OXFORD UNIVERSITY PRESS | European Journal of Orthodontics

Craniofacial growth and SITAR growth curve analysis

Journal:	*European Journal of Orthodontics*
Manuscript ID	EJO-2021-OA-0220.R2
Manuscript Type:	Original Article
Research Topics:	Growth, clinical, Statistics

Craniofacial growth and SITAR growth curve analysis

Abstract

Background: SITAR (SuperImposition by Translation And Rotation) is a shape invariant growth curve model that effectively summarises somatic growth in puberty.

Aim: To apply the SITAR model to longitudinal mandibular growth data to clarify its suitability to facial growth analysis.

Subjects and methods: 2D-cephalometric data on two mandibular measurements (ap: articulare-pogonion; cp: condylion-pogonion) were selected from the Denver Growth Study, consisting of longitudinal records (age range: 7.9 – 19.0 years) of females (sample size N: 21; number of radiographs n: 154) and males (N: 18; n: 137). The SITAR mixed effects model estimated, for each measurement and sex separately, a mean growth curve versus chronological age, along with mean age at peak velocity (APV) and peak velocity (PV), plus subject-specific random effects for PV and mean size. The models were also fitted versus Greulich-Pyle bone age.

Results: In males, mean APV occurred at 14.6 y (ap) and 14.4 y (cp), with mean PV 3.1 mm/y (ap) and 3.3 mm/y (cp). In females, APV occurred at 11.6 y (ap and cp), with mean PV 2.3 mm/y (ap) and 2.4 mm/y (cp). The models explained 95-96% of the cross-sectional variance for males, and 92-93% for females. The random effects demonstrated standard deviations (SD) in size of 5.6 mm for males and 3.9 mm for females, and SDs for PV between 0.3 mm/y and 0.5 mm/y. The bone age results were similar.

Conclusion: The SITAR model is a useful tool to analyse epidemiological craniofacial growth based on cephalometric data, and provides an array of information on pubertal mandibular growth and its variance in a concise manner.

Introduction

During the physical development of the human body, its component parts display very different growth rates, with great variability with regard to timing and intensity of the pubertal growth spurt. The necessity to estimate growth curves for the mandible, independently and separately to somatic growth, has long been acknowledged (1).

Numerous approaches have been considered in the past to describe growth curves at a population level. Initial efforts consisted of fitting growth curves to individual serial growth data, and then summarizing the results over individuals. Such mostly outdated mathematical models – varying in complexity – have been applied to craniofacial growth and comprise unsmoothed plots of annual increments, at times connected simply by crude linear graphs (2-5), undefined smoothed graphs (6), and logistic (7) or polynomial (8) models.

With the advent of multilevel modelling, the possibility emerged to estimate not only puberty timing for individuals by fitting a single model to the entire cohort, but also to identify the amount of individual variation in the pattern of growth. First attempts to apply such modelling on growth curves of craniofacial structures were based on two-level polynomial models (9, 10), and have evolved since to multilevel polynomial models (11-18), together with the computation of cubic smoothing splines (19-21). The investigations based on multilevel mixed models usually report the individual variations as polynomial coefficients, which are, however, difficult to interpret.

First described in 2010, SITAR (SuperImposition by Translation And Rotation) is a shape invariant curvilinear model based on a natural cubic spline that allows the representation of epidemiologic growth data as a single curve (22). The underlying assumption is – not unlike J. M. Tanner's doctrine - that the velocity curve is essentially constant in shape, and differs only between individuals in timing and intensity (23). Thus, the single mean curve is achieved by matching individual growth curves to the mean by shifting the curve up or down (representing differences in absolute size), left or right (for differences in timing of growth), and shrinking or stretching the age scale (for differences in the intensity of growth, i.e. velocity). These three parameters are presented as random effects while fitting the curve. The outcome is a mean curve and triplets of parameters per individual (size, timing, and intensity). The mean curve allows the average age at peak velocity (APV) and mean peak velocity (PV) to be determined, and the triplets of parameters summarize the departure of each individual curve from the mean in terms of three biologically meaningful variables. In some cases, the data are insufficient to estimate the timing random effect, and a reduced model is fitted without it. The elegance of the SITAR model lies in the simplicity of its comprehensive yet concise results, providing information not only on the mean growth curve, but also on individual growth patterns within the cohort.

The aim of this observational investigation was to apply for the first time the SITAR model to a craniofacial structure and to interpret its findings, and specifically to test whether a pubertal growth spurt could be detected. To this end, historical cephalometric data of a longitudinal growth study were analysed with the SITAR model to describe mandibular growth around the time of puberty.

