

Modeling Longitudinal Growth Data and Growth Percentiles

with Polynomial Gompertz Model in SAS® Software

Inna Perevozskaya, University of Maryland Baltimore County, Baltimore, MD

Olga M. Kuznetsova, Merck & Co., Inc., Rahway, NJ

ABSTRACT

Polynomial Gompertz growth model, that models the logarithm of the relative growth rate with a polynomial, provides a good fit to individual growth curves and growth percentiles. Since the expression of a Polynomial Gompertz growth function contains an integral, this nonlinear model cannot be fitted by SAS Procedure NLIN.

At present this model can be fitted to growth data by advanced nonlinear optimization subroutines within SAS/IML®. This paper describes a macro that uses NLPLM subroutine to fit a Polynomial Gompertz model. The macro also calculates the estimates of important growth characteristics such as final adult height (using QUAD call for numerical integration) and peak growth velocity, age at peak growth velocity, age and height at adolescent growth spurt takeoff etc. (using SAS/IML NLPNRR subroutine). Good initial estimates are essential for convergence and are calculated by approximating the relative growth rate (in given examples, through SAS/IML SPLINEC and SPLINEV calls) and fitting its logarithm with the polynomial of degree K. The macro outputs the values of predicted growth curve, velocity curve, relative growth rate curve and residuals ready to be plotted with SAS/GRAPH®. Examples of using the macro to fit individual height data and NCHS height, weight and length percentiles are given.

KEY WORDS: Growth model, relative growth rate, Polynomial Gompertz, NLPLM subroutine, NLPNRR, numerical integration, final adult height, peak growth velocity.

INTRODUCTION

Parametric modeling of human growth makes it possible to summarize the complex growth pattern in a reasonable number of parameters and offers numerous analytical advantages. That is why the field of growth modeling attracted so much attention [1 - 5], and many models that describe the human growth over a specified age segment or over the life span (from birth to adulthood) were developed and applied to growth data description and analysis. However, the availability of software that allows one to fit a model was always a limiting factor in the selection of a parametric family. Sometimes a model that has a simple foundation would not receive a wide use due to the absence of an easy way to fit the model to growth data.

That was the case with the Polynomial Gompertz model, a model that models the logarithm of the relative growth rate by a polynomial of a specified degree K. The third degree model was described and fit using a stochastic approach to the pre-adolescent height data as far back as in 1979 by Sandland and McGilchrist [6] (though the name "Polynomial Gompertz" comes from later works in modeling survival with cure by Cantor and Shuster [7]). Although the model provides a good fit for individual growth data (Figure 1) and growth percentiles (Figures 4-6), it has not become wide spread and was not thoroughly studied. The reason for that is that the expression for the Polynomial Gompertz growth function contains an integral, which complicates fitting the model. In particular, this model cannot be fit by SAS Procedure NLIN.

To facilitate the use of the Polynomial Gompertz growth model, the authors offer a SAS macro that fits the model to growth data using SAS/IML nonlinear optimization subroutine NLPLM.

POLYNOMIAL GOMPERTZ GROWTH MODEL

The relative growth rate is as fundamental a concept in growth analysis as hazard is in survival analysis. It is defined as the ratio of the growth rate dy/dx to achieved growth y :

$$r(x) = \frac{1}{y} \frac{dy}{dx} = \frac{d \ln y}{dx}$$

The k-th degree Polynomial Gompertz growth model is defined as a model where the logarithm of its relative growth rate is a polynomial of the k-th degree:

$$\ln r(x) = \sum_{i=0}^k b_i x^i$$

Then the expression for the k-th degree Polynomial Gompertz growth function is

$$y(x) = A_0 \exp\left(\int_0^x \exp\left(\sum_{i=0}^k b_i u^i\right) du\right) \tag{1}$$

The ordinary Gompertz function is the 1st-degree Polynomial Gompertz: the logarithm of its relative growth rate is a linear function of time. The usefulness of the Gompertz model in description of fetal and early infancy human growth is well demonstrated in works of A.K. Laird [5]. Recently, this model was applied to individual weight, length and head circumference data from term and pre-term infants [8]. The Gompertz function with a constant added to it was fit by J. Deming [9] to individual height during the adolescent cycle of growth.

For $k=1$ the integral in (1) can be expressed in a closed form, which is not the case for $k>1$. (For $k=2$ and $b_2<0$, the integral can be expressed through the normal probability function.) That makes it difficult to fit the Polynomial Gompertz function to growth data directly for $k\geq 3$. (Sandland and McGilchrist [6] fitted the 3rd degree polynomial to an approximated relative growth rate from birth to age 10.) However, although for $k\geq 3$ the ordinary least squares (OLS) fit cannot be obtained from SAS Procedure NLIN, such fit can now be obtained through SAS/IML non-linear optimization subroutine NLPLM.

The Polynomial Gompertz function provides a good fit to the complicated pattern of human growth. The human growth velocity (Figure 2) is high at birth, rapidly decelerates in infancy, slightly declines during a long period of juvenile growth, shoots up in adolescence and declines to 0 while the growth curve approaches final adult height. The pattern of the logarithm of the relative growth rate is simpler: from birth to adulthood it seems to be adequately modeled by a polynomial of the 5th degree (Figure 3). The Figures 1, 2 and 3 represent the 5th degree PG curve, fit to the famous 1759-1777 height data of the son of the Count de Montbeillard quoted by Scammon [10], and corresponding growth velocity and logarithm of the relative growth rate curves.

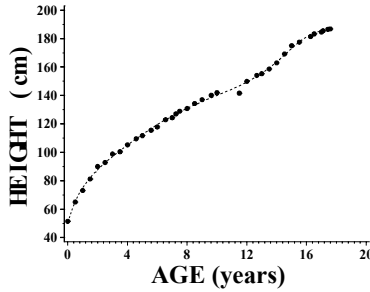


Figure 1. The 5th degree Polynomial Gompertz Model Fit to the Height Data of the Son of the Count de Montbeillard

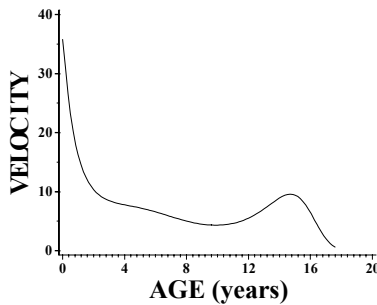


Figure 2. Height Velocity Curve Derived from the 5th degree PG fit to the Height Data of the Son of the Count de Montbeillard

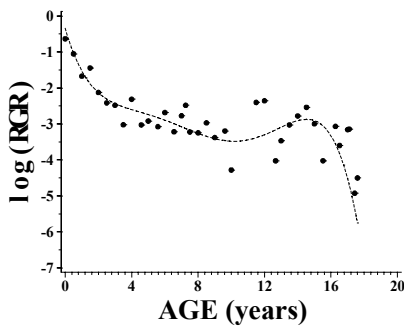


Figure 3. Approximated Logarithm of the Relative Growth Rate and the Logarithm of the Relative Growth Rate Curve Derived from the 5th degree PG fit to the Height Data of the Son of the Count de Montbeillard

The Polynomial Gompertz model gives a good fit to the length, height, weight and head circumference NCHS percentile curves. These percentile curves (the growth charts) [11] are routinely used by pediatricians to monitor growth. When height, length, weight or head circumference are studied in clinical trials, the individual growth data and population curves are often plotted against the percentiles. Also, the percentage of subjects below 3rd (or 5th) height or weight percentile is often used to compare the frequency of growth faltering in the treatment groups. Both tasks are greatly simplified if analytic expressions for percentile curves are available. As Figures 4, 5 and 6 show, Polynomial Gompertz family is a good candidate to model percentiles.

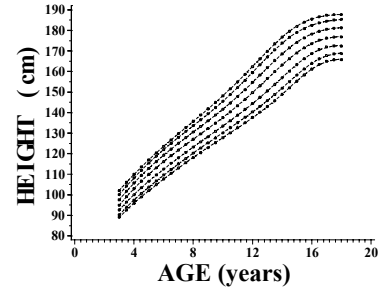


Figure 4. Height by Age 5th, 10th, 25th, 50th, 75th, 90th and 95th Reference Percentiles for Boys 3-18 Years Old (circles) and Fitted Polynomial Gompertz Curves (dashed lines)

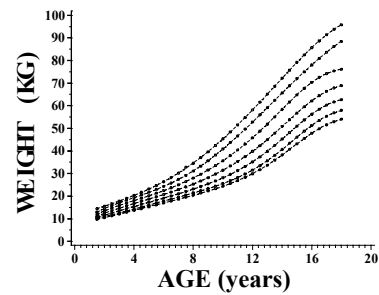


Figure 5. Weight by Age 5th, 10th, 25th, 50th, 75th, 90th and 95th Reference Percentiles for Boys 1.5-18 Years Old (circles) and Fitted Polynomial Gompertz Curves (dashed lines)

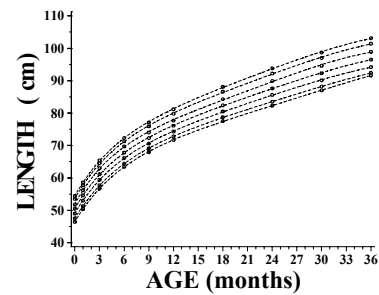


Figure 6. Length by Age 5th, 10th, 25th, 50th, 75th, 90th and 95th Reference Percentiles for Boys 0-36 Months Old (circles) and Fitted Polynomial Gompertz Curves (dashed lines)

Figure 4 represents seven height percentile curves for 3-18 years old boys as tabulated in [11] (circles) overlaid with the 5th degree PG curves fitted to each percentile tabulated data (dashed lines). It can be seen that the tabulated data are fit by the PG curves very closely. Figure 5 demonstrates the good fit of the 5th degree PG curves to the tabulated weight percentile curves for 1.5-18 years old boys. Figure 6 shows that the 3rd degree PG provides sufficiently good fit for the tabulated length percentiles for boys over 0-36 months interval.

MACRO %PGOLS DESCRIPTION

The SAS macro %PGOLS fits the Kth degree Polynomial Gompertz model to specified growth data &INDATA. It uses the Levenberg-Marquardt method implemented in SAS/IML subroutine NLPLM to find an ordinary least squares fit. The data set with the growth data consists of several observations with

values of AGE (an age variable) and &RESPVAR (a response variable: for example, height in cm). Good initial estimates of the parameters A_0, b_0, \dots, b_k are essential for convergence and are calculated prior to %PGOLS call by approximating the relative growth rate (in the given examples, through SAS/IML SPLINEC and SPLINEV calls) and fitting its logarithm with the polynomial of degree K. The values of initial estimates are then stored in the data set &PARM_INI.

The macro outputs a one-observation data set &OUTPARMS with the parameter estimates A_0, b_0, \dots, b_k . If the option to calculate the clinically important parameters of the growth curve is selected through DO_CLIN=1 in the macro %PGOLS call, the estimates of these parameters are added to the data set &OUTPARMS. The list of the clinical parameters currently includes: Final Adult Height (HF), Age at Peak Growth Velocity (AVP), Peak Growth Velocity (VP), Age at Adolescent Growth Spurt Takeoff (defined as the age of minimum growth velocity, AVL), Minimum Growth Velocity (VL), Height at Adolescent Growth Spurt Takeoff (HAVL). The full list of these characteristics can be obtained only for subjects observed throughout the adolescent growth spurt and is not relevant for pre-adolescent children. The macro default is DO_CLIN=0 (no calculation of the clinical parameters).

It is often of particular interest to interpolate growth at specified reference ages within the range of ages the measurements were taken. For example, in infants studies the exact ages at projected visits (2, 4, 8, 12 and 16 weeks) might form a set of such ages. The macro allows to calculate the values of fitted PG function at specified &REFNUM (up to 10) number of reference ages by assigning the ages to the macro variables REFAGE1, ... up to REFAGE10. The values of PG function FAGE1, ..., up to FAGE&REFNUM are then added to the data set &OUTPARM. The macro default is REFNUM=0 (no reference ages specified).

The scale coefficient in PG model does not have to be A_0 - a height at age 0 as in (1). When the individual measurements start at an age other than 0, it is convenient to express the function PG(x) in a form, where the integral is taken over the interval from the age of the first measurement AGE1 to x:

$$y(x) = A_F \exp\left(\int_{AGE1}^x \exp\left(\sum_{i=0}^k b_i u^i\right) du\right) \quad (2)$$

In this notation the scale coefficient A_F represents height at age AGE1.

When the model is fitted to several subjects data and the ages of the first measurement differ among the subjects, it is convenient to choose some common lower border of the integral STARTAGE for all the subjects. Then the PG(x) can be expressed as

$$y(x) = A_S \exp\left(\int_{STARTAGE}^x \exp\left(\sum_{i=0}^k b_i u^i\right) du\right), \quad (3)$$

where A_S is a height at age STARTAGE. Then

$$A_S = A_F \exp\left(\int_{STARTAGE}^{AGE1} \exp\left(\sum_{i=0}^k b_i u^i\right) du\right)$$

The macro calculates both AS and AF and places them in the data set &OUTPARM; the macro variable STARTAGE needs to be specified externally.

The macro also outputs the ready-to-plot data set &OUTPREDS with the variables AGE, PGOM (value of the PG fit), VELO (growth velocity), RGRATE (relative growth rate), POL (logarithm of the relative growth rate), &RESPVAR and the residual DIF (= &RESPVAR-PGOM) (ages are taken from the original data set &INDATA). If the ages are too far apart, the age grid for derived variables PGOM, VELO, RGRATE and POL can be refined for plotting by breaking each age interval into &AGEPART segments (the default value for AGEPART is 1).

The macro is written for the SAS System for UNIX (Release

6.12), but it can be used with the SAS System for PC as well. The SAS/IML NLPNRR subroutine is used to calculate ages at peak and minimum velocity. The SAS/IML QUAD calls are used for numerical evaluation of finite and infinite integrals. The SAS code that deals with the clinical parameters is placed in a macro %CLINPARM outside of macro %PGOLS to improve the readability of the code.

The macro re-scales the polynomial coefficients for NLPLM to improve convergence: without the re-scaling the gradients differ a lot in magnitude. The same re-scaling scheme used in the macro proved to be good for all longitudinal data tested, including the percentiles data. With the re-scaling, the optimization procedure demonstrated a very good convergence rate and required about 10 iterations.

EXAMPLES

The 5th degree PG fit for the height data of the son of the Count de Montbeillard, pictured in Figure 1, was obtained by the following macro call:

```
%let degree=5;
%let startage =0;
%PGOLS(indata= data.mont, respvar= htcm,
resplab=HEIGHT, outparm=olsparms, parm_ini=
mont_ini, outpreds=monpreds, do_clin=1,
refnum=10, refage1=1.0, refage2=2.0,
refage3=5.0, refage4=8.0,refage5=10.0,
refage6=11.0, refage7=12.0, refage8=13.0,
refage9=15.0, refage10=17.0,
agepart=1);
```

The 1-observation output data set OLSPARMS with the parameter estimates, clinical parameters and heights calculated at 10 reference ages is printed below (notice that AF=AS in this example, since STARTAGE and the first age are equal (both are 0):

OBS	B0	B1	B2	B3
1	-0.35130	-1.57177	0.45762	-0.068157
OBS	B4	B5	AF: HEIGHT at Age of 0	AS: HEIGHT at Age of 0.00
1	.0046446	-.00011463	50.7816	50.7816
OBS	Age at Peak Growth Velocity	HEIGHT Velocity at Growth Spurt Takeoff	Peak HEIGHT Velocity	HEIGHT at Adolescent Growth Spurt Takeoff
1	14.7137	4.30365	9.55298	139.503
OBS	Age at Adolescent Growth Spurt Takeoff	Final Adult HEIGHT	HEIGHT at Age 1.0	HEIGHT at Age 2.0
1	9.89544	186.198	74.9788	87.6964
OBS	HEIGHT at Age 5.0	HEIGHT at Age 8.0	HEIGHT at Age 10.0	HEIGHT at Age 11.0
1	112.460	130.891	139.953	144.363
OBS	HEIGHT at Age 12.0	HEIGHT at Age 13.0	HEIGHT at Age 15.0	HEIGHT at Age 17.0
1	149.351	155.557	172.912	185.285

The height plot for this subject (Figure 1) was obtained by PROC GPLOT of SAS/GRAPH with the following statements:

```
proc gplot data=monpreds;
plot htc*age pgom*age /overlay;
```

The velocity plot for the same subject (Figure 2) was obtained with:

```
proc gplot data=monpreds;
plot velo*age;
```

The 5th Height by Age Percentile for boys 3-18 years old pictured in Figure 4 was fitted by the 5th degree PG model by the following macro call:

```
%let degree=5;
%let startage =2;
%PGOLS(indata= data.hb318p, respvar= p5,
resplab=HEIGHT, outparm=hbp5prms, parm_ini=
hbp5_ini, outpreds=hbp5prds);
```

Similar calls were used to fit the other 6 percentile curves.

The 5th Length by Age Percentile for boys 0-36 months old pictured in Figure 6 was fitted by the 3rd degree PG model by the following macro call:

```
%let degree=3;
%let startage =0;
%PGOLS(indata = data.lb036p, respvar = p5,
resplab=LENGTH, outparm=lbp5prms, parm_ini =
lbp5_ini, outpreds= lbp5prds);
```

MACRO CODE

```
%macro PGOLS(indata=, respvar=, resplab=,
parm_ini=, outparm=, outpreds=, do_clin=0,
refnum=0, refage1=, refage2=, refage3=,
refage4=, refage5=, refage6=, refage7=,
refage8=, refage9=, refage10=, agepart=1);
```

/* This macro provides the OLS fit for the polynomial Gompertz model. The degree K of a polynomial and a common low border of the integral are specified by macro variables DEGREE and STARTAGE outside of macro.

REQUIRED parameters:
indata = name of the data set with growth data (AGE and &RESPVAR);
parm_ini = name of the data set with the initial values of coefficients;
respvar = name of the response variable in &indata;
resplab = label of the response variable;
outparm = name of the one-observation data set with the estimates of the parameters AF, AS,b0, b1, . . . , bk, and clinical characteristics (also: values of the fit curve at specified reference ages if requested);
outpreds = name of the ready-for-plot data set with the values of the PG fit, velocity, relative growth rate, etc. calculated at ages from &indata (age grid is finer if &agepart>1);

OPTIONAL parameters:
do_clin = 1, to calculate clinical parameters,
0, to skip clinical parameters calculation (default);
refnum = number of reference ages (default=0);
refage1-refage10 = reference ages (a total of &refnum) to calculate the growth curve at;
agepart = number of segments each interval between two consecutive ages from &indata is broken into for finer age grid in outpreds;*/
* #####;
proc iml;
/* Functions expol, PG, and object are used by main optimization routine */

```
start expol(u) global( beta_sc, scale, ageone);
if ncol(beta_sc)=1 then beta_sc=t(beta_sc);
b=beta_sc[,1:(&degree+1)];
sum=0;
do i=1 to (&degree+1);
if (u=0)&(i=1) then tmp2=1;
else tmp2=u##(i-1);
sum=sum+(b[i]/scale[i])*tmp2;
end;
if sum > 800 then print "ERROR IN SUM:"
beta_sc u sum;
return(exp(sum));
finish expol;
```

```
start PG(x) global (beta_sc, scale,ageone);
A=beta_sc[&degree+2];
pts={0 0};pts[1]=ageone; pts[2]=x;
r=0;
if x>ageone then do;
call quad(r,"expol", pts) ;
if r>900 then print "ERROR IN R" beta_sc r;
end;
if x<ageone then do;
pts[1]=x;pts[2]=ageone;
call quad(rr,"expol", pts) ;
if rr>900 then print "ERROR IN R" beta_sc rr;
r=-rr;
end;
return(A*exp(r));
finish PG;
```

```
start object(beta_obj) global ( age , resp,
N_OBS,beta_sc,scale,ageone);
beta_sc=beta_obj;
f=do(1,N_OBS,1);
do j=1 to N_OBS;
f[j]=resp[j]-PG(age[j]);
end;
return(f);
finish object;
```

```
start adjscale(A_AGEF, A_SAGE) global(ageone);
/*Calculates the scale parameter A_SAGE when the low limit of integral is set to &STARTAGE. A_AGEF is a scale coefficient when the lower limit of an integral is the lowest age AGEONE */
A_SAGE=A_AGEF;
pts={&STARTAGE 0};pts[2]=ageone;
if (pts[1] < pts[2]) then do;
call quad(r, "expol", pts) ;
A_SAGE=A_AGEF/exp(r);
end;
if (pts[1] > pts[2]) then do;
pts[1]=ageone;pts[2]=&startage;
call quad(r, "expol", pts) ;
A_SAGE=A_AGEF/exp(-r);
end;
finish adjscale;
```

```
#### The main part: OLS fit of PG model starts;
use &indata;
read all var{age} into age;
read all var{&respvar} into resp;
close &indata;
tage=t(age); ageone=tage[1];
```

```
use &parm_ini;
read all var _num_ into beta;
close &parm_ini;
```

```
scale=J(1,&degree+2,1); * scaling parameters;
scale[3]=10;
%if &degree=5 %then %do i=4 %to 6;
scale[&i]=10##(&i-3); %end;
%if &degree=3 %then %do i=4 %to 4;
scale[&i]=10##(&i-1); %end;
beta_sc=beta#scale; /* beta_sc contains scaled parameters */
```

```
N_OBS=nrow(resp);
```

```

beta_obj=beta_sc;
optn={50, 5};
optn[1]=N_OBS;
tcl={., ., ., 0, 0, 1e-3, 0, 0};
x0=beta_sc;
call nlpml(rc,xr,"object", x0,optn) tc=tcl;
beta_sc=xr; /* beta_sc has OLS estimates now */
beta=beta_sc/scale; /* unscaled parameters */

create one from beta;
append from beta; /* data set with betas */

create optim var {rc};
append; /* data set with return code */

call adjscale( beta[&degree+2], AS);
create startage var{AS};
append; /* data set with As */

%clinparm;
%if &do_clin=1 %then %do;call clinical;%end;

%if &refnum>0 %then %do;
  fage=J(1,&refnum,1);
  %do k=1 %to &refnum; fage[&k]=PG(&&refage&k);
  %end;
  create dat_fage from fage;
  append from fage; /* PG at reference ages */
%end;

/* generating the data set &outpreds with values
of age, response variable, PG fit, velocity, log
of relative growth rate and residuals */

dim_y=&agepart*(N_OBS-1)+1;
y=J(1,dim_y,1);
fineage=J(1,dim_y,1);
do i=1 to (N_OBS);
  if (i<N_OBS) then do;
    step=(age[i+1]-age[i])/&agepart;
    do j=1 to &agepart;
      index=&agepart*(i-1)+j;
      *print index;
      fineage[index]=age[i]+step*(j-1);
      y[index]=.;
    end;
  end;
  fineage[dim_y]=age[N_OBS];
  y[(i-1)*&agepart+1]=resp[i];
end; /* fine age grid is created */
pgom=y;
velo=y;
pol=y;
do i=1 to dim_y;
  pgom[i]=PG(fineage[i]);
  pol[i]=log(expol(fineage[i]));
  velo[i]=velocity(fineage[i]);
end;
dif=y-pgom;
create &outpreds var{y fineage pgom pol velo
dif};
append;
quit; * exiting Proc IML;

/* Combining the data sets for output */
proc print data=&outpreds;

data &outpreds;set &outpreds
(rename=(fineage=age));
rgrate=exp(pol);
run;

data two;set one; /* renames columns in data set
one to b0, b1, ..., AF */
%do ci=1 %to &degree+1;
  %let ind=%eval(&ci-1);
  rename COL&ci=b&ind;
%end;
%let deg2=%eval(&degree+2);
rename COL&deg2=AF;

run;
%if &refnum>0 %then %do;
data three; set dat_fage; /* renames columns in
dat_fage to fage1, fage2,... */
%do k=1 %to &refnum; rename COL&k=fage&k; %end;
run;
%end;

data &outparm; /*concatenates the data sets*/
set optim;set two;set startage;
%if &do_clin>0 %then %do;set dat_clin;%end;
%if &refnum>0 %then %do;set three; %end;
label rc="Return Code"
AS="AS: &resplab at Age of &STARTAGE"
AF="AF: &resplab at Age of &AGEF"
%if &do_clin>0 %then %do;
AVL="Age at Adolescent Growth Spurt Takeoff"
AVP="Age at Peak Growth Velocity"
VL="&resplab Velocity at Growth Spurt Takeoff"
VP="Peak &resplab Velocity"
HAVL="&resplab at Adolescent Growth Spurt
Takeoff"
HF="Final Adult &resplab"
%end;
%if &refnum>0 %then %do i=1 %to &refnum;
PAGE&i="&resplab at Age &&refage&i "
%end;;
run;
proc print data=&outparm label;title 'outparm';
%mend PGOLS; * #####;

%macro clinparm; * #####;
/* A macro with the definitions of the functions
called by module CLINICAL to calculate clinical
parameters */

start velocity(x) global (beta_sc,scale,
ageone);
  vel=PG(x)*expol(x);
  return(vel);
  finish velocity;

start gradient(x) global (beta_sc,scale,ageone);
  if ncol(beta_sc)=1 then beta_sc=t(beta_sc);
  b=beta_sc[,1:(&degree+1)]; sum0=0; do i=1 to
(&degree+1);
  sum0=sum0+(b[i]/scale[i])*x##(i-1); end;
  sum1=0;
  do i =1 to &degree;
    sum1=sum1+i*(b[i+1]/scale[i+1])*x**(i-1);
  end;
  tmp=exp(sum0)*(velocity(x)+PG(x)*sum1);
  return(tmp);
  finish gradient;

start minmax(argmin, argmax) global(age, N_OBS);
/* This subroutine computes initial guesses for
the latest local max and min of velocity #####*/
v=do(1,N_OBS,1);
do i=1 to N_OBS;
  v[i]=velocity(age[i]);
end;
*vt=t(v); *print vt;
argmin=.;
argmax=.;
do i=2 to N_OBS-1;
  difprev=v[i]-v[i-1];
  difnext=v[i]-v[i+1];
  sign=difprev*difnext;
  if (sign>0)&(difprev>0) then argmax=age[i];
  if (sign>0)&(difprev<0) then argmin=age[i];
end;
finish minmax;

start clinical;
/*puts clinical parameters in &dat_clin */
call minmax(x0_low,x0_peak);
AVL=.;AVP=.;VL=.;VP=.;HAVL=.;
/* 1. age of lowest growth velocity */

```

```

optn={0 0};
con={0 . . ,
      . . .};
if (x0_low ^= .) then call nlpnrr(rc, AVL,
"velocity",x0_low,optn,con) grd="gradient";
/* 2. age of peak growth velocity */
optn={1 0};
con={0 . . ,
      . . .};
if X0_peak ^=. then
call nlpnrr(rc,AVP,"velocity", x0_peak,optn,con)
grd="gradient";
if AVL ^=. then do;
  VL=velocity(AVL); HAVL=PG(AVL); end;
if AVP ^=. then VP=velocity(AVP);
HF=. ;
if beta_sc[&degree+1] < 0 then do;
  points={&AGEF .P};
  call quad(r,"expol",points) eps =1e-10;
  if r<=1000 then
HF=beta_sc[&degree+2]*exp(r)/scale[&degree+2];
else print "R is too big" r;
end;

create dat_clin var {AVL AVP VL VP HAVL HF};
append;
finish clinical;
%mend clinparm; *
#####;

```

CONCLUSION

Macro %PGOLS is a convenient SAS tool that provides an OLS Polynomial Gompertz fit to growth data. Thus, this macro effectively facilitates the use of PG family in growth modeling with SAS.

By applying this macro to growth data and growth percentiles data, it was established that the 5th degree PG model provides an adequate fit to a life-span of height and weight data (or 2 to 18 years growth percentiles) and that the PG model of the 3rd degree is sufficient to describe the length, weight and head circumference (or length, weight and head circumference percentiles) from birth to 36 months of age. The 3rd degree model also provides a good fit to pre-adolescent to final adult (for example, 7 to 18 years) height and weight data. Based on shapes of growth curves for many other body measurements (for example, trunk length, a limb length), it is very likely that a PG model provides a proper fit for those as well.

The macro can be used to compare different populations (for example, boys and girls, or treatment groups) in clinically meaningful characteristics of growth curves, such as final adult height, age at peak growth velocity, peak growth velocity, age at adolescent growth spurt takeoff, minimum growth velocity and height at adolescent growth spurt takeoff. A PG curve can be fitted to each subject's data by the macro %PGOLS with clinical parameters above output and collected for summarization and populations comparison.

The macro %PGOLS can be used to fit with a PG model available tabulated smoothed or empirical growth percentiles data for easy storage and retrieval of reference percentiles for plotting and analysis.

The macro %PGOLS will be used to compare a Polynomial Gompertz model with other existing growth models. The macro can be expanded to provide a generalized least squares fit. A non-linear mixed model based on a PG model is to be explored.

The complete SAS code that produces initial estimates of the PG model parameters, then calls the macro %PGOLS to obtain the OLS estimates of the model parameters, clinically meaningful

parameters and values of PG fit at specified reference ages, and, finally, generates the plot of the growth data overlaid with a PG fit as well as the plots of velocity, relative growth rate and logarithm of the relative growth rate curves and residuals, is available from the authors upon request.

REFERENCES

1. Bock, R. D. & Thissen, D. (1976). Fitting multi-component models for growth in stature. Proceedings of the 9th International Biometric Conference, 1, 432–42.
2. Peerson, J. M., Heining, M. J., Nommsen, L. A., Lonnerdal, B., and Dewey, K. G. (1973) Use of Growth Models to Describe Patterns of Length, Weight and Head Circumference among Breast-Fed and Formula-Fed Infants: The Darling Study. Human Biology, August 1993, v.65, no 4, pp. 611-626.
3. Preece, M.A., Baines, M.J., (1978) A New family of mathematical models describing the human growth curve. Annals of Human Biology, vol. 5, no. 1, 1-24.
4. Roche, A. F. (1980a). The analysis of serial data. Studies in Physical Anthropology, 6, 71–88.
5. Laird, A.K. (1967) Evolution of the human growth curve, Growth, 1967, 31, 345-355.
6. Sandland, R. L. , and McGilchrist, C.A. (1979). Stochastic growth curve analysis. Biometrics, 35, 255-271.
7. Olga M. Kuznetsova, Alan B. Cantor, Jonathan J. Shuster (1994), A Polynomial Gompertz Model for Survival with Cure, unpublished manuscript.
8. S. Zhang, O. Kuznetsova, J. Jacobs (1998), Use of a Gompertz curve to describe patterns of early growth in term and preterm infants, Human Biology Association Meeting, Salt Lake City, Uta.
9. 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.
10. Scammon, R.E. (1927). The first seriatim study of human growth. American Journal of Physical Anthropology 10, 329-336.
11. Hamill, P. V. V., Drizd, T. A., Johnson, C. L., Reed, R. B. & Roche, A. F. (1977). NCHS growth curves for children. Birth—18 years, United States. Vital and Health Statistics, National Center for Health Statistics, Series 11, No. 165. Washington, DC: US Government Printing Office.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dr. Jonathan J. Shuster of Pediatric Oncology Group Statistical Office and University of Florida, and Dr. Alan B. Cantor of Moffitt Cancer Center and University of South Florida, for introducing one of the authors to the Polynomial Gompertz model.

CONTACT INFORMATION

For further information, please contact:

Dr. Olga Kuznetsova
Merck & Co., Inc.
P.O. Box 2000, RY33-404
Rahway, NJ 07065-0900
Phone: (732) 594-5094,
Fax: (732) 594-6075
Email: olga_kuznetsova@merck.com