SCIENTIFIC REPORTS

OPEN

A time-varying biased random walk approach to human growth

Béla Suki[1] & Urs Frey[2]

Received: 15 March 2017
Accepted: 30 June 2017
Published online: 10 August 2017

Growth and development are dominated by gene-environment interactions. Many approaches have been proposed to model growth, but most are either descriptive or describe population level phenomena. We present a random walk-based growth model capable of predicting individual height, in which the growth increments are taken from time varying distributions mimicking the bursting behaviour of observed saltatory growth. We derive analytic equations and also develop a computational model of such growth that takes into account gene-environment interactions. Using an independent prospective birth cohort study of 190 infants, we predict height at 6 years of age. In a subset of 27 subjects, we adaptively train the model to account for growth between birth and 1 year of age using a Bayesian approach. The 5-year predicted heights compare well with actual data (measured height = 0.838*predicted height + 18.3; $R^2 = 0.51$) with an average error of 3.3%. In one patient, we also exemplify how our growth prediction model can be used for the early detection of growth deficiency and the evaluation of the effectiveness of growth hormone therapy.

Growth and development are complex nonlinear biological processes[1]. The growth process is driven by gene-environment interactions allowing for heterogeneity on a population level as well as adaptability to environmental conditions[2,3]. At the level of individual organisms, defects in growth can lead to the emergence of different health and disease phenotypes[4–6].

Various models have been proposed to describe growth as a function of time including exponential, sigmoid or even more complex mathematical functions most of which are descriptive in nature[7–10]. A universal mechanistic model was developed by West *et al.*[11] in which growth is related to total metabolic rate partitioned for cellular maintenance and cellular growth. This theory predicts an exponential growth for the quarter power of body mass. The key assumptions of the West theory[11] are based on the concept of continuous growth and allometric scaling but it does not fully explain human[12] and invertebrate[13] growth nor does it account for individual variability and adaptability. To better understand variations within a population as well as adaptability of the individual, it is necessary to relate mechanisms influencing growth operating on short time scales at the microscopic level to the overall growth of an organism and its adaptations to environmental fluctuations during development.

Observations of growth pattern in human infants provided evidence that growth is not continuous, but it occurs in saltatory bursts separated by waiting periods of nearly no growth[14]. The temporal dynamics of growth pattern suggested a non-random aperiodic behaviour[15] and mechanistically, this pattern[16–19] may be related to the pulsatile nature of growth hormone expression[20–22] or cell cycle dynamics of mitotic start and stop phases[23]. Although continuous and salutatory growth may co-exist but because of the limited accuracy of measurements, it is difficult to resolve whether the two growth patterns are clearly distinct[15,24–26]. Nevertheless, there is evidence that salutatory growth exists in infancy[14] and even in school age[24] although the exact distributions of salutatory intervals and growth amplitudes are still largely unknown.

Another general limitation of most growth models is that they are unable to fit population mean and variability curves[8,19,27], or the heterogeneous behaviour of growth trajectories of individual subjects[15,26]. Furthermore, there is also no clear association between the variability of individual growth and the variability of population growth, and thus population based growth charts are not suitable to predict the variability of individual growth trajectories. Nevertheless, the latter issue may be of critical importance in clinical decision making[28]. In clinical practice, individual growth trajectories, including confidence boundaries, would be needed to detect early signs of abnormal growth due to chronic disease, or to judge the efficiency of treatment strategies such as growth

[1]Department of Biomedical Engineering, Boston University, 44 Cummington Str., Boston, Massachusetts, 02215, USA. [2]University Children's Hospital Basel, UKBB, University of Basel, Spitalstrasse, PO Box 4031, Basel, Switzerland. Correspondence and requests for materials should be addressed to B.S. (email: bsuki@bu.edu) or U.F. (email: urs.frey@ukbb.ch)

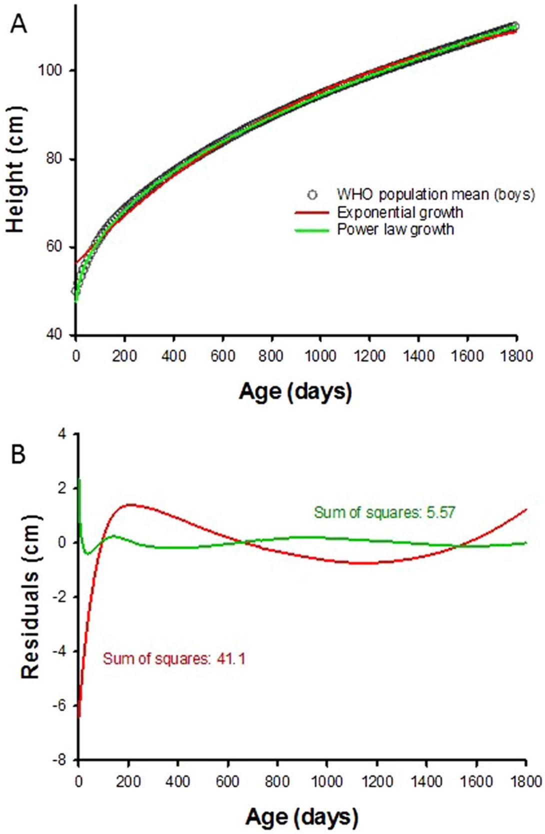

Figure 1. (**A**) Comparison of the fits of the exponential (red) and power law (green) growth models to the population mean of boy's height in the World Health Organization (WHO) data set. (**B**) Comparison of the sum of square residuals of the exponential and power law model fits. Note that the error of the exponential model is more than 7 times larger than that of the power law model.

hormone therapy. For clinical decisions in the context of precision medicine, we need individual growth prediction models including confidence boundaries rather than population based growth estimates.

In this study, we aim to link micro-scale events to overall growth pattern at the individual as well as population levels during preschool childhood development. Often referred to as a 'window of opportunity', preschool age is a critical developmental period[5,6,29], during which complex chronic diseases can arise due to interaction of genetic and environmental factors often accompanied by growth disturbance. To this end, we develop a biased random walk model of individual growth in which the growth increments are taken from time varying distributions mimicking the bursting behaviour of observed saltatory growth patterns. The model is then extended to include genetic factors as well as the effects of possible variations in environmental conditions both at the individual and population levels. We train the model based on existing World Health Organization (WHO) growth charts[27], and test its predictive ability by computing 5-year growth trajectories of individual subjects in an independent prospectively assessed longitudinal dataset from our laboratory. Finally, we exemplify the clinical utility of the modelling approach in a prematurely born small child with growth deficiency.

Results

We developed a stochastic model of biological growth based on the idea that growth occurs in discrete steps and can thus be described as a random walk. The steps of the random walk are taken from a distribution with time varying parameters with a positive mean. The model also includes a feedback system which determines when the distribution of step sizes changes. Specifically, a time window (w) is defined during which the parameters of the distribution do not change. The random walk starts with initializing the mean (μ) and variance (σ^2) of the distribution and generates w number of steps. Since the growth must approach a finite value, μ takes a smaller value after w steps gradually approaching 0. The cumulative running sum, which is also the position of the random walker, represents the growth curve. The model is solved analytically for two specific cases that describe how μ decreases in time: an exponential and a power law decrease (see Methods). Both models have 3 parameters that were adjusted to fit the mean growth curves of the WHO data set. The power law model provides a significantly better fit with a residual sum of squares 7 times smaller than the exponential model (Fig. 1). Thus, for the analysis of individual subject's growth curve, we only retain the power law model and extend it computationally as follows.

The computational model incorporates numerically a growth rate function that determines the time varying bias. Since the growth curve follows a power law with an exponent $1+\alpha$ (see Eq. 10, Methods), the growth increment, which is the derivative of the growth curve, is also a power law with an exponent α. We therefore

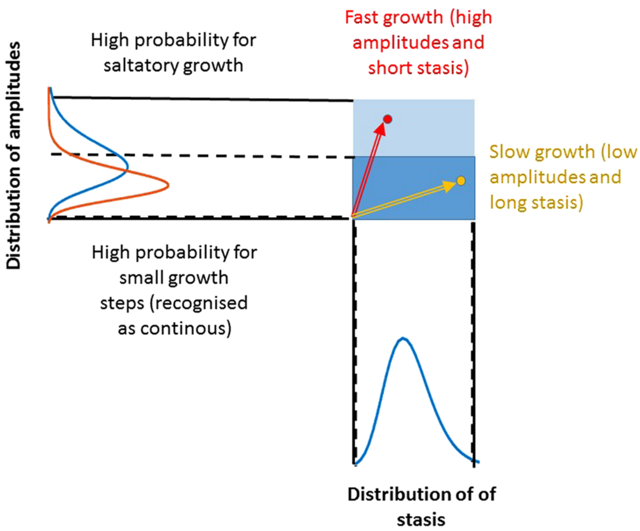

Figure 2. Schematic of the growth process modelled by the time-varying biased random walk. The distribution of stasis times is shown on the x axis by the blue lognormal curve. Two distributions of growth increment amplitudes are shown by the lognormal curves on the y axis. The blue represents a faster growth with larger increments at an younger age while the brown demonstrates a later stage of growth with smaller increments. For the stasis time, only one curve is shown as experiments do not suggest that this distribution changes with age. On this diagram, the red vector represents fast growth with short stasis time and large increments whereas the yellow vector corresponds to slow growth.

prescribe that the growth increments are taken from a lognormal distribution with a mean that decreases with time after w steps according to a power law with an exponent α which is obtained by fitting the derivative of the mean growth curve of the WHO data set. The variance of the lognormal (σ^2_{ampl}) is obtained from fitting (see below). Additionally, we introduce stochastic fluctuations in the time intervals between growth steps, the stasis times, which are also modelled as a lognormal with a given mean stasis time (T_{stasis}) and variance (σ^2_{stasis}). Figure 2 illustrates how these distributions determine the overall growth rate.

Using a grid search, we fit the mean as well as the mean ± 2SD height curves in the WHO data set with our computational model. Accordingly, we find that for boys, the optimal values for $\underline{T}_{stasis}$, σ_{stasis} and σ_{ampl} are 10 days, 0.1 days and 0.3 cm, respectively, whereas these numbers for girls are 11 days, 0.2 days, and 0.3 cm, respectively. While the model fit the mean height curve of the WHO data well, the 2SD curves were best fit if we also allowed an additional random variation in the genetic component. To do this, the value of α was randomized within ±2.5% of the value that corresponds to the height at birth for each individual run for a given subject. With these parameters, the overall error of the fit for population mean was 0.69 cm and for the ±2SD it was 0.99 cm. Figure 3 compares the simulated mean and ±2SD height growth curves with those in the WHO data set in boys (A) and girls (B).

For clinical decision making, individual growth trajectories with confidence intervals are needed. Our fully established computational time-varying biased random walk model allows us to simulate a family of random walks for a subject starting from any age if the value of α is known. To obtain α, we first carry out a regression between the α values obtained from the fitting of Eq. 10 to the WHO mean, the mean ±SD and the mean ± 2SD curves and height at 1 month of age. For any given child, this approach provides individual runs as well as an estimated mean trajectory and SD curves (Fig. 4). These trajectories are based on the assumption that the individual child statistically behaves similarly to the reference population in the training set. It can be seen that the bounds around the mean increase as a function of time. The individual trajectory is based on the height at the chosen initial age and the allowed variations correspond to the population-based variations derived from the WHO data. We next launch 500 trajectories using the height at 5 weeks of age in 190 subjects from the BILD (Bern Infant Lung Development) data set (Table 1)[29] and compare the predicted and measured heights at 72 months of age (Fig. 5). While there is no difference between the predicted and measured heights with an average error of 3.9%, the regression slope is 0.476 ($p < 0.0001$), the R^2 is only 0.253.

To improve on the prediction, we note that there are environmental influences that can also change both the initial and later growth rates. In order to allow adaption to such fluctuations, we build an adaptive correction mechanism into the model using serial observations. The difference between model prediction and measured height at any given age is likely a consequence of external influences such as feeding pattern or diseases. We account for such effects by adaptively correcting the value of α as follows. We first carry out a regression between the α values obtained from the fitting of Eq. 10 to the WHO mean, the mean ± SD and the mean ± 2SD curves and height at 1, 2, 4, 6 and 12 months of age. The procedure is the same as that described above for the initial birth height. This allows us to set up a look-up table of population α and intercept values at these time points. Next, given the height and the corresponding α at t_1 (e.g. 1 month of age), we use our model to predict the full distribution of height at the next measured age t_2 (e.g. 2 months). Comparing the measured height to the predicted

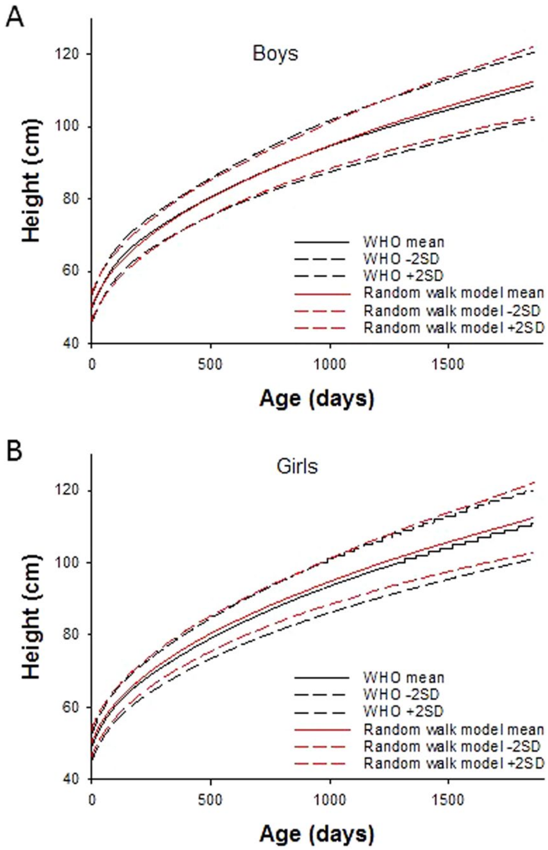

Figure 3. Simulated means and 2SDs of growth curves in comparison with the WHO training data set for boys (**A**) and girls (**B**).

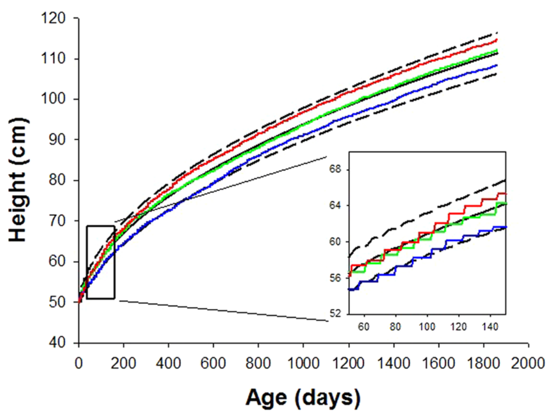

Figure 4. Simulation of a family of individual random walks (red, green and blue lines) of a given subject starting from birth. The probabilistic simulation allows us to calculate the mean (black solid line) as well as confidence intervals (+SD, black dashed line). The inset shows a magnified region of the growth with the 3 discrete random walk realizations. Note the staircase-like growth curves corresponding to the bursting behaviour.

distribution of height, we determine the Bayesian probability p_h that the measured height at t_2 is the actually observed value. This allows us to adjust the model and launch a set of new trajectories starting from t_2 as follows: (1) α is corrected by a small amount depending on the distance between the previous value, $\alpha(t_1)$, and the current value from the look-up table, $\alpha(t_2)$, as follows:

$$\alpha'(t_2) = \alpha(t_1) + c[\alpha(t_2) - \alpha(t_1)](1 - p_h)$$

	Mean	Median	SD	CI
N	190			
sex (boys/girls)	107	56.3%		
Gestational age (weeks)	39.82	40.00	1.19	0.03
Height at birth (cm)	49.57	50.00	2.06	0.04
Age at first measurement (days)	35.40	35.00	5.27	0.15
Height at first measurement (cm)	55.05	55.00	2.38	0.04
Z-Scores (SDS, WHO)	0.07	0.10	1.09	14.50
Age at second measurement (years)	6.02	6.05	0.33	0.05
Height at second measurement (cm)	117.65	118.00	5.56	0.05
Z-Scores (SDS, WHO)	0.39	0.43	1.03	2.62
Predicted height at second measurement (cm)	117.09	117.25	5.48	0.05
2 standard deviation of prediction (cm)	2.64	2.62	0.26	0.10

Table 1. Biometric data of the BILD cohort. Height at 6 years was predicted from length at 5 weeks of age ($n = 190$). SD: standard deviation, CI: confidence intervals, Z-Scores SDS: standard deviation scores from age related height based on the WHO reference values[27].

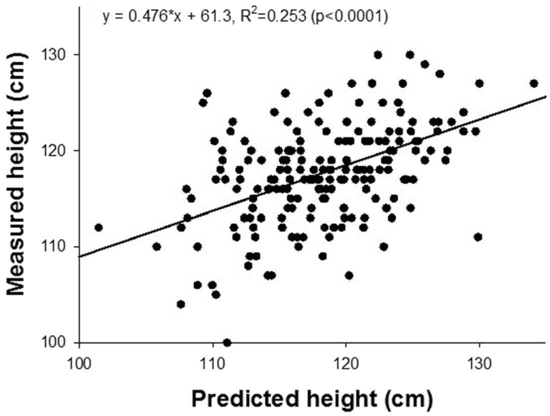

Figure 5. Measured height at 6 years of age as a function of predicted height computed from height at age of 1 month in 190 children from the BILD data set.

and (2) we start the new trajectories from the measured height at t_2 using the corrected α'. The value of $c = 0.05$ was found by comparing predictions and measured height. Note that if the predicted and measured heights are very close, p_h is nearly 1 and no correction is required. With any subsequent observation, we thus reset and narrow the probability distribution of observing future height values. Also notice that the entire history of measured height values influences the adaptive predictor through the above equation.

To test the predictive ability of the model on an individual basis, we calculate the individual growth trajectories and the corresponding $\pm$SD curves for a subset of 27 subjects of the BILD data set for which multiple measurements were available (Table 2). These simulations are based on 500 realizations of the random walk model. Figure 6A demonstrates the predictive ability of the model when the growth history is not taken into account. The slope of the regression line is somewhat reduced compared to the full data set but the R^2 increases to 0.35 ($p = 0.0012$). Next, we utilize the history of growth below one year to predict height at 6 years using our adaptive correction method applied at 1, 2, 4, 6 and 12 months of age (Fig. 6B). In this case, substantially more of the 27 children are found to grow within their individual bounds of growth trajectory. Furthermore, the regression line has a slope much closer to unity [measured height = 0.838*predicted height + 18.3] with a significantly increased $R^2 = 0.51$ ($p < 0.00003$). Additionally, paired t-test shows no difference between measured and predicted heights with an average error of only 3.3%.

Finally, we demonstrate how our model can be used to detect growth deficiency. Figure 7 compares the time course of measured heights with the predicted trajectory in a female subject born prematurely and small for gestation. The model captures well the measured time course until 39 months of age when the subject started receiving growth hormone therapy. First, when compared to the WHO mean-2SD, it is apparent that the model predicts that the subject would not reach normal height much earlier than she actually was given therapy. Second, the model can be used to detect the efficiency of the therapy as the model can predict the trajectory without therapy. Lastly, for comparison, the much simpler analytical model-based prediction of height and SD are also plotted demonstrating a good agreement with the full numerical model.

	Mean	Median	SD	CI
N	27			
sex (boys/girls)	12	44.4%		
Gestational age (weeks)	39.9	40	1.05	0.41
Height at birth (cm)	49.8	49.5	1.78	0.71
Height at 1 month (cm	53.2	53	2.4	0.95
Height at 12 months (cm)	74.5	73.5	2.48	0.98
Height at 6 years (cm)	116.9	115.9	4.94	1.95
Model predicted height at 6 years (cm)	117.6	117.3	4.21	1.67

Table 2. Height data of a subset of the BILD cohort and model predictions. Definitions are given in Table 1. There is no statistically significant difference between measured and model predicted height at 6 years (paired t-test).

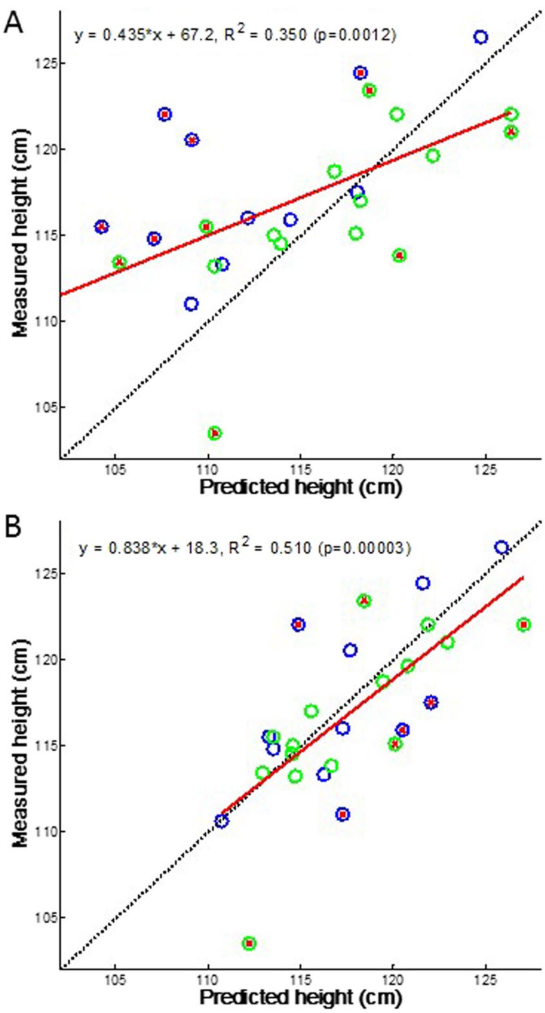

Figure 6. Prediction of measured height in the subset of 27 subjects who had complete data set between birth and 12 months of age. (**A**) Prediction of height at 72 months of age with the time-varying biased random walk model using 500 realizations and utilizing height only at 1 month. (**B**) Adaptive prediction of height at 72 months of age with the time-varying biased random walk model using 500 realizations as well as utilizing the growth history including height values at 1, 2, 4, 6 and 12 months of age in the same 27 subjects. Note that the adaptive Bayesian approach significantly improves the overall long term predictability of the target height by including a history of previous observation. Dotted line: line of identity; red line: regression line; blue circles: boys' height; green circle: girls' height; red cross in circles: those subjects whose measured height was outside the mean ± SD of the prediction.

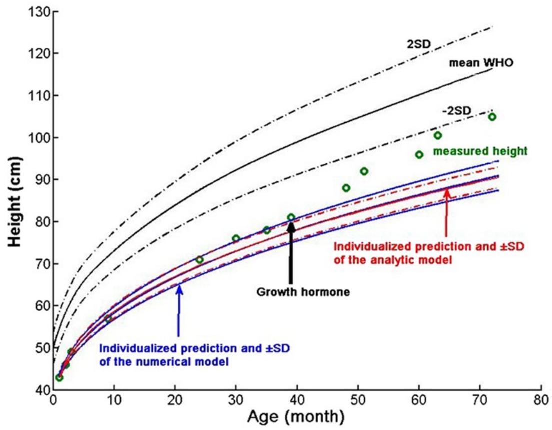

Figure 7. Measured (green circles) and predicted height (blue solid line) and $\pm$SD (blue dashed lines) in a girl born prematurely and small for gestation, following the upper limit of her personal growth trajectory and rapidly exceeding her predicted trajectory after starting growth hormone therapy at the age of 4 years. The WHO mean and $\pm$2SD curves for girls are also plotted (black). For comparison, the predictions of the analytical model are also shown (red).

Discussion

We have developed a probabilistic, adaptive growth model based on a time-varying biased random walk concept. The model was trained on the WHO population height curves. We also derived an analytical solution for the mean and SD of the growth model. Model parameters were estimated by fitting the WHO height curves including their population percentiles. To our knowledge, this is the first model to relate the discrete salutatory nature of individual growth trajectories to the variability of population growth curves. Using the model, we are also able to construct a family of potential growth trajectories for an individual child and hence we can estimate the entire probability distribution of predicted height and hence derive confidence intervals that a predicted target height will be reached within a given time period. In an independent dataset of 190 Swiss children from our birth cohort[29], we compared the predicted target heights at 6 years from measured data at 5 weeks and compared them to the prospectively measured heights at 6 years. We found a good correlation between prediction and measured outcome with an average error of only 3.9% and in a subset of 27 subjects the error dropped to 3.3% when an adative approach was applied.

Clinical implications. As an example of the power of modelling approach, we analysed the growth of a single child born prematurely and small for gestation. Her natural growth followed the upper limit of the trajectory predicted by our model. It is clinically important that the predicted trajectory did not reach the WHO population's 3rd percentile. This implies that the child had limited natural chances to reach normal height during growth. The decision to use growth hormone therapy could have been made earlier had such a prediction been available for the clinician. Additionally, after the start of growth hormone therapy, the measured height immediately exceeded the upper bound of the model predicted height. Given that growth hormone is not always effective in infants born prematurely and small for gestation, obtaining immediate evidence of treatment response is critical since potential side effects of treatment and costs need to be considered.

Classical growth models are based on normative age related height or weight distributions and the prediction of individual growth is always related to the population's age related height distribution. Population based growth charts are not able to account for the large variability of individual growth patterns[26]. Although regression based models are available[7, 9, 10, 30], to our knowledge, no adaptive mechanistically based prediction model exists which could predict an individual's growth from any given starting age and which also provides individualized probability that the target height will be reached. In addition to growth hormone therapy, the latter is particularly important in other clinical decision making processes. For example, in paediatric pharmacology[28], a long standing problem is that the general practitioner needs to dose medication without the knowledge of current height or weight, but with biometric measurements made at the last visit in the general paediatric practice. To estimate the risk that the dosage is incorrect, it is necessary to know the probability that the prediction of the current height or weight is true. If the probability of a correct prediction is low, the patient's actual height or weight needs to be assessed before medication. Similarly, in situations of treating sick children with growth deficits (Fig. 7), probabilistic estimates of the time interval beyond which a significant improvement in the child's growth pattern is expected helps medication planning. Furthermore, environmental conditions (e.g. nutrition, toxic substances) may change during a child's development. In such cases, the estimated growth cannot not easily be derived from WHO growth data[27]. In current clinical practice, the prediction of target-height is only based on paternal height measures and end-height predictions are based on bone X-rays. Both methods, however, do target height in adolescence. For early preschool age, however, our adaptive, probabilistic growth prediction may be more useful and can adjust for environmental influences and take into account the information content of the history of serial past height measurements.

Limitations of the methods and the model. The concept of time-varying biased random walk is useful and can be extended to predict other growth curves. For simplicity, we have restricted the model to height prediction within the preschool age range although prediction of weight gain is equally important. Future studies should address the issue of the more complex weight gain characteristics in this age range. Nevertheless, the concept could be adapted to older age ranges where hormonal factors and puberty significantly complicate the prediction. Such predictions may require more complex time-varying increment distributions than the power law applied here. While there are proposed mechanisms for the salutatory growth such as pulsatile growth hormone expression[20–22] and cell cycle dynamics[23], our model only incorporate the discontinuous nature of growth at the level of the organism.

We have also made several assumptions, which need to be considered in future work generalising the method for other age ranges. At birth, length within the WHO data set is assumed to be normally distributed. We also assumed that the birth height is determined genetically although feeding and environmental fluctuations probably contribute significantly. To account for the population variability, it was necessary to introduce a small intra-individual random variability (2.5%) of growth rate for each subject. Furthermore, most parameters of the model were derived from fitting the training set of the WHO data. Although this is a limitation of the approach, the WHO data set has been shown to be appropriate for a wide range of situations[27]. If data specific to the country or region were available, these parameters may be fine-tuned for a given type of population in future studies. This limitation together with the unknown environmental factors such as feeding pattern may explain the fact that the model is able to capture only about half of the variance of the individual data (Fig. 6B). Additionally, the biological evidence and quantitative description of saltatory growth and waiting time distributions is limited especially in human subjects. The best evidence is in infancy and early childhood[14, 17, 24] but data for older age groups to allow quantification of the width of the random walk distribution are scarce. Any new data, however, can be easily incorporated into the framework of the current model by changing the distributions or improving on the adaptive correction. For example, in case the growth is more continuous in a certain age, the distribution of increments would need to peak around a small value and the stasis times should be in the same order of magnitude as the time resolution of the measurements. In this case, the computational model simply reduces to a continuous growth model without discrete steps. The analytical model, however, has a general form that does not depend on the knowledge of the exact distributions of the salutatory steps (Eq. 10). If more specific information on salutatory and continuous growth patterns become available, future studies could improve the predictive power of our computational model. The use of lognormal distribution of waiting times may be related to how random walks with a drift lead to lognormal escape times[31]. Finally, the calculation of the SD in the analytical model assumes that the increments are uncorrelated. In its current form, the analytical model can only be used to predict the SD and hence it is not possible to apply in an adaptive manner which is why it was only used for the individual child's case in Fig. 7.

In summary, our probabilistic, adaptive growth model represents a novel approach to understand and quantitatively predict growth and development in young children. The model links discrete growth in individuals to continuous population growth curves and is of high relevance for clinical decisions in the context of precision medicine. To our knowledge, this is the first individualized growth prediction algorithm, which can be performed based on previous history of observations. This improves the safety of growth predictions for clinical decision making such as drug dosing. Furthermore, the statistical predictions by the model could be automated for tele-monitoring purposes combined with automatic learning from each observed height. Such an approach may also help detect and better quantify the probability of upcoming growth deficits in chronic diseases at an early stage and the effects of subsequent treatment.

Materials and Methods

Random walk models of growth. The model can be used to mimic the growth of an organ or the entire body with fluctuations during the process. The model is a feedback system including a random walker whose steps x are taken from a distribution with time varying parameters. We first define a window of length w during which the parameters of the distribution do not change. The random walk starts with initializing the mean (μ) and variance (σ^2) of the distribution with values $\mu(1)$ and $\sigma^2(1)$, respectively, and generating w number of values for x. During the ith window of length w, the walker takes values from a probability distribution written as p(i,x) with mean and variance of $\mu(i)$ and $\sigma^2(i)$, respectively. Since the growth must approach a finite value, we require that $\mu(i)$ gradually decrease approaching 0 with increasing i. Once a long series of x is obtained, the final output of the model is the cumulative running sum of x:

$$g(k) = \sum_{1}^{k} x_j \tag{1}$$

For convenience, we assume that k is an integer multiple of w, that is, $k = nw$. The growth process is then represented by the time series of g(k). The value $\mu(1)$ can be thought of as the initial level of growth factors (e.g. hormones) present in the system at the beginning of the biological growth process. The window w can be considered as a sensor that measures the average level of growth and the speed of growth is reduced at w intervals. Thus, as $\mu(i)$ monotonously decreases in successive windows, the growth slows down. The $\sigma^2(i)$ can then be used to generate variabilities representing fluctuations within an individual subject.

To obtain an analytic model, we partition the sum in Eq. 1 into segments of length w.

$$g(k) = \sum_{1}^{w} x_j + \sum_{w+1}^{2w} x_j + \ldots + \sum_{(n-1)w+1}^{k=nw} x_j \tag{2}$$

If w is large, we can approximate each sum as the corresponding μ times w:

$$g(k) = w\mu(1) + w\mu(2) + \ldots + w\mu(n) \tag{3}$$

which is also the expected value of g(k) in Eq. 2. Furthermore, if the increments in Eq. 2 are not correlated in time, g(k) is a simple sum of independent random variables. Therefore, the variance of g(k) is simply the sum of the variances of each random variables which is constant within the time window of w and hence it can be written as:

$$\sigma_g^2(k) = w\sigma^2(1) + w\sigma^2(2) + \ldots + w\sigma^2(n) \tag{4}$$

For simple choices of $\mu(i)$, Eq. 3 can be solved analytically. For example, if $\mu(i) = \delta\mu(i-1)$ with $0 < \delta < 1$, we can write Eq. 3 as:

$$g(k) = \mu_0 w[1 + \delta + \delta^2 + \ldots + \delta^n] \tag{5}$$

where we introduced $\mu_0 = \mu(1)$. This geometric series can be summed up in a closed form

$$g(k) = \mu_0 w \frac{1 - \delta^{n+1}}{1 - \delta} \tag{6}$$

Finally, for large $k = nw$, we can replace k with continuous time t and after some manipulations we obtain the following solution:

$$g(t) = \frac{\mu_0 w}{1 - \delta}(1 - \delta e^{-\gamma t}) \tag{7}$$

where $\gamma = \ln(1/\delta^{1/w})$ and hence $\tau = 1/\gamma$ is the time constant of the system. Eq. 7 shows that the time course of the continuous biased and time varying random walk is an exponential saturation. The w can be considered as a time scale of the sensor which smooths the microscopic fluctuations of the growth process. The initial growth rate, μ_0, only affects the saturation level. Note that the strength of the feedback, δ, enters both the saturation level and the velocity with which the saturation is approached.

Another simple choice for the mean to decrease is a power law form $\mu(i) = \mu_0 i^\alpha$ with $\alpha < 0$. In this case, Eq. 3 becomes

$$g(k) = \mu_0 w[1^\alpha + 2^\alpha + \ldots + n^\alpha] = \mu_0 w \sum_{1}^{n} i^\alpha \tag{8}$$

Using $n = k/w$ and transitioning to the continuous time domain, Eq. 8 is written as

$$g(t) = w\mu_0 \int_1^{t/w} \tau^\alpha d\tau \tag{9}$$

Eq. 8 is then written in its final form as

$$g(t) = At^{1+\alpha} + B \tag{10}$$

which is a power law growth function with $A = \mu_0 w^{-\alpha}/(1+\alpha)$ and $B = -\mu_0 w/(1+\alpha)$. However, due to the relation among the 4 parameters given by $B = w^{1+\alpha}A$, only 3 of them are independent. Furthermore, Eq. 10 is similar to the extension of the Faulhaber formula to non-integer powers[32] but it provides an improved approximation of Eq. 8 with an error <0.6% for a 6 year prediction of g(t) when $\alpha = -0.5$.

To derive a formula for the variance of the growth curve, we note that Eq. 4, can be written in the same form as Eq. 8 by assuming that $\sigma(i) = s\mu(i)$ where s is a constant. The same arguments can be applied as for the mean and the final result is:

$$\sigma^2(t) = Ct^{1+2\alpha} + D \tag{11}$$

where $C = s^2\mu_0^2 w^{-2\alpha}/(1+2\alpha)$ and $D = -ws^2\mu_0^2/(1+2\alpha)$.

A computational model of height. To extend the analytical model computationally, we note that the derivative of the power law in Eq. 10 defines the increment distribution of growth. Taking the derivative of the WHO height data and applying a linear regression in the log-log domain, we find that the exponent α is −0.543 with an intercept of 1.037 for the average height of boys whereas these numbers for girls are −0.511 and 0.93, respectively. The WHO data reveal a range of growth rates within the population of both boys and girls; therefore, we also fit the mean ± SD and the mean ± 2SD increment curves, with the power law model. For boys, this gives an α of −0.505 for the mean + 2SD curve and −0.589 for the mean − 2 SD curve whereas for girls, these exponents are −0.483 and −0.546, respectively.

While the fully deterministic power law model captures the mean growth rates, it does not account for the distribution of growth increment amplitudes and waiting times (stasis) at each step of the random walk. We incorporate these stochastic elements into a computational model and demonstrate that in the statistical sense, this model also describes the average height curves in the WHO data[27] which is then taken as a training set for the model. The overall growth velocity is determined by variations in both the growth increments and the stasis intervals as demonstrated schematically in Fig. 2. The relationship between the amplitude of the random walk steps and length of stasis periods determine the growth vector. Long stasis times and low amplitudes are related to slow growth and via versa. For growth in height, the amplitudes must be non-negative. If the model was applied to describe weight, an occasional loss can also be allowed and hence the amplitudes can take negative values.

Growth increment amplitudes, stasis times and their variations have been described in the literature only in small selected populations of different age groups. In infancy, peak amplitudes of salutatory growth were described to be between 0.46 and 1 cm[14, 24] and their variations between 10 and 40% while in early school-age these numbers were between 0.2 and 0.42 cm with variations of 10 to 25%[14, 24, 25]. The corresponding stasis times were reported to be between 10 and 17 days with variations of ~10%. However, there is no systematic data on developmental changes of growth increments and mean stasis time, $\underline{T}_{stasis}$, and their variances σ_{ampl} and σ_{stasis}, respectively. These parameters can, however, be estimated by fitting model to the WHO population data. Additionally, due to the strong genetic component of growth[33], the initial height at birth will be taken to represent genetic predisposition to growth. Thus, a linear regression between α values of the mean and mean $\pm$ SD and $\pm$2SD growth curves and the corresponding initial heights at birth provides the genetic background in the model.

To construct a computational random walk model, we use the following assumptions: (*1*) the time variation of the mean of the growth increment distribution is obtained by differentiating the growth function in Eq. 10 with respect to time which is a power law with an exponent α. Note that since the power law analytic model provides a better description of the data (see Fig. 1), only this model is retained for further analysis. (*2*) Since experimental data show a right skewed distribution of height increments[34], we assume that the increment distribution is lognormal which also guarantees that all increments are positive. (*3*) We assume that the distribution of stasis times is also lognormal with a mean $\underline{T}_{stasis}$ and variance σ_{stasis}. Assigning specific values for these parameters and knowing the height at birth that determines α through the regression as described above, we can simulate individual growth curves. Thus, using 200 random walks for a single subject, we compute the expected growth curve of the subject by ensemble averaging. Since the population distribution in the WHO data set for heights is normal[27], we mimic the growth of 100 infants with a normal distribution of initial heights at birth (mean height is 49.88 cm and SD is 2.1 cm). Summing up all these growth curves, we determine the least square error between the simulated population and the mean and $\pm$2SD height trajectories of the WHO data which allows us to optimize these parameters to best describe the WHO population. Once the model is set up, we introduce an additional feature of the model, an adaptive correction mechanism that accounts for environmental influences as described in the Results section. Finally, the model is used to predict height at 6 years of age in an independent population data set.

Experimental data. In a prospective birth cohort (BILD)[29] of preschool children (40% boys), height, weight, BMI was measured at the age of 5 weeks and at the age of 6 years using standard measurement techniques (Table 1). The cantonal ethics committee of the University of Basel and the Children's Hospital approved the study. All methods were performed in accordance with the relevant guidelines and regulations and written informed consent was obtained from all parents. 40% of the children were boys.

References

1. Bertalanffy, L. v. Quantitative laws in metabolism and growth. *Q Rev Biol* **32**, 217–231 (1957).
2. Koch, L. G. & Britton, S. L. Development of animal models to test the fundamental basis of gene-environment interactions. *Obesity (Silver Spring)* **16**(Suppl 3), S28–32, doi:10.1038/oby.2008.513 (2008).
3. Skinner, M. K. Environmental Epigenetics and a Unified Theory of the Molecular Aspects of Evolution: A Neo-Lamarckian Concept that Facilitates Neo-Darwinian Evolution. *Genome Biol Evol* **7**, 1296–1302, doi:10.1093/gbe/evv073 (2015).
4. Cummings, A. M. & Kavlock, R. J. Gene-environment interactions: a review of effects on reproduction and development. *Crit Rev Toxicol* **34**, 461–485 (2004).
5. Custovic, A., Marinho, S. & Simpson, A. Gene-environment interactions in the development of asthma and atopy. *Expert Rev Respir Med* **6**, 301–308, doi:10.1586/ers.12.24 (2012).
6. Ramos, R. G. & Olden, K. Gene-environment interactions in the development of complex disease phenotypes. *Int J Environ Res Public Health* **5**, 4–11 (2008).
7. Gliozzi, A. S., Guiot, C., Delsanto, P. P. & Iordache, D. A. A novel approach to the analysis of human growth. *Theor Biol Med Model* **9**, 17, doi:10.1186/1742-4682-9-17 (2012).
8. Lampl, M. Perspectives on modelling human growth: mathematical models and growth biology. *Ann Hum Biol* **39**, 342–351, doi:10.3109/03014460.2012.704072 (2012).
9. Nierop, A. F. *et al*. Modelling individual longitudinal human growth from fetal to adult life - QEPS I. *J Theor Biol* **406**, 143–165, doi:10.1016/j.jtbi.2016.06.007 (2016).
10. Rahmandad, H. Human growth and body weight dynamics: an integrative systems model. *PLoS One* **9**, e114609, doi:10.1371/journal.pone.0114609 (2014).
11. West, G. B., Brown, J. H. & Enquist, B. J. A general model for ontogenetic growth. *Nature* **413**, 628–631, doi:10.1038/35098076 (2001).
12. Makarieva, A. M., Gorshkov, V. G. & Li, B. L. Ontogenetic growth: models and theory. *Ecological Modelling* **176**, 15–26, doi:10.1016/j.ecolmodel.2003.09.037 (2004).
13. Hirst, A. G. & Forster, J. When growth models are not universal: evidence from marine invertebrates. *Proc Biol Sci* **280**, 20131546, doi:10.1098/rspb.2013.1546 (2013).
14. Lampl, M., Veldhuis, J. D. & Johnson, M. L. Saltation and stasis: a model of human growth. *Science* **258**, 801–803 (1992).
15. Lampl, M., Ashizawa, K., Kawabata, M. & Johnson, M. L. An example of variation and pattern in saltation and stasis growth dynamics. *Annals of Human Biology* **25**, 203–219, doi:10.1080/03014469800005582 (1998).
16. Hermanussen, M., De Lama, M. D. R., Tresguerres, J. A. F., Grasedyck, L. & Burmeister, J. Short-term growth: Evidence, for chaotic series of mini growth spurts in rat growth. *Physiology & Behavior* **64**, 7–13, doi:10.1016/S0031-9384(98)00023-7 (1998).
17. Hermanussen, M. *et al*. Micro and macro perspectives in auxology: findings and considerations upon the variability of short term and individual growth and the stability of population derived parameters. *Annals of Human Biology* **25**, 359–385, doi:10.1080/03014469800005702 (1998).

18. Johnson, M. L., Straume, M. & Lampl, M. The use of regularity as estimated by approximate entropy to distinguish saltatory growth. *Annals of Human Biology* **28**, 491–504 (2001).
19. Lampl, M. Human growth from the cell to the organism: saltations and integrative physiology. *Ann Hum Biol* **36**, 478–495, doi:10.1080/03014460902911670 (2009).
20. Martin, J. B., Renaud, L. P. & Brazeau, P. Jr. Pulsatile growth hormone secretion: suppression by hypothalamic ventromedial lesions and by long-acting somatostatin. *Science* **186**, 538–540 (1974).
21. Veldhuis, J. D., Roemmich, J. N., Richmond, E. J. & Bowers, C. Y. Somatotropic and gonadotropic axes linkages in infancy, childhood, and the puberty-adult transition. *Endocr Rev* **27**, 101–140, doi:10.1210/er.2005-0006 (2006).
22. Veldhuis, J. D. *et al*. Gender modulates sequential suppression and recovery of pulsatile growth hormone secretion by physiological feedback signals in young adults. *J Clin Endocrinol Metab* **90**, 2874–2881, doi:10.1210/jc.2004-1363 (2005).
23. Lee, M. G., Norbury, C. J., Spurr, N. K. & Nurse, P. Regulated Expression and Phosphorylation of a Possible Mammalian Cell-Cycle Control Protein. *Nature* **333**, 676–679, doi:10.1038/333676a0 (1988).
24. Caino, S., Kelmansky, D., Adamo, P. & Lejarraga, H. Short-term growth in healthy infants, schoolchildren and adolescent girls. *Annals of Human Biology* **33**, 213–226, doi:10.1080/03014460500525369 (2006).
25. Lampl, M. & Johnson, M. L. Normal human growth as saltatory: Adaptation through irregularity. *Applications of Nonlinear Dynamics to Developmental Process Modeling* 15–38 (1998).
26. Lampl, M. & Thompson, A. L. Growth chart curves do not describe individual growth biology. *American Journal of Human Biology* **19**, 643–653, doi:10.1002/ajhb.20707 (2007).
27. Group, W. M. G. R. S. WHO Child Growth Standards: length/height-for-age, weight-for-age, weight-for length, weight-for-height and body mass index-for-age: methods and development. *Geneva*: *World Health Organization* (2006).
28. Bielicki, J. A. *et al*. Not too little, not too much: problems of selecting oral antibiotic dose for children. *BMJ* **351**, h5447, doi:10.1136/bmj.h5447 (2015).
29. Fuchs, O., Latzin, P., Kuehni, C. E. & Frey, U. Cohort profile: the Bern infant lung development cohort. *Int J Epidemiol* **41**, 366–376, doi:10.1093/ije/dyq239 (2012).
30. Ranke, M. B. *et al*. Accurate long-term prediction of height during the first four years of growth hormone treatment in prepubertal children with growth hormone deficiency or Turner Syndrome. *Horm Res Paediatr* **78**, 8–17, doi:10.1159/000339468 (2012).
31. Kish, E. A., Granqvist, C. G., Der, A. & Kish, L. B. Lognormal distribution of firing time and rate from a single neuron? *Cogn Neurodyn* **9**, 459–462, doi:10.1007/s11571-015-9332-6 (2015).
32. Parks, H. R. Sums of non-integral powers. *J Math Anal Appl* **297**, 343–349 (2004).
33. Brescianini, S., Giampietro, S., Cotichini, R., Lucchini, R. & De Curtis, M. Genetic and environmental components of neonatal weight gain in preterm infants. *Pediatrics* **129**, e455–459, doi:10.1542/peds.2010-0510 (2012).
34. Lampl, M. & Johnson, M. L. Distribution methods and analysis of nonlinear longitudinal data. *Methods Enzymol* **321**, 182–195 (2000).

Acknowledgements

The authors thank Prof. Urs Zumsteg, Pediatric Endocrinology, UKBB, Basel for helpful discussions. This study was partially funded by NIH HL-126040 and SNF # 320030-163311.

Author Contributions

B.S. developed the analytical and computational model and wrote manuscript. U.F. developed computational model, collected data and wrote manuscript.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.