cm^2.

A more satisfactory summary of the whole 58 children is shown in table 1, where the pooled within-child RMS are shown for each model and sex together with their appropriate degrees-of-freedom (d.f.).

It is immediately evident that Model 4 gives a far less satisfactory fit than any of the other three, with an RMS approximately fourfold that obtained with Models 1–3. Although Models 2 and 3 give a lower RMS than Model 1, the effect is less marked. Care must be taken in drawing conclusions concerning Model 2 in view of the likely bias introduced by the problems of fitting. The RMS values quoted in the table only represent the 20 boys and 17 girls where the fitting was successful and it is quite possible that the missing data could alter the picture considerably.

The usual requirements for inference from least squares estimates are that the residuals are normally distributed, with zero mean and constant variance. Further, they should be independent. The assumption of normality was valid; this was seen by studying the normal order plots for each child for each model. The plots of residual against fitted h were, however, very suggestive of non-independence of the residuals; there was no evidence of heteroscedasticity.

Model	RSS	d.f.	RMS	F^*
1	268·1	903	0·30	
	122·4	637	0·19	
2	95·9	472	0·20	6·6
	64·5	455	0·14	5·6
3	148·3	868	0·17	20·0
	89·9	614	0·15	9·4
4	1088·8	903	1·21	
	483·0	637	0·76	

Table 1. Within-child residual sums of squares (RSS), pooled over each model, sexes separate. Degrees-of-freedom (d.f.) and residual mean squares (RMS) are also shown. Boys in top rows.

* "F" calculated from Extra SS due to the sixth parameter in Models 2 and 3. Thus,

$$F = \frac{\text{ESS/d.f.}}{\text{RSS/d.f.}},$$ where RSS is that for the model being tested.

The ESS was calculated as, e.g. $RSS_1 - RSS_2$, giving the ESS due to the parameter of Model 2. Because of the non-linearity, the F statistics should be considered only as a guide.

In the case of Model 2, the RSS were only based on the 20 boys and 17 girls in which satisfactory parameter estimates were obtained (see text).

The distribution of the residuals was studied further in two ways. Firstly a runs test was carried out for each fit. As, in most cases, there were more than 20 observations of one or other sign, the large sample normal approximation of the runs test was used, which yields a standard normal deviate as the test statistic.

The results of this are summarized in table 2. It can be seen that for all models the means of the standard normal deviates were significantly below zero, implying that

Model	N	Mean standard normal deviate	Standard error of the mean	Number of normal deviates between −1·96 and +1·96
1	35	−2·08	0·24	11
	23	−1·25	0·25	17
2	20	−1·18	0·35	12
	17	−0·66	0·26	15
3	35	−1·04	0·20	27
	23	−0·81	0·24	19
4	35	−3·43	0·20	5
	23	−3·23	0·22	3

Table 2. The mean results of the runs tests carried out for each child, for each fitted model. A high negative value in column 3 reflects an inappropriate low number of sign changes of residuals about the fitted curve. Column 5 gives the number of individuals for whom the standard normal deviate fell between −1·96 and +1·96, therefore showing the proportion of individuals who would have had a "negative" runs test if considered separately.

Upper rows show the results for boys; lower rows represent girls.

there were too few changes of the residuals about the fitted curve. By this criterion, Models 2 and 3 were better than Model 1, but all three were significantly better than Model 4. Conclusions about Model 2 must be guarded for the reasons given above.

Since it was important to know if any particular part of the growth curve was especially badly fitted in each model, the residuals for every child were grouped according to the percentage of adult height attained. This acted as a rough estimate of maturity. The group widths were 5 per cent as indicated in tables 3–6. Each table shows the results for one model, the sexes being considered separately. In the case of Model 1, there was a tendency for under-estimation of fitted values of *h* until about 65 per cent of adult height, then a period of over-estimation until 75 per cent adult height but no clear pattern thereafter. Models 2 and 3 showed no consistent pattern but Model 4 under-estimated until 65 per cent adult height, over-estimated until 80 per cent, then repeated this pattern.

Per cent adult height	*N*	Mean residual	Standard error of mean	Standard deviation of residuals	Coefficient of skewness
Boys					
50–54·9	9	0·33	0·16	0·48	−0·30
55–59·9	21	0·41**	0·59	0·59	−0·65
60–64·9	54	0·26**	0·08	0·61	0·62*
65–69·9	81	−0·34*	0·05	0·48	0·43
70–74·9	97	−0·39**	0·04	0·36	0·44
75–79·9	122	0·07	0·04	0·44	−0·22
80–84·9	187	0·32**	0·03	0·43	−0·18
85–89·9	165	−0·21**	0·03	0·39	0·21
90–94·9	123	0·07	0·04	0·46	−0·09
95–100·0	221	−0·01	0·03	0·44	0·41*
Girls					
50–54·9	5	0·25	0·21	0·46	0·17
55–59·9	10	0·30*	0·14	0·43	0·22
60–64·9	22	0·05	0·12	0·58	−0·17
65–69·9	38	−0·18	0·09	0·58	−0·07
70–74·9	61	−0·13*	0·06	0·43	−0·83*
75–79·9	74	0·06	0·05	0·40	0·00
80–84·9	86	0·15**	0·04	0·39	0·02
85–89·9	89	−0·07	0·04	0·36	−0·33
90–94·9	102	−0·07	0·04	0·36	−0·52*
95–100·0	266	0·02	0·02	0·36	−0·12

Table 3. Grouped residual data for all children fitted with Model 1.

* Indicates significance at 5 per cent and ** at 1 per cent levels.

Coefficient of skewness, column 6, was calculated by $g_1 = M_3 M_2^{3/2}$ where M_2 and M_3 are the second and third moments of the distributions.

Tables 3–6 also show the standard deviations of the residuals (column 5) and coefficients of skewness (column 6). In Models 1–3 there was clearly no significant heterogeneity of variance. Model 4 was less regular, but still showed no genuine trend. Skewness was not a major problem; what there was was greater in Model 4.

Analysis of parameters

Model 1: Table 7 shows the mean values for each parameter together with the standard deviations. The correlation matrix for the parameters between children is

Per cent adult height	N	Mean residual	Standard error of mean	Standard deviation of residuals	Coefficient of skewness
Boys					
50–54·9	2	−0·08	0·64	0·90	—
55–59·9	10	0·37*	0·15	0·48	0·43
60–64·9	22	0·23*	0·11	0·50	1·20**
65–69·9	38	−0·18*	0·08	0·50	0·42
70–74·9	48	−0·25**	0·05	0·37	0·18
75–79·9	65	−0·03	0·05	0·40	−0·29
80–84·9	96	0·18**	0·04	0·39	−0·01
85–89·9	94	−0·09*	0·04	0·39	−0·07
90–94·9	77	0·05	0·04	0·37	−0·28
95–100·0	141	−0·01	0·03	0·31	0·18
Girls					
50–54·9	2	0·34	0·34	0·48	—
55–59·9	4	0·08	0·17	0·34	0·23
60–64·9	12	0·02	0·15	0·53	−0·23
65–69·9	28	0·07	0·10	0·51	−0·44
70–74·9	44	−0·10	0·06	0·40	−0·54
75–79·9	51	−0·09	0·05	0·36	0·11
80–84·9	65	0·15	0·04	0·33	0·08
85–89·9	61	−0·02	0·05	0·37	−0·05
90–94·9	73	−0·06	0·03	0·28	0·20
95–100·0	217	0·01	0·02	0·28	0·13

Table 4. Grouped residual data for the children successfully fitted by Model 2. Conventions as for table 3.