Material and Methods

The data were drawn from the University of Oklahoma Denver Growth Study, commonly known as the Denver Growth Study. Conducted between 1927 and 1967, its inventory comprises longitudinal records (mainly models, photos and radiographs) of 158 females and 155 males of healthy and non-syndromic children of Caucasian origin. None of the recruited individuals were exposed to any type of orthodontic treatment or extraction of permanent teeth prior to or during the observation period. The socio-economic background of the participants' families was reportedly stable and overall slightly above average.

For the entire observation period, the same cephalostat was used to adhere to a consistent focus-to-coronal plane distance. All lateral cephalograms were produced with the head stabilized by ear rods and teeth in centric occlusion.

The entire dataset of the Denver growth study was considered and screened. Subjects with at least four dated serial cephalograms and four dated serial hand-wrist radiographs (HWR), starting between 8 and 9 years and spanning the growth spurt period, were included in the present investigation. After eliminating cephalograms of inadequate quality, individuals with too few radiographs were excluded. Twenty-one females with 154 cephalograms and 128 HWR and 18 males with 137 cephalograms and 112 HWR were included. The measurement occasions were nominally yearly, and the vast majority (208 / 252 or 83%) of adjacent measurement pairs were spaced roughly a year apart (6-18 months), with a further 32 two years apart. There were no missing data as such, just varying time gaps between measurements.

The HWR were assessed for bone age according to Greulich and Pyle (24), in accordance to the protocol outlined in the atlas: after first finding the standard best matching the film to be assessed, the individual bones were compared in sequential order, using the descriptions that accompany the standards in the atlas.

The cephalograms were traced with the following mandibular landmarks: Condylion (Co), defined as the most posterior-superior point of the mandibular condyle, Articulare (Ar) as the point of intersection of the dorsal contours of the processus articularis mandibulae and os temporale, and Pogonion (Pg) as the most anterior point of the symphysis of the mandible with the head viewed in the Frankfort plane (Figure 1). The tracings were digitized, using a tablet digitizer with a resolution of 1 milli-inch (AccuGrid, Numonics, Landsdale, PA) to obtain the mandibular measures cp (condylion-pogonion) and ap (articular-pogonion), and recorded after adjustment for the magnification factor of 96% (25).

Two board certified orthodontists [H.K., G.M.] shared the workload of tracing and measuring the cephalometric landmarks (each a complete subset independently) and one [P.B.] determined bone age. Thirty randomly selected radiographs were assessed a second time at least two months later to determine intra- and inter-observer reliability.

Legal and ethical approval for releasing the data was obtained from the Federal Commission of Experts for Professional Secrecy in Medical Research (XXX).

Statistical analysis

Descriptives were reviewed in IBM SPSS version 25 (IBM Corp., Armonk, NY, USA) and the SITAR model computed in R version 3.6.3 (R Core Team. 2019. R: a language and environment for statistical computing. Version 3.6.3 ed. Vienna: R Foundation for Statistical Computing), with SITAR package (Cole T. 2019. Super Imposition by Translation and Rotation Growth Curve Analysis R package sitar. Version 1.1.2. Comprehensive R Archive Network).

Intra- and interobserver reliability was assessed by reviewing the absolute differences and computing the intraclass correlation coefficients (ICC) for absolute agreement based on two-way random effects analysis of variance.

The HWRs and cephalometric measurements were taken at different times, so the bone ages based on the HWRs were adjusted to the cephalometric measurement ages using cubic interpolation. The within subject correlations between age and bone age all exceeded 0.97.

SITAR models were computed for the two sexes and the two measurements, according to chronological age and bone age separately, using all the available data. The model including all three random effects did not converge so the reduced model was fitted, with just the random intercept (size) and random slope (intensity). Mean peak velocity and age at peak velocity were obtained by differentiating the mean growth curve, and their standard errors were obtained using the bootstrap. The numbers of spline curve degrees of freedom were optimised, as were the options of linear or log transformation for the measurement and age, to minimise the Bayesian Information Criterion (BIC).

Fitting the models with the measurement log-transformed allowed the size random effect SD and the peak velocity to be estimated in fractional (i.e. %/100) units, which is useful when comparing them to those for height. The model's goodness of fit was assessed by the percentage of variance explained, calculated as $100\left(1-\left(\sigma_r / \sigma_f\right)^2\right)$ where σ_f is the residual standard deviation (RSD) after fitting just the fixed effects mean curve to the data, and σ_r is the RSD after adding the SITAR random effects.

Results

Table 1 summarises the dataset by sex. The median number of cephalograms per subject was 8 (min. 4; max. 9), the mean age at first cephalogram was 8.7 years (range 7.9 y to 9.0 y), and the mean age at last cephalogram was 16.6 years of age (range 15.0 y to 19.0 y). After interpolation, mean bone age was almost one year earlier than mean chronological age in males, while in females the two ages were similar.

ICC scores confirmed the reproducibility of the cephalometric measurements, both of condylion-related distances (ICC intra-observer: 0.98 [95% CI: 0.96 – 0.99]; inter-observer: 0.97 [95% CI: 0.93 – 0.99]) and articulare-related distances (ICC intra-observer: 0.99 [95% CI: 0.98 – 0.99]; inter-observer: 0.97 [95% CI: 0.94 – 0.98]). The mean absolute differences were slightly smaller for conylion-based measurements (intra-observer: 0.8 mm [95% CI: 0.4 mm – 1.2 mm]; inter-observer: 1.5 mm [95% CI: 1.0 mm – 1.9 mm]) than for articulare-based measurements (intra-observer: 1.2 mm [95% CI: 0.8 mm – 1.6 mm]; inter-observer: 2.2 mm [95% CI: 1.7 mm – 2.7 mm]). The HWR bone age assessment was repeatable with high consistency (ICC score: 0.99 [95% CI: 0.98 – 1.00]).

The plots in Figure 2 show the raw data, i.e. the growth curves colour-coded by individual plotted against chronological age, by sex (rows) and measurement (columns). The plots show considerable variability, and the shape of the underlying mean curve is far from obvious.

Optimal SITAR models were fitted to each dataset, with three degrees of freedom for the spline curve and the measurement log transformed. Figure 3 shows plots summarising the SITAR models, depicting the individual curves both raw (in grey) and after SITAR adjustment (colour-coded), i.e. with their size and intensity random effects suitably adjusted to best superimpose them on (or rather under) the mean curve, shown as a heavy black line. In addition, the mean velocity curve is shown as a dashed line, with the age at peak velocity marked with a vertical dotted line.

Table 2 summarises the models. The adjusted curves were close to the mean curve, with the variance explained by the models 95-96% for males and 92-93% for females. The velocity curves all had an obvious peak, at age 14.5 years in males and 11.6 years – three years earlier – in females. Peak velocity was also higher in males than females, both in absolute terms, at around 3.2 mm versus 2.4 mm per year, and in fractional terms, at 2.8% versus 2.4% per year.

The random effects summarise the intercept (i.e., individual difference in size) and the age scaling or intensity (i.e., individual difference in velocity). Their standard deviations (SD) are given in Table 2. The SD for size was 5.6 mm (5%) for males and 3.8 mm (3.6%) for females. The SD for intensity ranged between 0.12 and 0.16, corresponding to 12-16% or 0.3-0.5 mm/y variation in peak velocity. The two random effects had a correlation of 0.3 to 0.4 in males and -0.1 to 0.0 in females.

The results for the bone age models were broadly similar to those for chronological age (Supplementary Table 1), though the percentage of variance explained was smaller, for males (92-94%) and females alike (87-90%). Mean bone age at peak velocity was almost one year earlier than for chronological age in males, at 13.7 years, and similar for females at 11.7

years, while peak velocity was the same. The random effects for size were similar, but those for intensity were somewhat larger at 19-22%.

Supplementary figures 1 and 2 show respectively the individual growth curves plotted against bone age, and the fitted SITAR models versus bone age.

Discussion

SITAR (SuperImposition by Translation And Rotation) is a shape invariant model for growth curves that has lately seen increasing popularity as a methodology to assess somatic height growth and has also been applied successfully to various other anthropometric measurements (23). While its validity has been ascertained in numerous previous reports (26), the present findings are a useful addition to the literature, as they corroborate SITAR's suitability to craniofacial growth analysis, which differs by being predominantly based on radiographic-derived measurements of millimetre increments. Considering the plethora of craniofacial studies available and the richness of the pertinent literature, it is-perhaps surprising that SITAR has not been applied to craniofacial growth before.

In large growth studies, it is crucial to reduce the dimensionality of the data, yet at the same time retain indicative (i.e. biologically meaningful) information from the observations. SITAR helps to achieve this because it models the underlying biology of pubertal timing and intensity. It also allows individual growth curves to be generated, including estimates of their peak velocity and age at peak velocity. It is not only a tool to uncover the average growth pattern of a cohort and the relationship of the individual towards the mean, but it opens up the prospect to examine the heterochrony within the scope of comparative studies as well. With many preserved and readily accessible historical craniofacial studies (27), a wider request for SITAR is to be expected.

Using CP and AP as craniofacial measurements

This investigation focuses on *one* craniofacial feature. CP and AP are basically two mandibular traits which are largely surrogates for each other, and were only used to demonstrate that each's shortcoming as landmark is of no relevance. The intention was to "duplicate" the model to investigate the impact of articulare versus condylion. Although Figure 3 portrays a slight differential landmark behaviour, the small differences do not reduce the validity of the models and have no discernible impact on the model fit (i.e. the variance explained by the models). These minor differences probably emanate from the difficulty to trace condylion, and from articulare being influenced by the growth-related displacement of the clivus or the mandible's rotation around the condyle (Björk's 'matrix rotation'). Now, having established SITAR for mandibular length, the modelling of even smaller growth-related growth changes such as Nasion – Sella would be interesting and relevant. Additionally, the comparison of specific growth patterns (e.g. low versus high angle subjects, or subjects with different Angle classifications) could be of interest.

Comparing SITAR-based results to previous investigations of the Denver data

Some remarks can be made relating to published works. SITAR was fitted here in its reduced form omitting the timing random effect, with a subject-specific random intercept (size) and random slope (intensity). It differs from other growth curve models in that the "slope" adjustment is achieved by stretching / shrinking the age scale, which conserves the adult plateau in the mean curve. Conventional models, which rotate the curve in the plane, fail to do this. Here SITAR identified biologically plausible mean curves with age at peak mandibular velocity at around 14.5 years in males and 11.6 years in females, and peak mandibular velocity of 3.2 mm/y in males and 2.4 mm/y in females. These results are in line with those from a previous investigation on the same Denver data, with the sole difference of a later peak in females (19).

It has been suggested that PV values obtained with SITAR tend to be greater than those established with other population average models and cubic spline methodology, owing to the fact that the latter do usually not account for longitudinal data from a single individual (28). Yet, the comparison of PV obtained here by SITAR to PV as described in the literature does not seem to echo this assumption. While the disparity is subtle, the reported SITAR mandibular PVs fall well within the range of other PV scores based on population average models of mandibular growth (2, 12, 19, 21, 29).

Comparing the SITAR model for mandibular growth to other published SITAR models

- Variance explained by the model: The variance explained by the models is 95-96% for males and 92-93% for females. Although impressive, these numbers are not as close to 100% as the SITAR models published for somatic height (22, 23). This is probably because the reduced SITAR model was fitted, and including the timing random effect might have added a few percentage points to the variance explained. However, even so, the variance explained is relatively high given that the timing random effect is missing, and it raises the prospect that timing plays less of a part in mandibular growth than it does in somatic growth. The following rationale might be submitted as a possible interpretation of this observation: the growth pattern of the mandible, which doesn't evolve from the primordial skeleton, is manifestly distinct from the growth pattern of endochondral ossification seen in epiphyseal plates of long bones, and mandibular growth is known to be less regulated by genes and more responsive to environmental factors than long bones (30). As such, the weaker role of timing on peak velocity as a random effect could be assumed as an outcome of the lower level of genetic regulation.

- The size random effect: The standard deviations for the size random effect were 5.6 mm and 3.9 mm for males and females respectively, while the fractional variation for intensity i.e. peak velocity was 12-16% in both sexes. It is interesting to compare these results with those for height (23). As the mandible is obviously far smaller than stature, the size random effect and peak velocity (which both depend on mean size) must be measured in fractional units, i.e. adjusted for mean size, to allow a comparison. With the model fitted to the log of the measurements, the size random effect is estimated as mm per mm or %/100, while PV is estimated as mm/year per mm, or %/100 per year. This gives the percentage SDs for the size random effect in table 2 of 5% and 3.6% by sex, which are broadly similar to those published for height (3.7%) and leg length (5.4%) (23). The percentage PVs for ap and cp in males and females were respectively 2.8%, 2.8%, 2.4% and 2.3% per year, about half the reported PVs of 4-8% per year for ten linear anthropometric measures (23), indicating that at its peak, growth in the mandible is appreciably less intense than in the long bones. This, again, is in agreement with the fact that mandibular bone growth regulation differs from the classic genetic control governing endochondral ossification processes of the long bones (30).

- The intensity random effect: On the other hand, the comparison of the *intensity* random effect, with a standard deviation of 14%, corresponds well to published values for the intensity random effect SD for height and other linear measures (22, 23).

SITAR models based on interpolated bone age

The models based on bone age yielded very similar results, the main differences being an earlier age at PV, particularly for males, and greater variability in the intensity random effect. The former is due to mean bone age in the males being almost one year earlier than mean chronological age, while the latter may simply be due to greater measurement error in bone age compared to chronological age.

The added benefits of an assessment based on skeletal age that must be interpolated seem limited, and even at a correlation level of 0.97 caution must be applied when interpreting the results. This is the main reason why these data were only added as supplementary files. Since chronological age might have been regarded as limitation, it was nevertheless deemed beneficial to present models based on skeletal age, to allow a comparison.

Limitations

Obvious limitations were the known constraints inherent in cephalometric data analysis, most notably the two-dimensional representation of three-dimensional structures and the use of ill-defined landmarks (31, 32). Moreover, the small sample size, with just 39 subjects and 291 measurements could be criticized, though the number of measurements per subject was satisfactory for SITAR (33). In addition, the measurements were inherently noisier than height or weight, which are the more typical outcomes in SITAR models. The combination of a small sample and noisy data means that SITAR can be hard to fit, and it requires careful tuning to achieve model convergence. Despite this, the cohort was sufficiently large to generate meaningful results, albeit with some uncertainty.

Conclusion

The growth curve in puberty contains considerable information. SITAR is a shape invariant model for growth curve analysis of entire cohorts that also sheds information on the individual by portraying the variance as simple, biologically interpretable effects. This investigation is the first to report the application of SITAR to a craniofacial structure and it gives testimony to the suitability of SITAR for mandibular growth analysis during puberty, based on radiographic material.

Supplementary material

Supplementary Table 1.

Supplementary Figures 1 and 2: Growth curves and SITAR models according to Greulich-Pyle bone age.

Data availability

The Denver Growth Study is part of the AAOF Craniofacial Growth Legacy Collection that is openly available in a public repository (www.aaoflegacycollection.org). The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of interest

None to declare.

Acknowledgments

The authors are indebted to the late Dr. [XXX] who archived the data of the Denver growth study in [XXX] for the benefit of future research.

Literature

1. Hellman, M. (1935) The face in its developmental career. *The Dental Cosmos*, 77, 685-699.
2. Lewis, A.B., Roche, A.F. and Wagner, B. (1982) Growth of the mandible during pubescence. *The Angle Orthodontist*, 52, 325-42.
3. Hunter, C.J. (1966) The correlation of facial growth with body height and skeletal maturation at adolescence. *The Angle Orthodontist*, 36, 44-54.
4. Buschang, P.H., Jacob, H.B. and Demirjian, A. (2013) Female adolescent craniofacial growth spurts: real or fiction? *European Journal of Orthodontics*, 35, 819-25.
5. Hunter, W.S., Baumrind, S., Popovich, F. and Jorgensen, G. (2007) Forecasting the timing of peak mandibular growth in males by using skeletal age. *American Journal of Orthodontics and Dentofacial Orthopedics*, 131, 327-33.
6. Lewis, A.B., Roche, A.F. and Wagner, B. (1985) Pubertal spurts in cranial base and mandible. Comparisons within individuals. *The Angle Orthodontist*, 55, 17-30.
7. Roche, A.F., Lewis, A.B., Wainer, H. and McCartin, R. (1977) Late Elongation of the Cranial Base. *Journal of Dental Research*, 56, 802-808.
8. Buschang, P.H., Tanguay, R., Turkewicz, J., Demirjian, A. and La Palme, L. (1986) A polynomial approach to craniofacial growth: description and comparison of adolescent males with normal occlusion and those with untreated Class II malocclusion. *American Journal of Orthodontics and Dentofacial Orthopedics*, 90, 437-42.
9. Buschang, P.H., Tanguay, R., Demirjian, A., LaPalme, L. and Goldstein, H. (1988) Pubertal growth of the cephalometric point gnathion: multilevel models for boys and girls. *American Journal of Physical Anthropology*, 77, 347-54.
10. Buschang, P.H., Tanguay, R., Demirjian, A., LaPalme, L. and Turkewicz, J. (1988) Mathematical models of longitudinal mandibular growth for children with normal and untreated Class II, division 1 malocclusion. *European Journal of Orthodontics*, 10, 227-34.
11. Buschang, P.H., Santos-Pinto, A. and Demirjian, A. (1999) Incremental growth charts for condylar growth between 6 and 16 years of age. *European Journal of Orthodontics*, 21, 167-73.
12. Nahhas, R.W., Valiathan, M. and Sherwood, R.J. (2014) Variation in timing, duration, intensity, and direction of adolescent growth in the mandible, maxilla, and cranial base: the Fels longitudinal study. *Anatomical Record*, 297, 1195-207.
13. Chvatal, B.A., Behrents, R.G., