A Mathematical Procedure for Estimating the **Spatial Relationships between Lung Function**, Somatic Growth, and Maturation

DUANE L. SHERRILL, WAYNE J. MORGAN,¹ LYNN M. TAUSSIG, LOUIS I. LANDAU, AND **BENJAMIN BURROWS**

Westend Research Laboratory, Division of Respiratory Sciences, University of Arizona College of Medicine, Tucson, Arizona 85724

ABSTRACT. A mathematical procedure is described for examining growth velocity rates of lung function (FEV₁) in relationship to somatic growth and maturation. This procedure uses a polynomial smoothing spline model to yield a fitted curve through the data and to estimate the process first derivatives (i.e. growth velocity curves). We demonstrate this technique using data from children and adolescents enrolled in the Tucson Epidemiological Study of Airway Obstructive Diseases. The study group consisted of 772 healthy normal subjects, aged 3 to 25 y. The results for the normal subjects (male and female) indicate that the growth velocity peak (GVP) of somatic growth leads the GVP of functional growth (FEV₁) by approximately 7 and 11 mo for females and male subjects, respectively, and that the GVP of maturation lags behind that of functional growth (FEV₁) by approximately 1 y (male and female subjects). In addition, the normal subjects' growth velocity curves for FVC and FEV₁/FVC were examined. The FEV₁/ FVC ratio was consistently high, and its growth velocity was not significantly different than 0 over the age range studied. This suggests that in normal children, the GVP seen in FEV₁ is primarily due to a GVP in vital capacity or lung vol, rather than to a direct effect on expiratory flow rates alone. From this study we concluded that the polynomial smoothing spline procedure can adequately model the inherently noisy pulmonary function data and additionally yield an accurate estimate of the process first derivative. (Pediatr Res 25:316-321, 1989)

Abbreviations

FVC, forced vital capacity FEV₁, forced expired volume in one second SG, somatic growth M, maturation GVP, growth velocity peak

The relationships between lung function and anthropometric growth in adolescents have been studied by several investigators (1-9). Although some reports have focused on the reproducibility of lung function and anthropometric variables (6), others have explored growth in lung function and its relation to indices of

Received August 1, 1988; accepted November 10, 1988. Correspondence Dr. Duane L. Sherrill, Division of Respiratory Sciences, University of Arizona College of Medicine, Tucson, AZ 85724. Supported by NHLBI SCOR Grant HL 14136.

thoracic growth and stature (2, 3, 5, 6). The latter studies have all reported an apparent lag time between growth in stature and subsequent growth in lung function. That is to say, increases in lung function continue to occur for at least 1 y after cessation of increase in stature. These results were also suggested by Simon et al. (8), who showed that the growth spurt for lung length (determined from chest radiographs) occurred 6 mo after that for standing height, for males and females ages 6-15 y.

The two major contributing factors of functional lung development are: 1) SG assessed by a power function of height, which as mentioned above, leads functional growth; and 2) M, which we propose lags behind the functional growth component. For our purposes, lung M is defined as that portion of functional lung development which is not directly related to changes in SG. An estimate of M is then determined as that portion of the observed FEV_1 not predicted, using a power function of simple allometry. This approach assumes that FEV₁ growth can be described by a simple allometry model and that any deviations from the predicted model can be attributed to lung M. If SG and M do account for the overall changes in functional lung growth (*i.e.* growth in FEV_1 or FVC), then the sum of the their corresponding growth velocity curves should, when properly scaled, be similar to that of lung function.

In the past, prediction of expected pulmonary function variables were based primarily on multiple regression models. As these were limited for describing the complex data of functional growth, other parametric modeling methods were developed to examine growth during specific age intervals (NHLBI Workshop, 1986). Some authors modeled these age intervals using ordinary least squares methods to produce predictive equations (7, 9); others used mathematical fitting techniques to yield relationships that would have more reasonable biologic interpretations (3). Most of these approaches were designed to insure continuous solutions throughout the range of data. Further, exponential models of velocity (2) require arbitrary assumptions concerning the underlying shape of the velocity curve. Thus, nonparametric regression models (i.e. smoothing splines, which do not require that the true underlying function be known) should prove useful (10-12).

In this report, we present a mathematical smoothing procedure to examine the spatial relationships between functional lung growth, somatic growth and maturation. This procedure yields a smoothing curve through the data and estimates the process first derivatives (*i.e.* growth velocity curves) for each of the response variables. The lead-lag relationships between these factors are then determined by the temporal relationships between their GVP. We demonstrate this approach using pulmonary function test response data from children and adolescents enrolled in the Tucson Epidemiological Study of Airway Obstructive Diseases (13).

¹ Supported by NHLBI Clinical Investigator Award HL 01377.

Study population. Subjects selected to illustrate this procedure were participants in the Tucson Epidemiological Study of Airways Obstructive Disease. Details of selection of the study population and of the methods for obtaining expiratory flow information have been described (13). Only technically satisfactory spirometric tests are included in the selected groups. In total, 879 tests in 398 different male subjects and 793 tests in 374 female subjects aged 3-25 y have been used in the normal nonsmoking groups. Subjects were selected for the healthy normal group if they had never been diagnosed as having asthma or chronic bronchitis. In addition, subjects who had ever had an observed FEV1 less than 80% of predicted, based on previously published prediction equations (3), were also excluded. The distributions of the number of observations/subject and total number of data points for the two groups (males and females) are tabulated in Table 1. These data were collected over eight surveys, which included pulmonary function testing. Thus, a subject could have a maximum of eight pulmonary function measures.

Statistical model. We chose a polynomial smoothing spline as the mathematical model for describing this complex biologic data. This polynomial smoothing spline model fits an Nth degree polynomial (where N is an odd integer) to the intervals between adjacent data points. It uses the remainder of the data for estimating the polynomial coefficients (10-12). The individual polynomials are then "grafted" together (i.e. as a "spline") in a manner which yields continuous curves for the first derivatives of the model. This particular smoothing procedure yields maximum likelihood estimates of the "true" values corresponding to the observed data. The conditional maximum likelihood function (L), which contains both linear and nonlinear parameters, is calculated recursively using a Kalmen Filter State Space Model (14). (The "state space" model also yields estimates of the process derivatives.) The natural log of L function can be minimized by estimating the linear and nonlinear parameters separately, the former by ordinary least squares and the latter by a nonlinear optimization procedure (15). The solution to L can also be used for selecting the degree of polynomial spline that yields the best possible fit to the data: this then determines the degree of the derivatives obtained. (For example, a first degree polynomial is simply a straight line, and the model only yields the fitted spline. A third degree polynomial yields both the fitted spline and the first derivative.) Confidence intervals for the splines, and corresponding derivatives, are estimated from the resulting covariance matrix. In addition, because the smoothing spline procedure is a nonparametric technique and thus makes no assumptions concerning the distribution of the data, the spline fits are less effected by the irregular distribution of the data set (Table 1).

In the context of this report, M was determined as that portion of the lung function response (*e.g.* FEV₁ or FVC) that could not be attributed to somatic growth. Hence, it was estimated by dividing each observed functional response by its height-predicted estimate. These predicted values were estimated using the power function $Y = a \times k$, as described in a previous report (3). Briefly, the gender-specific exponents were determined using log-

Table 1. Distribution of number of data points/normal subject

No. of points	Male	Female
1	194 (194)	184 (184
2	130 (65)	160 (80
3	186 (62)	147 (49
4	152 (38)	112 (28)
5	120 (24)	85 (17)
6	60 (10)	54 (9)
>7	37 (5)	51 (7)
Totals	879 (398)	793 (374)

log regression analysis between FEV₁ or FVC (dependent variable Y) and height (independent variable X), over specified age intervals. These age intervals were selected to minimize age-height interactions that became evident in the residuals for the older ages (3). In addition, an estimate of SG assessed by height was determined using the power function exponent to yield SG = HT^k. These power function estimates of SG yield estimates of body size which have vol U (inches**k) that are more appropriate for comparisons with lung function measures (*i.e.* FEV₁ or FVC).

Then for comparative purposes we assumed that dM and dSG were linearly related to $dFEV_1$. This relationship can be expressed by the equation

$$dFEV_1 = k1dSG + k2dM + e$$
(1)

where d indicates the first derivative, k1 and k2 are coefficients to be estimated using multivariate regression analysis, and e is the error term that is assumed to be normally distributed with 0 mean and variance matrix sigma (16).

To facilitate comparisons between groups (males and females), the GVP are represented as vectors and plotted using vector diagrams (17), the length of each vector being proportional to its adjusted GVP amplitude (adjusted using k1 and k2 from equation 1) and its angle of rotation being proportional to its time shift relative to the normal subjects' FEV₁ GVP, this being arbitrarily set to 45° for both males and females. For example, if the M GVP occurred 1 y after the FEV₁ GVP in normal subjects, this would be represented by a vector at 65° (*i.e.* lagging functional growth by 20°). In this scheme, a vector at 45° would represent no phase shift, but a vector coincident with the x axis would represent a lead time of 2.25 y and a vector coincident with the y axis would represent a lag time of 2.25 y. Thus, each degree of rotation for a vector is equivalent to a time shift of 0.05 y, with larger angles representing longer time lags.

RESULTS

Figure 1 illustrates the raw data (dots) with the fitted third degree smoothing spline models (upper solid lines). Corresponding growth velocity curves (lower solid lines) for normal subjects' pulmonary function (FEV₁), somatic growth, and M data derived as the first derivatives of the spline are also shown. These plots show that the male subjects experience much higher GVP than females and that these peaks occur later in life. The growth velocity curves can in general be characterized by an initial steady-state period (*i.e.* period of constant growth), followed by an acceleration-deceleration period, and ending with another quasi-steady-state period. The M velocity curves differ from the others in that their initial as well as their ending steady-state phases (Fig. 1 E and F) are not significantly different than 0. This suggests that maturation, as we have defined it, is only contributing to functional development during its accelerationdeceleration period. In contrast, somatic (Fig. 1 C and D) and lung function growth velocity curves (Fig. 1 A and B) have significantly elevated initial steady-state phases. In addition, the M estimates appear to have more variability, which is likely a consequence of taking the ratio of two relatively noisy variables (observed FEV_1 and height predicted FEV_1), but could also result from subjects experiencing cessation of growth in height and onset of M at different times.

The normal subjects data set (Table 1) included a large crosssectional sample (no. studies/subject = 1) of 22% and 23% for males and females, respectively. To insure that these individuals' observations did not alter the longitudinal trends in the response data, we repeated the smoothing spline analysis, excluding all subjects that had only one observation. The 95% confidence intervals of the smoothing spline curves were then used to compare the fitted response curves between the two groups. The results indicate that the spline curves did not differ significantly



Fig. 1. FEV₁, SG, and M data (*dots*) for normal subjects (allometric exponent estimates were k = 2.45 for males and k = 2.34 for females). Fitted third degree smoothing splines (*upper solid lines*) and corresponding first derivatives (*lower solid lines*) for male (A, C, E) and female (B, D, F) subjects. Broken line, 0 growth velocity level.

when subjects with a single observation were excluded. Therefore, we elected to include those subjects in all subsequent analyses.

The results of adjusting the normal subjects growth velocity data using equation 1 are shown in Figure 2. For both normal male and female subjects, it becomes clear that the somatic GVP leads the FEV₁ GVP and that the maturation GVP lags behind the FEV₁ GVP, as expected. The ages at which the GVP occurred and adjusted amplitudes of the peaks are listed in Table 2. The lead times for the somatic GVP were approximately 7 and 11 mo for female and male subjects, respectively, and the lag time for the maturation GVP was approximately 1 y (male and female subjects). The adjusted growth velocity curves also demonstrate that functional growth during the preadolescent period can almost totally be accounted for by SG. Maturation, by contrast, makes its largest contribution starting at the onset of the adolescent growth spurt and continues developing after the SG spurt has ceased. The maturation factor only surpasses the SG contribution at a point approximately halfway through its deceleration phase.

The amplitude and phase relationships between these GVP become clearer when they are represented as vectors. Figure 3 shows the vector plots for normal male and female subjects. The absolute age of occurrence of the GVP can be verified on the vector plot (Fig. 3) by using the circular age axis (male or female). For example, the male dFEV₁ vector is pointing to 14.25 y on the male axis; for females, the dFEV₁ vector points to 11.83 y on the female axis. This circular axis also illustrates that even though the female dSG vector lags behind the male dSG vector, when dFEV₁ are referenced to 45°, it actually occurs at a younger age (11.25 versus 13.50), as has frequently been reported (17).

One can see (Figure 3) that with the lung function vectors both

referenced to the same point in time (*i.e.* 45°), the M vectors actually coincide and differ only in amplitude. Likewise, the SG vectors only differ slightly in phase, with females having a GVP in SG closer to their GVP in FEV₁. Females in general do not achieve as high a growth rate as their male counterparts.

This procedure was also applied to the FVC and FEV_1/FVC data, for both male and female subjects. The ages at which the FVC GVP occurred and adjusted amplitudes of the peaks are listed in Table 3. The amplitudes of the dFVC GVP were larger then those estimated for dFEV₁, as expected, for both male and female subjects. In contrast, the dM GVP were smaller for both sexes. The ages at which the FVC GVP occurred were similar to those estimated using the FEV₁ data. The FEV₁/FVC ratio was consistently high, and its growth velocity was not significantly different from 0 over the age range studied.

The spatial relationships between lung function, SG (assessed by height), and M are also apparent in the plot of the fitted spline curves (Fig. 4). Here, the SG and maturation curves were adjusted using the constants estimated from equation 1. For both male and female normal subjects (Fig. 4 A and B), SG tracks closely with FEV₁ during the younger ages and appears to uncouple at the older ages. During this latter period of functional development, M contributes as much as one liter to functional growth for male subjects; for females the contribution is slightly less. However, the proportion of FEV₁ contributed during this period by maturation is similar for both sexes (25% males and 26% females). The temporal relationships between the curves are still evident, but are not as pronounced as they are in the growth velocity (Fig. 2) and vector plots (Fig. 3).

The smoothing spline procedure, in addition to yielding an optimal model through the data and first derivatives, also com-



Fig. 2. Growth velocity curves adjusted using equation 1 as described in text, for normal male (A) and female (B) subjects. The linear coefficient estimates were k1 = 0.109 and k2 = 4.24 for males and k1 = 0.158 and k2 = 3.66 for females.

Table 2. *FEV*₁: *GVP and age at GVP for normal subjects*

	Males		Females		
	GVP (L/y) (SEM)	Age at GVP (y)	GVP (L/y) (SEM)	Age at GVP (y)	
dSG	0.342	13.50	0.217	11.25	
$dFEV_1$	0.449 (0.053)	14.25	0.316 (0.040)	11.83	
dM	0.192 (0.037)	15.21	0.110 (0.027)	12.80	

putes 95% confidence intervals for the spline curve and corresponding derivatives. These confidence intervals were not included in the raw data plots (Fig. 1) for clarity purposes. However, the derivative confidence intervals were used to determine the age at which functional growth ceases (*i.e.* dFEV₁ or dFVC not significantly different from 0). For normal male subjects, the age at which the growth rate can no longer be distinguished from 0 occurred at 18.75 y, and for normal female subjects it was 15.85 y. These ages were the same for dFEV₁ and dFVC in male and female groups. As Detels *et al.* (2) point out, the growth velocity slope in this region is very small, thus making the starting age of functional decline difficult to determine accurately.

DISCUSSION

The lag time estimates between lung function and somatic GVP for our subjects are similar to those reported by Detels *et*



Fig. 3. Vector representation of normal subjects GVP. The length of each vector is proportional to the adjusted GVP amplitude, and the angle of rotation is proportional to the age at which the GVP occurred. Normal subjects' dFEV₁ vectors (males and females) are arbitrarily set to 45°. The actual ages at which the GVP occurred can be read by extending the vectors to the appropriate age axis (male or female).

Table 3. FVC: GVP and age at GVP for normal subjects

	Males		Females	
	GVP (L/y) (SEM)	Age at GVP (y)	GVP (L/y) (SEM)	Age at GVP (y)
dSG	0.452	13.50	0.304	11.25
dFVC	0.539	14.17	0.325	11.75
dM	(0.059) 0.120	15.58	(0.021) 0.030	12.88
	(0.025)		(0.005)	

al. (2). However, in their report, the ages at which the GVP occurred were consistently less then those estimated for our subjects (Table 4). There may be several reasons for the discrepancy. First, Detels et al. had only two pulmonary function testing points/subject, which were 5 y apart. Thus, each subject's velocity estimate was arbitrarily set to the age at midpoint of the 5-y followup. These interpolated values were then averaged at yearly intervals and fit using nonlinear least squares regression techniques. This interpolating and averaging scheme induces uncertainty in both the growth velocity and age data. Second, their lack of data in the younger age groups make the GVP determination difficult (see Fig. 3 in Ref. 2). Perhaps this may also explain the apparent discrepancy in their FEV_1 growth velocity curves, where female subjects actually exhibit larger GVP than their male counterparts. This discrepancy appears to be the result of an exceptionally low FEV_1 GVP for their male subjects, as the female FEV₁ GVP agrees well with our results for normal females (Table 4).

The observation that maximal lung maturation effects lag behind somatic and maximal functional (FEV₁) growth is consistent with the findings of other investigators. Tanner *et al.* (18) reported that in general the GVP for thorax width and height lag behind SG peak by about 6 and 12 mo, respectively. This suggests that at least some portion of the component we termed M may, in fact, be the indirect result of this continued thoracic growth. To investigate this proposal, one would like to examine the 320



Fig. 4. Adjusted spline curves, scaled using growth velocity constants k1 and k2 from equation 1, for normal male (*A*) and female (*B*) subjects.

Table 4. Study comparisons: GVP and age at GVP

	UCLA*			Tucson				
	Males		Females		Males		Females	
	GVP (L/y)	Age (y)	GVP (L/y)	Age (y)	GVP (L/y)	Age (y)	GVP (L/y)	Age (y)
dFEV ₁ dFVC dSG	0.260 0.420	13.0 13.0 12.0	0.360 0.280	11.0 11.0 10.0	0.449 0.539	14.25 14.17 13.40	0.316 0.325	11.83 11.75 11.25

* Detels et al. (2).

temporal relationships between the GVP associated with thorax width and height. Unfortunately, thoracic measurements were not obtained on these subjects.

Additional evidence supporting the proposal of delayed lung growth was reported recently by Griscom *et al.* (19). In their study, they measured the lengths, anteroposterior diameters, transverse diameters and cross-sectional areas and contained vol of the tracheas of 130 subjects whose ages were <20 y. They concluded that there were no differences between boys and girls until age 14, when the tracheas of girls stopped growing. Their data also indicate that male tracheas continue to enlarge in vol, but not in length, after cessation of growth in height. This tracheal growth is a process indicative of the airway growth that one would expect to accompany the increase in lung vol associated with the maturation component.

We chose to investigate FEV_1 because of its reproducibility (20, 21) and to avoid the possible effects of chest wall limitation on the complete expiration of the vital capacity. It is important to note, however, that FVC behaved essentially the same as FEV_1 (Table 3). Indeed, the FEV_1/FVC ratio was consistently high, and its growth velocity was not significantly different from 0 over the age range studied. This suggests that in normal children, the GVP seen in FEV_1 is functionally due to a GVP in vital capacity or lung vol, rather than to an increase in size-corrected

forced expiratory flow. The continued increase in lung vol was probably due to continued thoracic (i.e. chest wall) growth after increase in stature had ceased. Increased muscle force after cessation of growth in stature may have led to further thoracic growth. The effect of increased strength, however, could only have been mediated through the progressive increase of thoracic and lung vol with growth. At lung vol near TLC, lung compliance drops precipitously; hence only an extreme increase in transpulmonary pressure could have caused a further increase in FVC. Similarly, it is doubtful that increased strength led to a more complete forced exhalation with a lower residual vol. If this were the case, one would predict that the FEV1/FVC would decrease during the GVP in FVC, and this was not observed. Thus, it is unlikely that the increased muscle strength could have acutely increased FVC at the time of testing; rather, any effect of increased muscle mass with age would had to be mediated through growth.

In this study we have demonstrated that a nonparametric regression technique (*i.e.* polynomial smoothing spline) can be used for modeling the inherently noisy pulmonary function and somatic growth data. With this method, we could discern genderrelated differences in GVP for SG, FEV1, and M. This procedure has several advantages over more conventional modeling methods such as polynomial regression and exponential models. First, this technique does not require assumptions concerning the underlying structure of the data. This is important as making assumptions concerning the "true" model can influence the adequacy or ability of the model to fit the data and also places constraints on the form of the process derivatives. Second, a recent computer simulation study (22) has shown the polynominal smoothing spline to be more sensitive than regression techniques for detecting small slope changes. Without this technique, estimation of the growth velocity curves for maturation data, where the signal to noise ratio was exceptionally high, would have been most difficult. In addition, since smoothing procedures automatically yield continuous first derivatives (10), they are well suited for growth studies where growth velocity is usually of interest.

REFERENCES

- Hurwitz S, Allen J, Liben A, Becklake MR 1980 Lung function in young adults: evidence for differences in the chronological age at which various functions start to decline. Thorax 35:615-619
- Detels R, Sayre JW, Tashkin DP, Massey JR, Coulson AH, Rokaw SN 1984 The UCLA population studies of chronic obstructive respiratory disease. VI Relationship of physiologic factors to rate of change in forced expiratory volume in one second and forced vital capacity. Am Rev Respir Dis 129:533– 537
- Burrows B, Cline MG, Knudson RJ, Taussig LM, Lebowitz MD 1983 A descriptive analysis of the growth and decline of the FVC and FEV₁. Chest 83:717-724
- Degroodt EG, Van Pelt W, Borsboom G, Quanjer PH, Van Zomeren BC 1988 Growth of lung and thorax dimensions during the pubertal growth spurt. Eur J Respir Dis 1:102-108
- Schrader PC, Quanjer PH, Van Zomeren BC, Wise ME 1984 Changes in the FEV₁-height relationship during pubertal growth. Bull Eur Physiopathol Respir 20:381–388
- Schrader PC, Quanjer PH, Borsboom G, Wise ME 1984 Evaluating lung function and antropometric growth data in a longitudinal study on adolescents. Hum Biol 56:365-381
- Knudson RJ, Slatin R, Lebowitz MD, Burrows B 1976 The maximum expiratory flow-volume curve: normal standards, variability, and effects of age. Am Rev Respir Dis 113:587-600
- Simon G, Reid L, Tanner JM, Goldstein H, Benjamin B 1972 Growth of radiologically determined heart diameter, lung width, and lung length from 5-19 years, with standards for clinical use. Arch Dis Child 47:373-381
- Dockery DW, Berkey CS, Ware JH, Speizer FE, Ferris JR 1983 Distribution of forced vital capacity and forced expiratory volume in one second in children 6 to 11 years of age. Am Rev Respir Dis 128:405-412
- Wade TD, Anderson SJ, Bondy J, Ramadevi VA, Jones RH, Swanson GD 1988 Using smoothing splines to make inferences about the shape of gasexchange curves. Comput Biomed Res 21:16-26
- 11. Wahba G 1978 Improper priors, spline smoothing and the problem of guarding against model errors in regression. J R Statist Soc 40:364-372
- Eubank RL 1988 Spline Smoothing and Non-Parametric Regression, vol. 90. Statistics: Textbooks and Monographs. Marcel Dekker Inc, New York

- 13. Lebowitz MD, Knudson RJ, Burrows B 1975 The Tucson epidemiological study of chronic obstructive lung disease: I: methodology and prevalence of disease. Am J Epidemiol 102:137-152
- Kalman RE, Bucy RS 1961 New results in linear prediction and filtering theory. Trans Am Soc Mech Eng Ser D (J Basic Eng) 83:95-107
 Draper NR, Smith H 1966 Applied regression analysis. John Wiley and Sons, Inc, New York
- 16. Graybill FA 1976 Theory and Application of the Linear Model. Duxbury Press, North Scituate, MA
- 17. Kolman B 1970 Elementary Linear Algebra. Macmillan Publishing Co Inc, New York
- 18. Tanner JM 1978 Growth at Adolescence, 2nd ed. Blackwell, Oxford 19. Griscom NT, Wohl MB 1986 Dimensions of the growing trachea related to
- age and gender. AJR 146:233-237 20. Lebowitz MD, Quackenboss J, Camilli AE, Bronnimann D, Holberg CJ, Boyer
- B 1988 The epidemiological importance of intraindividual changes in objective pulmonary responses. Eur J Epidemiol 3:390–398
 Pennock BE, Rodgers RM, McCaffree DR 1981 Changes in measured spirometric indices. Chest 80:97–99
- Anderson SJ 1987 Smoothing polynomial splines with applications to modeling univariate, bivariate, and growth curve data. PhD thesis, University of Colorado Health Sciences Center, Denver, CO