Per cent adult height	N	Mean residual	Standard error of mean	Standard deviation of residuals	Coefficient of skewness
Boys					
50–54·9	7	−0·44*	0·20	0·54	0·85
55–59·9	22	0·16	0·10	0·45	0·03
60–64·9	55	0·15**	0·05	0·34	0·54
65–69·9	81	−0·10**	0·03	0·31	0·00
70–74·9	97	−0·10**	0·03	0·30	0·33
75–79·9	115	0·05	0·03	0·34	−0·39
80–84 9	180	0·08*	0·03	0·34	−0·34
85–89·9	177	−0·14**	0·03	0·33	0·00
90–94·9	123	0·19**	0·03	0·36	0·22
95–100·0	222	−0·04	0·03	0·41	−0·04
Girls					
50–54·9	4	−0·19	0·24	0·47	0·45
55–59·9	10	0·22	0·12	0·37	−0·14
60–64·9	22	0·02	0·17	0·41	0·29
65–69·9	39	−0·08	0·07	0·43	−0·04
70–74·9	62	−0·02	0·05	0·37	0·67*
75–79·9	70	0·02	0·04	0·32	−0·20
80–84·9	84	0·04	0·04	0·33	−0·23
85–89·9	92	−0·04	0·04	0·35	−0·30
90–94·9	100	0·00	0·03	0·32	−0·22
95–100·0	269	0·00	0·02	0·34	−0·07

Table 5. Grouped residual data for all children fitted with Model 3. Conventions as for table 3.

Per cent adult height	N	Mean residual	Standard error of mean	Standard deviation of residuals	Coefficient of skewness
Boys					
50–54·9	9	1·13**	0·15	0·44	−0·37
55–59·9	20	1·16**	0·16	0·70	−0·84
60–64·9	52	0·54**	0·14	1·00	0·51
65–69·9	68	−0·57**	0·11	0·89	0·91**
70–74·9	96	−0·98**	0·05	0·53	0·70*
75–79·9	132	−0·06	0·07	0·80	0·18
80–84·9	179	0·73**	0·05	0·61	−0·37
85–89·9	157	−0·24**	0·05	0·67	0·30
90–94·9	141	−0·43**	0·06	0·70	−0·14
95–100·0	225	0·97**	0·06	0·83	0·71**
Girls					
50–54·9	3	1·19**	0·15	0·21	0·00
55–59·9	9	1·02**	0·18	0·53	−0·06
60–64·9	22	0·23	0·19	0·90	0·07
65–69·9	38	−0·30	0·17	1·03	0·17
70–74·9	59	−0·55**	0·09	0·67	0·12
75–79·9	75	0·03	0·08	0·73	0·38
80–84·9	89	0·57**	0·05	0·48	0·11
85–89·9	78	−0·16*	0·07	0·63	−0·81**
90–94·9	105	−0·52**	0·05	0·51	−0·61*
95–100·0	274	0·53**	0·04	0·63	1·07**

Table 6. Grouped individual data for all children fitted with Model 4. Conventions as for table 3.

	Boys (N=35)		Girls (N=23)		Boys less Girls	
	Mean	SD	Mean	SD	Difference	SE
h_1	174·6	6·0	163·4	5·1	10·2**	1·5
h_θ	162·9	5·6	152·7	5·2	10·2**	1·4
s_0	0·1124	0·0126	0·1320	0·0181	−0·0196**	0·0040
s_1	1·2397	0·1683	1·1785	0·1553	0·0612	0·0438
θ	14·60	0·93	12·49	0·74	2·11**	0·23

	h_1	h_θ	s_θ	s_1	θ	
h_1		0·97	0·08	−0·09	0·07	r (α=0·01)
h_θ	0·97		0·29	0·04	0·06	Boys, 0·43
s_0	−0·17	−0·07		0·60	−0·33	Girls, 0·53
s_1	−0·25	−0·20	0·66		−0·27	
θ	0·11	0·26	−0·33	−0·34		

Table 7. Mean values for parameters of Model 1. In the correlation matrix the upper right triangle contains the correlation coefficients for the girls, the lower left triangle those for the boys.

* and ** indicate significant levels of 5 per cent and 1 per cent respectively.

also shown. The sexes were treated separately. Note that h_1, h_θ and θ were significantly greater in boys whereas s_0 was greater in girls; s_1 did not differ between the sexes. The most striking correlations between parameters were h_1 and h_θ (0·97 both sexes) and s_0 and s_1 (0·60 and 0·66).

Model 2: Comparable data to those above are given for Model 2 in table 8. Only the 20 boys and 17 girls in whom stable fits were achieved have been considered.

	Boys (N=20)		Girls (N=17)		Boys less Girls	
	Mean	SD	Mean	SD	Difference	SE
h_1	175·8	6·6	163·7	4·1	12·1**	1·8
h_θ	159·5	7·4	151·4	6·7	8·1**	2·3
s_0	0·1210	0·0198	0·1379	0·0194	−0·0169*	0·0065
s_1'	1·8780	0·7615	1·4879	0·4888	0·3901	0·2148
θ	13·94	1·19	12·13	1·11	1·81**	0·38
γ	2·2145	1·5062	1·4843	0·9243	0·7302	0·4201

	h_1	h_θ	s_0	s_1'	θ	γ	
h_1		0·72	0·24	−0·08	0·13	−0·11	
h_θ	0·53		0·09	−0·67	0·62	−0·72	r (α=0·01)
s_0	−0·36	−0·33		0·36	−0·41	0·24	Boys, 0·56
s_1'	−0·09	−0·65	0·70		−0·73	0·96	Girls, 0·61
θ	−0·09	0·63	−0·30	−0·60		−0·74	
γ	0·23	−0·60	0·29	0·84	−0·64		

Table 8. Mean values for the parameters of Model 2, for the 37 children successfully fitted.
Conventions as for table 7.

	Boys (N=35)		Girls (N=23)		Boys less Girls	
	Mean	SD	Mean	SD	Difference	SE
h_1	174·0	5·8	163·2	4·9	10·8**	1·5
h_θ	164·0	5·7	153·8	5·2	10·2**	1·5
p_0	0·0880	0·0257	0·1103	0·0260	−0·0223**	0·0069
p_1	0·2245	0·0795	0·2351	0·1122	−0·0106	0·0345
q_1	1·3676	0·1743	1·1513	0·1794	0·2163**	0·0473
θ	14·75	0·98	12·66	0·86	2·09**	0·31

	h_1	h_θ	p_0	p_1	q_1	θ	
h_1		0·94	0·21	−0·09	0·02	−0·01	
h_θ	0·95		0·11	0·23	0·21	0·19	r (α=0·01)
p_0	−0·05	−0·05		−0·55	−0·20	−0·55	Boys, 0·43
p_1	−0·05	0·17	−0·70		0·58	0·63	Girls, 0·53
q_1	−0·24	−0·10	−0·31	0·41		0·33	
θ	0·10	0·32	−0·32	0·35	0·24		

Table 9. Mean values for parameters of Model 3.
For description of correlation matrix, see legend to table 7.
Conventions as for table 7.

Much the same pattern as for Model 1 was seen with h_1, h_θ and θ being greater in boys but s_θ being greater in girls. Once again, s_1' was not different between sexes. The value of γ was not significantly different between boys and girls, but in both sexes it was significantly greater than 1.

The correlation matrix showed much larger correlations between the parameters in this sub-sample, with the greatest dependence being between s_1' and γ in both sexes.

Model 3: The data for this model are shown in table 9. The pattern was again very similar to Model 1 with the same type of differences between the sexes. Once again, the correlations between parameters showed the same pattern as for Model 1, with a return to more orthogonality than in Model 2.

In eight boys and nine girls there was no improvement in fit compared to Model 1, and in these individuals substitution of the parameters into Model 3 and simplification led to the same parameter estimates as were obtained in Model 1.

Model 4: The data from the last model are shown in table 10. In the case of the girls, one individual has been left out of the calculations as, in her case, the parameter estimates were extremely discordant with those of the other girls and the variance was spuriously elevated. There was no very clear pattern to the parameters; f, a_1 and c_2 were greater in boys, b_1 was greater in girls and c_1 did not differ between sexes.

Overall, the numbers were too small to draw strong conclusions about the distributions of the parameters in any of the models. From normal plots, however, it could be seen that there was no strong reason to reject a normal distribution of the

	Boys (N=35)		Girls (N=22)†		Boys less Girls	
	Mean	SD	Mean	SD	Difference	SE
f	174·1	6·0	163·3	4·9 (163·2)	10·8**	1·5
a_1	148·1	5·9	136·6	6·1 (156·8)	11·5**	1·6
b_1	0·3089	0·0619	0·3860	0·0705 (0·3931)	−0·0771**	0·0175
c_1	2·1435	1·1010	1·8175	0·9434 (1·9304)	0·3260	0·2335
b_2	1·0712	0·1871	1·0125	0·1379 (0·9980)	0·0587	0·0456
c_2	13·7324	0·8373	11·6542	0·6171 (11·3567)	2·0782**	0·2042

	f	a_1	b_1	c_1	b_2	c_2	
f		0·88	−0·18	−0·20	0·03	−0·16	
a_1	0·80		−0·05	0·10	0·31	0·01	r (α=0·01)
b_1	0·06	−0·10		0·79	0·38	−0·23	Boys, 0·43
c_1	0·11	0·13	0·87		0·51	0·28	Girls, 0·54
b_2	−0·16	0·04	0·53	0·60		−0·08	
c_2	0·14	0·45	−0·04	−0·04	−0·30		

Table 10. Mean values for parameters of Model 4.

Conventions as for table 7.

† One observation left out (F121) as parameter estimates, especially a_1, were very non-typical. Means calculated with all 23 values are shown in brackets. The correlation matrix is composed of estimates based on 22 values for girls.

population parameters except for γ in Model 2, p_0 in Model 3 and b_1 in Model 4 where log-normal distributions seemed more suitable. In further analysis, however, conventional normal theory statistics and tests were employed without any transformation.

As the parameters of the four models were estimated separately for each model, and therefore had no *a priori* relationships between models, despite sharing the same nomenclature in some cases, immediate comparison between models was relatively meaningless. A more useful approach was to convert the parameters into a selection of biological parameters (which are effectively a series of linear combinations of the model parameters) and then compare these estimates.

Tables 11 and 12 list the biological parameters calculated and give the mean, standard deviation and extreme values for each parameter, as calculated via each of the four models. One-way analysis of variance was used to compare the means of each parameter from each model. Only five showed significant differences; these were velocity at Take-Off (TO); height at Peak Height Velocity (PHV); height increase TO to PHV and percentage of adult height at PHV (these in both sexes) and velocity increase TO to PHV in girls only. By inspection, it was clear that in all cases, the extreme values were obtained from Model 4 and therefore a series of linear contrasts were set up to specifically show this. In this way estimates from Model 4

	Model 1	Model 2	Model 3	Model 4
Adult height (cm)	174·6 (6·0) 164·0 –184·5	175·8 (6·6) 164·2 –190·1	174·0 (5·8) 163·7 –182·8	174·1 (5·0) 163·6 –184·0
Age at take-off (years)	10·71–(0·85) 8·63– 12·43	10·85 (0·85) 8·75– 12·41	11·15 (1·05) 8·61– 13·44	10·46 (0·77) 8–67– 11·93
Height at take-off (cm)	138·9 (5·9) 126·5– 148·3	139·7 (6·2) 127·2 –148·6	141·0 (6·5) 126·4 –152·9	137·7 (5·8) 126·7 –146·6
Velocity at take-off (cm/yr)	4·51 (0·60) 3·40– 5·94	4·50 (0·52) 3·57– 5·59	4·66 (0·58) 3·66– 5·94	3·98 (0·52) 3·05– 5·01
Age at peak height velocity (years)	14·17 (0·91) 11·90– 16·24	14·23 (1·01) 11·85– 15·52	14·36 (0·99) 11·94– 16·70	13·62 (0·82) 11·60– 15·32
Height at peak height velocity (cm)	159·5 (5·5) 149·4 –168·4	159·7 (5·6) 149·2 –166·6	160·8 (5·6) 149·5 –170·7	155·8 (5·5) 150·7 –163·8
Peak height velocity (cm/yr)	8·23 (1·17) 5·63– 10·02	8·38 (1·40) 4·90– 10·69	8·72 (1·03) 5·63– 10·25	8·22 (1·33) 5·52– 10·42
Height increase, take-off to PHV (cm)	20·5 (2·6) 16·3 – 26·0	20·1 (4·0) 11·5 – 24·1	19·7 (2·9) 15·4 – 25·4	18·1 (2·1) 13·9 – 22·4
Velocity increase, take-off to PHV (cm/yr)	3·71 (0·93) 0·86– 5·54	3·88 (1·17) 0·39– 6·16	4·05 (0·85) 0·75– 5·54	4·24 (1·13) 1·38– 5·97
Percentage adult height at take-off	79·6 (2·3) 75·3 – 83·8	79·9 (2·5) 74·7– 85·0	81·1 (2·6) 75·3– 85·4	79·1 (2·1) 75·3 – 83·8
Percentage adult height at PHV	91·3 (1·0) 89·4 – 93·2	91·4 (1·8) 84·5 – 92·8	92·4 (1·2) 89·4 – 94·2	89·5 (1·0) 88·1 – 91·9
N	35	20	35	35

Table 11. Biological parameters for the boys; calculated from the function parameters of each model. Values given are the mean, with the standard deviation of the distribution in brackets. Under each mean the range of values is given. As before, only the children in whom successful fits were achieved are reported under Model 2.

	Model 1	Model 2	Model 3	Model 4
Adult height (cm)	163·4 (5·1) 153·6 –172·8	163·6 (4·1) 154·8 –168·7	163·2 (4·9) 153·5 –171·3	163·5 (4·8) 152·9 –169·1
Age at take-off (years)	8·96 (0·70) 7·7 – 10·01	8·93 (0·62) 7·89– 9·97	9·05 (0·82) 8·05– 10·92	8·66 (0·89) 7·63– 9·60
Height at take-off (cm)	129·9 (6·3) 118·2 –142·1	130·2 (6·3) 120·3 –141·9	130·9 (6·7) 118·6 –144·8	127·9 (6·2) 115·8 –139·9
Velocity at take-off (cm/yr)	5·19 (0·43) 4·33– 6·24	5·26 (0·49) 4·50– 6·46	5·29 (0·39) 4·66– 6·24	4·58 (0·43) 3·74– 5·66
Age at peak height velocity (years)	11·87 (0·74) 10·31– 13·21	11·90 (0·82) 10·32– 13·44	12·01 (0·85) 10·35– 13·75	11·36 (0·94) 10·51– 12·69
Height at peak height velocity (cm)	148·3 (5·1) 137·0 –157·1	148·4 (5·1) 139·3 –157·4	149·2 (5·2) 137·4 –155·4	145·0 (5·5) 133·8 –154·2
Peak height velocity (cm/yr)	7·47 (0·76) 6·12– 9·27	7·56 (0·95) 6·09– 9·52	7·50 (0·76) 6·12– 9·27	7·91 (0·75) 6·60– 9·59
Height increase, take-off to PHV (cm)	18·2 (2·6) 13·9 – 22·0	18·2 (3·4) 13·6 – 24·8	18·2 (2·8) 13·9 – 24·0	17·1 (2·0) 13·3 – 20·1
Velocity increase, take-off to PHV (cm/yr)	2·28 (0·86) 0·94– 3·86	2·38 (1·08) 0·90– 4·14	2·21 (0·86) 0·73– 3·86	3·33 (0·85) 1·84– 5·01
Percentage adult height at take-off	79·4 (2·4) 76·3 – 84·2	79·8 (2·4) 75·8 – 84·1	80·2 (2·5) 76·3 – 85·9	78·4 (2·5) 75·2 – 82·7
Percentage adult height at PHV	90·7 (1·0) 89·2 – 92·3	90·9 (1·0) 88·9 – 93·3	91·4 (1·2) 90·2 – 94·2	88·8 (1·7) 87·4 – 91·2
N	23	17	23	22

Table 12. Biological parameters for the girls; calculated from the function parameters of each model. Values given are the mean, with the standard deviation of the distribution in brackets. Under each mean the range of values is given. As before, only the children in whom successful fits were achieved are reported under Model 2 and Subject 121 is excluded from the values under Model 4.

were compared to a weighted mean of the estimates from Models 1–3. These took the form

$$\text{Contrast} = \sum \lambda_i x_i$$

with $\lambda_1 = \lambda_2 = \lambda_3 = +1$ and $\lambda_4 = -3$.

The standard error (SE) of the contrast was given by

$$\text{SE} = \sqrt{(s^2 \sum \lambda_i^2/n_i)},$$

where s^2 was the pooled estimate of the within-model variance for any particular parameter. Table 13 gives the results of these contrasts. Clearly, Model 4 gave consistently lower estimates for velocity at TO; height at PHV; height increase, TO to PHV; and percentage of adult height at PHV, but higher estimates of the velocity increase, TO to PHV.

Apart from these differences, the estimates were remarkably consistent from all four models.

One parameter in particular has generally been poorly represented by these types of models; this is the magnitude of PHV (Tanner *et al.*, 1976). Graphical estimates of age at PHV and PHV itself were available on 18 boys and 9 girls of the overall sample (kindly provided by Mr. R. H. Whitehouse). The mean values of

	Contrast	S.E. Contrast	*t*
Boys			
Velocity at take-off (cm/yr)	1·73	0·38	4·48
Height at peak height velocity (cm)	12·6	3·29	3·83
Height increase, take-off to PHV (cm)	6·0	1·61	3·73
Velocity increase, take-off to PHV (cm/yr)	−1·08	0·62	−1·75
Percentage adult height at PHV	6·6	0·73	9·03
Girls			
Velocity at take-off (cm/yr)	2·00	0·19	10·50
Height at peak height velocity (cm)	10·90	2·28	4·78
Height increase, take-off to PHV (cm)	3·3	1·18	2·80
Velocity increases, take-off to PHV (cm/yr)	−3·12	0·45	−6·93
Percentage adult height at PHV	6·6	0·55	12·00

Table 13. Comparison of Model 4 estimates with those from Models 1, 2 and 3 for the five "biological" parameters showing evidence of different means between models by one-way analysis of variance. See text for details.

these estimates are shown in table 14, together with the means of the same individuals obtained from the four models. Concordance was very good for *age* at PHV but, once again, the estimates of PHV from the models tended to be rather lower than the graphical estimates. The difference was significant at $P<0{\cdot}001$ by paired *t*-test for both sexes when the highest model mean was compared to the graphical mean.

Finally, the relationship between the model parameters and the biological parameters was investigated. The simple approach was to study the correlations between the two types of parameter, within each model. Interpretation had to be cautious, however, as many of the biological parameters were linear combinations of each other. If large numbers of such dependent variables were considered, there would have been linear dependence in the correlation matrix. This potentially leads to considerable errors and, for this reason, a limited number only were considered.

	Graphical	Model 1	Model 2	Model 3	Model 4
Boys					
At at PHV (years)	13·83 (0·87)	13·77 (0·32)	13·63 (0·87)	13·92 (0·89)	13·32 (0·75)
PHV (cm/yr)	9·62 (1·11)	8·69 (1·21)	9·11 (1·59)	8·93 (1·11)	8·62 (1·35)
Girls					
Age at PHV (years)	11·82 (0·82)	11·80 (0·77)	11·83 (0·85)	11·97 (0·90)	11·34 (0·95)
PHV (cm/yr)	8·32 (0·85)	7·58 (0·84)	7·62 (0·98)	7·61 (0·80)	7·92 (0·81)

Table 14. Comparison of means of graphical estimates of PHV and Age at PHV with estimates from the four models. Standard deviations are shown in brackets.

Tables 15–18 show such correlations for each model. It can be seen that in Model 1, h_1 and h_θ were principally associated with the height at PHV and, to a lesser extent, height at TO, whereas s_0 was mainly associated with growth velocity at TO. The other rate constant, s_1, was related to the velocity at PHV; θ, although highly correlated with several parameters, was principally associated with age at PHV.

In Model 2, the picture was more obscure, possibly because of the smaller numbers of individuals, and there was less concordance between the sexes. The only really firm associations were those of h_1 and h_θ, as in Model 1. The same was also true in Model 3 but p_0 and p_1 were not very definitely associated with any biological parameter, especially in boys. The third rate constant, q_1, was correlated with ages of TO and PHV as was θ.

There was, therefore, no very clear picture to be drawn from these three related models, although it seemed that in Models 1 and 2, s_0 may reflect behaviour of the growth curve near TO and s_1 (and s_1') that near PHV. The parameters h_1 and h_θ were clearly the main determinants of absolute size, as would be expected.

In Model 4, f and a_1 were associated with the two height parameters (at TO and PHV), b_1 essentially represented velocity at TO and b_2 that at PHV. There was no ready interpretation of c_1, but c_2 was identical with age at PHV.

	Age at TO	Ht at TO	Vel at TO	Age at PHV	Ht at PHV	Vel at PHV
h_1	0·01	0·73	0·25	0·08	0·95	0·09
	0·09	0·78	0·26	0·01	0·95	−0·02
h_θ	0·21	0·87	0·11	0·23	0·99	−0·06
	0·20	0·91	0·29	0·01	0·99	−0·12
s_0	0·04	0·12	0·44	−0·30	−0·06	0·25
	0·24	0·60	0·55	−0·34	0·30	−0·14
s_1	0·08	0·01	0·09	−0·20	−0·10	0·60
	0·21	0·29	−0·01	−0·10	0·18	0·53
θ	0·90	0·51	−0·78	0·99	0·30	−0·78
	0·82	0·04	−0·56	0·97	0·06	−0·38

Table 15. Correlations between Model 1 parameters and a selection of "biological" parameters.

The upper values represent boys and the lower girls.

r ($\alpha=0·01$) for the boys and girls was 0·43 and 0·53 respectively.

	Age at TO	Ht at TO	Vel at TO	Age at PHV	Ht at PHV	Vel at PHV
h_1	0·05	0·71	0·27	−0·03	0·83	0·13
	0·11	0·81	−0·09	−0·04	0·95	0·05
h_θ	0·17	0·91	−0·02	0·18	0·93	0·14
	0·12	0·81	0·17	0·19	0·86	−0·02
s_0	0·23	0·24	0·35	−0·39	−0·20	0·46
	0·17	0·26	0·63	0·31	−0·48	0·21
s_1	0·22	0·24	0·21	−0·51	−0·27	0·66
	0·03	0·01	0·23	−0·06	−0·54	−0·13
θ	0·29	0·04	−0·43	0·83	0·32	−0·87
	0·17	0·21	−0·23	−0·11	0·72	0·14
γ	0·14	0·25	0·18	−0·54	−0·20	0·57
	0·01	−0·07	0·17	−0·01	−0·57	−0·19

Table 16. Correlations between Model 2 parameters and a selection of "biological" parameters.

Conventions as for table 15.

r ($\alpha=0·01$) for the boys and girls was 0·56 and 0·61 respectively.

	Age at TO	Ht at TO	Vel at TO	Age at PHV	Ht at PHV	Vel at PHV
h_1	0·05	0·71	0·22	0·08	0·93	0·08
	0·16	0·82	0·22	−0·02	0·94	0·06
h_θ	0·28	0·86	0·07	0·29	0·99	−0·13
	0·40	0·92	0·22	0·16	0·98	−0·13
P_0	−0·20	−0·09	0·18	−0·31	−0·16	0·11
	−0·23	0·28	0·46	−0·55	0·12	0·04
P_1	0·26	0·17	0·04	0·29	0·14	−0·25
	0·61	0·21	0·03	0·57	0·18	−0·29
q_1	0·44	0·16	−0·31	0·34	0·01	0·25
	0·56	0·30	−0·17	0·45	0·31	0·35
θ	0·94	0·60	−0·73	0·99	0·37	−0·71
	0·85	0·13	−0·41	0·98	0·18	−0·34

Table 17. Correlations between Model 3 parameters and a selection of "biological" parameters.

Conventions as for table 15.

r (α=0·01) for the boys and girls was 0·43 and 0·53 respectively.

	Age at TO	Ht at TO	Vel at TO	Age at PHV	Ht at PHV	Vel at PHV
f	0·11	0·78	0·27	0·14	0·95	0·14
	−0·10	0·79	0·30	−0·17	0·93	0·06
a_1	0·52	0·95	−0·05	0·45	0·91	−0·15
	0·23	0·94	0·27	0·00	0·96	−0·05
b_1	−0·04	0·13	0·71	−0·40	0·13	0·63
	0·23	0·26	0·41	−0·22	0·05	0·14
c_1	0·38	0·33	0·54	−0·01	0·25	0·49
	0·73	0·33	0·31	0·30	0·09	0·01
b_2	0·16	0·12	0·18	−0·25	0·01	0·68
	0·33	0·37	−0·02	−0·02	0·27	0·63
c_2	0·87	0·43	−0·62	1·00	0·29	−0·65
	0·83	−0·01	−0·46	1·00	−0·08	−0·43

Table 18. Correlations between Model 4 parameters and a selection of "biological" parameters.

Conventions as for Table 15.

r (α=0·01) for the boys and girls was 0·43 and 0·54 respectively.

4. Discussion

It has been shown that it is possible to develop a family of mathematical functions that will describe the human growth curve. The method of derivation has been purely empirical and has made no pretence to true biological meaning. Nevertheless, the parameters do relate, in some instances closely, to empirically determined biological ones. The whole family (Models 1–3) generally simulated the shape of the individual growth curve better than anything else available at the present time, and certainly better than the only other model explicitly studied, that of Bock *et al.* (1973).

Model 2, in its present parameterization was really too ill-conditioned for general use. Even with the re-parameterization mentioned in the Appendix, the RSS surface was clearly irregular with many local minima and probably an extremely attenuated "trough" leading to the true minimum. This latter point was suggested by the slow convergence in some individuals. It is possible that manipulation of the minimization routine could have circumvented the problems. Alternatively, use of a different algorithm, such as the simplex method of Nelder and Mead (1965), might have improved the situation. However, the purpose of this project was to develop robust models that were readily fitted, rather than *prima donnas* that required great and individual care. For this reason we have not pursued the matter further.

From the results of the successful fits with Model 2, it would appear that the *a priori* assumption of $\gamma=1$ in Model 1 was invalid. However, this could be looked at from another angle; the two models may be considered as quite distinct, one with five parameters and the other with six. The merits of each can be assessed by their performance when applied to data and from this standpoint, the merits of Model 1 are obvious and Model 3, with one extra parameter, did at least as well.

Model 1 was remarkably well conditioned, converging rapidly in all cases. No parameter was ever less than four times its standard error and, in most cases, (especially h_1, h_θ and θ) the estimates were 20 or more times their standard errors. Model 3 was similar except when p_0 and p_1 were equal and the model reduced to Model 1.

Model 4 was generally well behaved but was slow to converge if the data were scanty near adult height. In only six cases did it give a better fit than the Model 1, as judged by the RMS.

In all four models, some degree of auto-correlation of the residuals was found. It was worst in Model 4 and least in Model 3. The runs test was not without its drawbacks in this situation, as it did not take account of the size of the residuals—in most cases in this study they were very small and often within the measuring error of the original data. Some clarification was sought utilizing the Durbin–Watson statistic (Durbin and Watson, 1971) but the results were such as to alter the picture only marginally, suggesting that correlation among the residuals was slightly less frequent than the simple runs test suggested.

The very nature of this type of data will tend to encourage auto-correlation. By definition, $h_t \geq h_{t-1}$ and if this constraint was not met at the time the data was collected, then the measurement would have been checked. On the other hand, a rather high measurement may have gone unchecked and this differential could have led to some bias in the data. To this should be added the generally observed cyclical fluctuations in growth rate, some seasonal and some not, which no model could hope to imitate unless there were a very large number of parameters. For these reasons, it was thought more useful to study whether there was any gross pattern to the residuals, such that some particular part of the curve was under or over-estimated. In general, Model 2 and 3 did not show any particular trend and Model 1 tended to under-estimate in the early part of the curve and then over-estimate until 75 per cent of adult height. Model 4 was worst, both in terms of regular trends and the absolute magnitude of the residuals.

On balance, if a degree of auto-correlation could be accepted, then all the first three models gave acceptable representations of the growth curve. The main problem of the auto-correlation is the effect upon variance estimates of the parameters and the subsequent statistical tests. Watson (1955) indicated that ignored correlations could lead to deceptively small variance estimates. For this reason, all the significance levels quoted here should be considered as approximate and, potentially, as "over-significant".

Turning to the actual parameter estimates, a few general points may be made. As already said, the estimates were generally of low residual variance, especially in Model 1, and consistent when convergence was started from different values, (with the exception of the 21 individuals who failed to converge in Model 2). Little can be said about the distribution of the parameters because of the low numbers, but there seemed to be no evidence of marked non-normality.

As stated under Results, it was difficult to attribute simple biological meaning to the parameters in any of the models. It was possible, however, to make a few observations. In Models 1 and 2, s_0 and s_1 were certainly related to prepubertal velocity and PHV respectively, as might be expected from the original structure of the model. These parameters were the lower and upper asymptotes of s and therefore governed the changing rate at which h approached h_1.

The height parameters h_1 and h_θ were clear in their biological meaning and θ was obviously very near to age at PHV, with which it was highly correlated. This pattern held in all the first three models.

In conclusion, it would appear that there were two candidates for consideration as useful, new, mathematical models for growth curves. These were Models 1 and 3. Model 3 offered somewhat greater flexibility and had the useful property of reducing to Model 1 when the extra parameter was unnecessary. It never gave a less satisfactory fit to the data than Model 1. However, the attraction of Model 1 remains its extreme robustness and simplicity; it contains only five parameters and has a particularly simple functional form. Moreover, over the time domain, two years to maturity, it can be expressed in the differential equation form:

$$\frac{dh}{dt}=s(t)\,.\,(h_1-h)$$

$$\frac{ds}{dt}=(s_1-s)(s-s_0)$$

in which the number of unknown parameters is reduced to three; h_1, s_0 and s_1. Exploration of younger ages was not possible with the data used but would be a worthwhile area for more investigation.

Acknowledgments

We would like to thank Professor J. M. Tanner for helpful discussions and for allowing us to use the Harpenden data. Useful comments on our ideas were given by Professors R. N. Curnow, M. J. R. Healy and W. A. Marshall. We would like to thank Jan, wife of M.A.P. and sister of M.J.B., for bringing us together!

Appendix

Derivation of a family of descriptive growth functions

The derivation of the growth functions is described here in a series of stages to indicate the reasons behind the various assumptions that have been made. Because of the previous successful fitting of the upper, adolescent region of the growth curve by the logistic function, this is taken as the starting point of the analysis.

The logistic function has usually been fitted in the form:

$$h = P + \frac{K}{1 + \exp(a - bt)} \tag{1}$$

P, K, a and b being the parameters. However, for comparison with later analysis, it is convenient to state it in the following form:

$$y = \frac{y_0 \exp[\gamma y_0(t-\theta] + y_1 \exp[\gamma y_1(t-\theta)]}{\exp[\gamma y_0(t-\theta)] + \exp[\gamma y_1(t-\theta)]} \tag{2}$$

where γ, y_0, y_1 and θ are parameters related to the previous ones by:

$$y_0 = P,\ y_1 = P + K,\ \gamma = b/K \text{ and } \theta = a/b$$

The curve described by this function is shown in figure A.1 where y_0, y_1 and θ are explicit in their meaning; γ is proportional to the slope of y at $t = \theta$, the value about which y is centred.

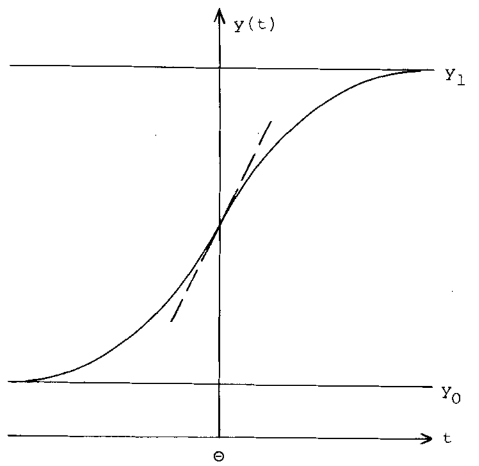


Figure A.1. The general form of the logistic function defined in equation (6). Note that $\left.\frac{dy}{dt}\right|_{t=\theta} = \gamma(y_0 - y_1)^2$, which is represented by the broken line.

In what follows, we rely strongly on the fact that the expression (2) is the solution of the differential equation:

$$\frac{dy}{dt} = \gamma(y_1 - y)(y - y_0) \tag{3}$$

Since the logistic function of (1) or (2) fits the latter part of the growth curve well, the height at time t, $h(t)$ may be expected similarly to satisfy:

$$\frac{dh}{dt} = \gamma(h_1 - h)(h - h_0) \tag{4}$$

in this region. The constant h_1, is immediately identifiable as adult height and the solution is the expression (2) with y replaced by h. The growth curve derived from this model is shown in figure A.1 and obviously cannot adequately describe the lower, pre-adolescent part of the curve. It might, however, be expected that the differential equation (4) would remain valid for large values of t.

Let us write (4) as:

$$\frac{dh}{dt}=s(h)(h_1-h) \quad (5)$$

where $s(h)=\gamma(h-h_0)$.
The form of $s(h)$ is a logistic function with zero as its lower asymptote. However, studying the behaviour of $s(h)$ from a sample of real data, using the relationship:

$$s(h)=\frac{dh}{dt}\cdot\frac{1}{(h_1-h)} \quad (6)$$

(dh/dt being derived numerically) the form of $s(h)$ was found to roughly approximate a sigmoid shape with a non-zero lower asymptote. Examples are shown in figure A.2. This suggested that the function s could be taken as a logistic function.

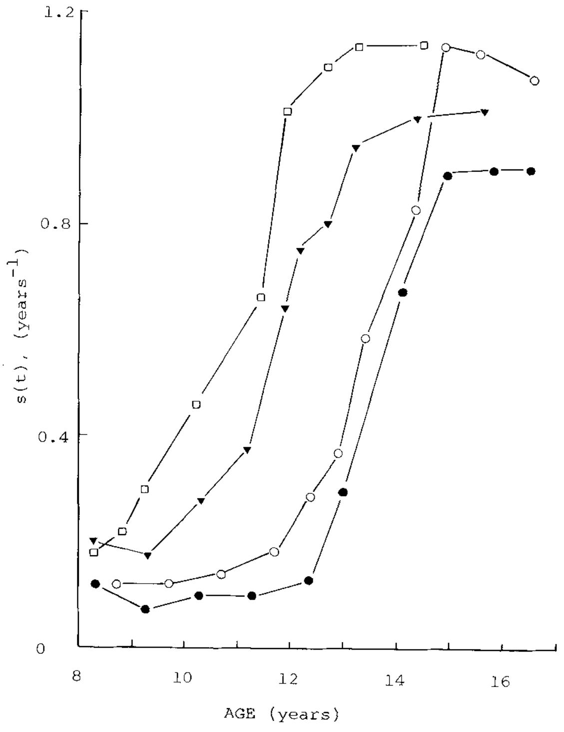


Figure A.2. Examples indicating the rough shape of the function $s(t)=s(h)$, as defined by equation (15). Each curve represents values for $s(t)$ obtained for a single individual.

Considering now s as a function of time, rather than of height, and therefore expressing it as $s(t)$, the following system of differential equations is obtained:

$$\frac{dh}{dt}=s(t)\,.\,(h_1-h) \quad (7)$$

$$\frac{ds}{dt}=\gamma(s_1-s)(s-s_0) \quad (8)$$

In words, if s is given logistic properties then in (7) the rate at which h approaches h_1 (i.e. dh/dt) is governed by the distance h is below h_1 and a function, dependent on time, that has a central period of rapid acceleration. Therefore, at t near zero the growth velocity is the product of some small constant and (h_1-h), but subsequently s rapidly increases to a new, upper asymptote producing the appearance of the adolescent growth spurt.

To obtain s as a function of t, explicitly, equation (8) is integrated (cf. equations (2) and (3)) to give:

$$s=\frac{s_0 \exp [\gamma s_0(t-\theta)]+s_1 \exp [\gamma s_1(t-\theta)]}{\exp [\gamma s_0(t-\theta)]+\exp [\gamma s_1(t-\theta)]}. \qquad (9)$$

Integration of equation (7) is straightforward since

$$\frac{d}{dt}\left[\ln\left(\frac{1}{h_1-h}\right)\right]=s=\frac{d}{dt}\,[\ln\{\exp [\gamma s_0(t-\theta)]+\exp [\gamma s_1(t-\theta)]\}^{1/\gamma}] \qquad (10)$$

which leads to:

$$\ln\left(\frac{1}{h_1-h}\right)=\ln\{\exp [\gamma s_0(t-\theta)]+\exp [\gamma s_1(t-\theta)]\}^{1/\gamma}+\ln\left\{\frac{1}{2^{1/\gamma}(h_1-h_\theta)}\right\}, \qquad (11)$$

the last term being the constant of integration, for which $h=h_\theta$ when $t=\theta$.

Gathering terms and simplifying produces:

$$h=h_1-\frac{(h_1-h_\theta)}{\{\frac{1}{2}\exp [\gamma s_0(t-\theta)]+\frac{1}{2}\exp [\gamma s_1(t-\theta)]\}^{1/\gamma}} \qquad (12)$$

We refer to this function as Model 2. The corresponding model, when $\gamma=1$, namely,

$$h=h_1-\frac{2(h_1-h_\theta)}{\exp [s_0(t-\theta)]+\exp [s_1(t-\theta)]} \qquad (13)$$

will be referred to as Model 1. It can be shown that γ must be close to 1 to make Model 2 approach the logistic form expected for large values of t, and taking $\gamma=1$ in fact produces a very good model, irrespective of its connection with Model 2.

There are only five parameters in Model 1, two height parameters h_θ and h_1 (adult height); a time parameter, θ, and two rate constants s_0 and s_1 having the dimensions of inverse time. The function s for this model has the form

$$s=\frac{s_0 \exp [s_0(t-\theta)]+s_1 \exp [s_1(t-\theta)]}{\exp [s_0(t-\theta)]+\exp [s_1(t-\theta)]} \qquad (14)$$

and the velocity function is

$$\frac{dh}{dt}=s(t)(h_1-h)=\frac{2(h_1-h_\theta)\{s_0 \exp [s_0(t-\theta]+s_1 \exp [s_1(t-\theta)]\}}{\{\exp [s_0(t-\theta)]+\exp [s_1(t-\theta)]\}^2} \qquad (15)$$

For both Models 1 and 2 the acceleration function is

$$\frac{d^2h}{dt^2}=\frac{ds}{dt}(h_1-h)+s\left(-\frac{dh}{dt}\right) \qquad (16)$$

$$=\gamma(h_1-h)[-2s^2+(s_0+s_1)s-s_0s_1] \qquad (17)$$

The positions of maximum and minimum growth velocity are obtained when the r.h.s. of the acceleration function is put equal to zero. Apart from the trivial solution, when $h=h_1$, the values of s at these points are the solutions of

$$s^2-\tfrac{1}{2}(s_0+s_1)s+\tfrac{1}{2}s_0s_1=0 \tag{18}$$

giving:

$$s=\tfrac{1}{4}(s_0+s_1)\pm\sqrt{\{\tfrac{1}{4}(s_0+s_1)^2-\tfrac{1}{2}s_0s_1\}} \tag{19}$$

We can therefore obtain analytically, from (9) or (14) age at "take-off" and peak height velocity and, subsequently, height, velocity, etc. at these points.

As preliminary studies showed high correlations between s_1 and γ a minor re-parameterization was carried out such that Model 2 was finally studied in the form:

$$h=h_1-\frac{(h_1-h_\theta)}{\{\frac{1}{2}\exp[\gamma s_0(t-\theta)]+\frac{1}{2}\exp[s_1'(t-\theta)]\}^{1/\gamma}} \tag{12 a}$$

Improvements to Model 1 can be made in a different way. Further flexibility in the model may be expected if the function $s(t)$ is not necessarily symmetric. If s in (9) is taken as the sum of two logistic functions with different parameters, then it will be asymmetric in general and the fit might be improved; the mathematical analysis remains straightforward.

Writing $s=p+q$ where

$$\frac{dp}{dt}=(p_1-p)(p-p_0),\quad \frac{dq}{dt}=(q_1-q)(q-q_0) \tag{20), (21}$$

and assuming unitary constants of proportionality, as in Model 1, we obtain

$$p=\frac{d}{dt}[\ln\{\exp[p_0(t-\theta)]+\exp[p_1(t-\theta)]\}] \tag{22}$$

$$q=\frac{d}{dt}[\ln\{\exp[q_0(t-\theta)]+\exp[q_1(t-\theta)]\}], \tag{23}$$

although strictly the θ need not be the same for both p and q. Solving (7) with the modifications to s, now yields:

$$h=h_1-\frac{4(h_1-h_\theta)}{\{\exp[p_0(t-\theta)]+\exp[p_1(t-\theta)]\}\{\exp[q_0(t-\theta)+\exp[q_1(t-\theta)]\}} \tag{24}$$

Although there are two further parameters in (24), it was discovered that one (e.g. q_0) could be set equal to zero without seriously affecting the generality.

$$h=h_1-\frac{4(h_1-h_\theta)}{\{\exp[p_0(t-\theta)]+\exp[p_1(t-\theta)]\}\{1+\exp[q_1(t-\theta)]\}} \tag{25}$$

We refer to this model as Model 3; note that if $p_0=p_1$, Model 1 is regained.

This function does not, however, provide an analytical solution for the maximum and minimum velocities, as there is no simple, analytical expression for the roots of the equation

$$s=\frac{p_0\exp[p_0(t-\theta)+p_1\exp[p_1(t-\theta)]}{\exp[p_0(t-\theta)]+\exp[p_1(t-\theta)]}+\frac{q_0\exp[q_0(t-\theta)]+q_1\exp[q_1(t-\theta)]}{\exp[q_0(t-\theta)]+\exp[q_1(t-\theta)]} \tag{26}$$

when s is given by (19).

References

Bock, R. D., Wainer, H., Peterson, A., Thissen, D., Murray, J., and Roche, A. (1973). A parameterization for individual human growth curves. *Human Biology*, **45**, 63–80.

Bock, R. D., and Thissen, D. (1976). Fitting multi-component models for growth in stature. *Proceedings of the 9th International Biometrics Conference*, Biometric Society, Boston.

Count, E. W. (1942). A quantitative analysis of growth in certain human skull dimensions. *Human Biology*, **14**, 143–165.

Count, E. W. (1943). Growth patterns of the human physique: an approach to kinetic anthropometry. *Human Biology*, **15**, 1–32.

Daniel, C., and Wood, F. S. (1971). *Fitting Equations to Data*. New York: John Wiley and Sons, Inc.

Deming, J. (1957). Application of the Gompertz curve to the observed pattern of growth in length of 48 individual boys and girls during the adolescent cycle of growth. *Human Biology*, **29**, 83–122.

Draper, N. R., and Smith, H. (1966). *Applied Regression Analysis*. New York: John Wiley and Sons, Inc.

Durbin, J., and Watson, G. S. (1971). Testing for serial correlation in least-squares regression. *Biometrika*, **58**, 1–19.

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

Joossens, J. J., and Brems-Heyns, E. (1975). High power polynomial regression for the study of distance, velocity and acceleration of growth. *Growth*, **39**, 535–551.

Marquadt, D. W. (1963). An algorithm for least squares estimation of non-linear parameters. *Journal of the Society of Industrial and Applied Mathematics*, **11**, 431–441.

Marubini, E., Resele, L. F., and Barghini, G. (1971). A comparative fitting of the Gompertz and logistic functions to longitudinal height data during adolescence in girls. *Human Biology*, **43**, 237–252.

Marubini, E., Resele, L. F., Tanner, J. M., and Whitehouse, R. H. (1972). The fit of the Gompertz and logistic curves to longitudinal data during adolescence on height, sitting height, and biacromial diameter in boys and girls of the Harpenden Growth Study. *Human Biology*, **44**, 511–524.

Nelder, J. A., and Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, **7**, 308–313.

Stützle, W., Gasser, Th., and Largo, R. (1966). Analysis of the adolescent growth spurt using smoothing spline functions. *Réunion de Co-ordination des Recherches sur la Croissance et le Développement*. Rennes: Centre International de l'Enfance.

Tanner, J. M., Whitehouse, R. H., Marubini, E., and Resele, L. F. (1976). The adolescent growth spurt of boys and girls of the Harpenden Growth Study. *Annals of Human Biology*, **3**, 109–126.

Watson, G. S. (1955). Serial correlation in regression analysis. I. *Biometrika*, **42**, 327–341.

Address correspondence to: Dr. M. A. Preece, Department of Growth and Development, Institute of Child Health, 30 Guilford Street, London WC1N 1EH.

Zusammenfassung. Eine neue Familie von mathematischen Funktionen für Längsschnitt-Wachstumsdaten wird beschrieben. Alle ihre Mitglieder leiten sich ab von der Differentialgleichung $\mathrm{d}h/\mathrm{d}t = s(t)\,.\,(h_1 - h)$, wobei h_1 die erwachsene Größe und $s(t)$ eine Zeitfunktion ist. Die Form von $s(t)$ wird durch eine von vielen Funktionen gegeben, wobei alle Differentialgleichungen sind, so daß sie eine Familie von verschiedenen Modellen ergibt.

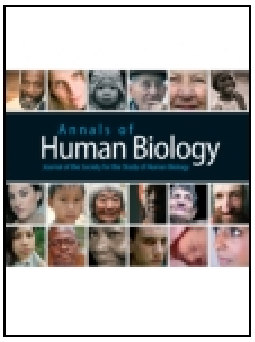
Drei Versionen werden verglichen. Bei allen wurde gefunden, daß sie den früheren Modellen überlegen sind. Modell 1, bei dem $s(t)$ durch $\mathrm{d}s/\mathrm{d}t = (s_1 - s)\ (s - s_0)$ definiert war, war besonders genau und robust und enthielt nur 5 Parameter zur Beschreibung des Körperhöhenwachstums vom Alter zwei bis zur Reife.

Abgeleitete „biologische" Parameter wie höchster Körperhöhenzuwachs waren bei diesen drei Mitgliedern der Familie sehr stabil, unterschieden sich aber in einigen Fällen signifikant von vorhergehenden Schätzungen.

Résumé. Une nouvelle famille de fonctions mathématiques destinées à s'ajuster aux données longitudinales de croissance est décrite. Tous ses membres décrivent de l'équation différencielle $\mathrm{d}h/\mathrm{d}t = s(t)\,.\,(h_1 - h)$ où h_1 est le format adulte et $s(t)$ une fonction du temps. La forme de $s(t)$ est donné par l'une de nombreuses fonctions, toutes des équations différencielles, et engendre ainsi une famille de modèles différents.

Trois versions ont été comparées. Toutes ont été trouvées supérieures aux modèles précédemment décrits. Le modèle 1, dans lequel $s(t)$ est défini par $\mathrm{d}s/\mathrm{d}t = (s_1 - s)(s - s_0)$, est particulièrement adéquat et robuste tout en ne comportant que cinq paramètres pour décrire la croissance staturale de deux ans à la maturité.

Des paramètres "biologiques" dérivés tels que la vitesse au pic de croissance en hauteur se montraient très consistants de l'un à l'autre des trois membres de cette famille, mais dans certains cas ils différaient significativement des estimations antérieures.



Annals of Human Biology

ISSN: 0301-4460 (Print) 1464-5033 (Online) Journal homepage: www.tandfonline.com/journals/iahb20

A new family of mathematical models describing the human growth curve—Erratum: Direct calculation of peak height velocity, age at take-off and associated quantities

Adrian Sayers, Mike Baines & Kate Tilling

To cite this article: Adrian Sayers, Mike Baines & Kate Tilling (2013) A new family of mathematical models describing the human growth curve—Erratum: Direct calculation of peak height velocity, age at take-off and associated quantities, Annals of Human Biology, 40:3, 298-299, DOI: 10.3109/03014460.2013.772655

To link to this article: https://doi.org/10.3109/03014460.2013.772655

Published online: 06 Mar 2013.

Annals of Human Biology, May–June 2013; 40(3): 298–299

ISSN 0301-4460 print/ISSN 1464-5033 online
DOI: 10.3109/03014460.2013.772655

informa
healthcare

LETTER TO THE EDITOR

A new family of mathematical models describing the human growth curve—Erratum: Direct calculation of peak height velocity, age at take-off and associated quantities

Adrian Sayers[1], Mike Baines[2] & Kate Tilling[1]

[1]*School of Social and Community Based Medicine, Faculty of Medicine and Dentistry, University of Bristol, Bristol, UK, and* [2]*Department of Mathematics and Statistics, University of Reading, Reading, UK*

A new family of mathematical functions to fit longitudinal growth data was described in 1978. The ability of researchers to directly use parameters as estimates of age at peak height velocity resulted in them overlooking the possibility of directly calculating these quantities after model estimation. This erratum has corrected three mistakes in the original manuscript in the direct calculation of peak height velocity and age at take-off and has implemented the solutions in a STATA program which directly calculates the estimates, standard errors and confidence intervals for age, height and velocity at peak height velocity.

Keywords: Growth model, peak height velocity, Preece Baines

ERRATUM

Preece and Baines, in their original 1978 article (Preece and Baines 1978), proposed three new mathematical models of human growth. Model 1, described as their simple and robust model, has become popular and widespread in its simulation of childhood growth, and their original paper has been cited in excess of 300 times.

Uniquely, the models were derived from a differential equation $dh/dt = s(t)\cdot(h_1 - h)$, where h_1 is adult size and $s(t)$ is a function of time satisfying the logistic differential equation $ds/dt = (s_1 - s(t))\cdot(s(t) - s_0)$, where s_0 and s_1 are parameters.

The utility of these models was partly due to the correlation between the estimated parameters and suggested biologically meaningful time-points including age at peak height velocity (PHV) and age at 'take off' (TO, onset of rapid pubertal growth). Suggested interpretations of each parameter are given in the original paper (Preece and Baines 1978). Of the three models proposed, only two had analytical solutions to these meaningful time points, and the errata concerns these calculations.

The two models of $h =$ height and $t =$ time which we address are described as model 1 and model 2 in the original paper. Model 2, see below, describes height in six parameters, and Model 1, see below, is a special case of model 2 when $\gamma = 1$ (as γ is now constant it ceases to be a parameter and, therefore, model 1 describes height in five parameters). For Model 1, the form was

$$h = h_1 - \frac{2(h_1 - h_\theta)}{\left(\exp[s_0(t-\theta)] + \exp[s_1(t-\theta)]\right)} \quad (1)$$

while for Model 2,

$$h = h_1 - \frac{(h_1 - h_\theta)}{\left(\frac{1}{2}\exp[\gamma s_0(t-\theta)] + \frac{1}{2}\exp[\gamma s_1(t-\theta)]\right)^{1/\gamma}} \quad (2)$$

Model 2 was less successful in reaching convergence and, therefore, the use of the more parsimonious model 1 was preferred.

Whilst applying these models to the analysis of height data, we have found three errors in the original paper. The first error, which concerns the acceleration function, is applicable to both models 1 and 2.

The acceleration function, obtained by differentiating dh/dt, is

$$\frac{d^2h}{dt^2} = \frac{ds}{dt}(h_1 - h) + s\left(-\frac{dh}{dt}\right) \quad (3)$$

The original paper suggests that, after the substitution of dh/dt (the derivative of height with respect to time, t, i.e velocity),

$$\frac{dh}{dt} = s(t)\cdot(h_1 - h) \quad (4)$$

Correspondence: Adrian Sayers, University of Bristol, School of Social and Community Based Medicine, Canygne Hall, 39 Whatley Road, Bristol BS8 2PS, UK. E-mail: adrian.sayers@bristol.ac.uk

(*Received 19 November 2012; revised 11 January 2013; accepted 22 January 2013*)

and ds/dt (derivative of s with respect to time, t)

$$\frac{ds}{dt} = \gamma \cdot (s_1 - s)(s - h_0) \tag{5}$$

the acceleration function is equal to

$$\frac{d^2h}{dt^2} = \gamma(h_1 - h)\left[-2s^2 + (s_0 + s_1)s - s_0s_1\right] \tag{6}$$

However, the correct expression for the acceleration is in fact

$$\frac{d^2h}{dt^2} = (h_1 - h)\left\{\left(\gamma\left[(s_0 + s_1)s - s_0s_1 - s^2\right]\right) - s^2\right\} \tag{7}$$

This acceleration expression simplifies in the special case of model 1 (when $\gamma = 1$) to the expression

$$\frac{d^2h}{dt^2} = (h_1 - h)\left[-2s^2 + (s_0 + s_1)s - s_0s_1\right] \tag{8}$$

given in the original paper (note that this expression does not include the parameter γ).

The second error we have noted, which is applicable to model 1, occurs in the process of calculating age at peak height velocity/take-off.

In the case of model 1 the acceleration function (Equation 8) is set equal to zero and solved for s. Apart from the trivial solution when $h = h_1$, it follows that

$$s^2 - \frac{1}{2}(s_0 + s_1)s + \frac{1}{2}s_0s_1 = 0 \tag{9}$$

which the original paper suggests can be solved to give

$$s = \frac{1}{4}(s_0 + s_1) \pm \sqrt{\left[\frac{1}{4}(s_0 + s_1)^2 - \frac{1}{2}s_0s_1\right]} \tag{10}$$

However, after applying the general solution to a quadratic equation and gathering and simplifying, we find that the correct solution is:

$$s = \frac{1}{4}(s_0 + s_1) \pm \sqrt{\left(\frac{1}{4}(s_0 + s_1)\right)^2 - \frac{1}{2}s_0s_1} \tag{11}$$

As Equation (8) does not generalize to model 2, Equation (11) is also not appropriate for model 2.

The third error stems from this lack of generalization. The acceleration function (equation 7) (which is generalizable to both model 2 and model 1 assuming $\gamma = 1$), can be further manipulated into the form

$$\frac{d^2h}{dt^2} = (h_1 - h)\left\{s^2\left(-1 - \gamma\right) + s\gamma(s_0 + s_1) - \gamma s_0s_1\right\} \tag{12}$$

which is easily solved. After setting the right hand side to zero and ignoring the trivial solution when $h = h_1$, the general solution to a quadratic equation is applied and, after subsequent simplification and gathering procedures, we find that

$$s = \frac{-\gamma(s_0 + s_1) \pm \sqrt{\gamma^2(s_0 - s_1)^2 + 4\gamma s_0s_1}}{\left(-2 - 2\gamma\right)} \tag{13}$$

Substituting s from Equation (13) into Equation (14),

$$s = \frac{s_0 \exp\left(\gamma s_0(t - \theta)\right) + s_1 \exp\left(\gamma s_1(t - \theta)\right)}{\exp\left(\gamma s_0(t - \theta)\right) + \exp\left(\gamma s_1(t - \theta)\right)} \tag{14}$$

and solving for t,

$$t = \theta + \frac{\log_e\left(-\frac{(s - s_1)}{(s - s_0)}\right)}{\gamma(s_0 - s_1)} \tag{15}$$

allows the calculation of time (age) at PHV to be calculated.

Once the time of PHV/TO is known, t may be substituted into either model 1 or 2 (Equation 1 or 2) and, thus, height at PHV/TO may be calculated. Similarly once the height at PHV/TO is known and S from Equation (13) is known, these can be substituted into Equation (4) to calculate velocity at TO and PHV.

Thus, the growth traits previously mentioned can be calculated directly, rather than by using the parameters (and their standard errors) estimated from the model as proxies.

With the advances in statistical computation it is now possible to automatically calculate standard errors of these derived parameters by applying the delta method, instead of conducting a one-step Taylor series expansion manually. Therefore, intervals and limits are easily available for the analytic calculations of age at peak height velocity and age at take-off. The solutions to the previous expressions have now been implemented into STATA (STATA 2011) package, which utilizes `nl` and `nlcom` routines, which yield parameters, standard errors and confidence intervals.

These corrections and advances in statistical software should enhance the ability of researchers to derive biologically meaningful growth parameters from sequential height data using either of the first two original Preece-Baines models.

The STATA program `pbreg` is available from the authors and from the SSC archive.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES

Preece MA, Baines MJ. 1978. A new family of mathematical models describing the human growth curve. Ann Hum Biol 5:1–24.

STATA, StataCorp. 2011. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP.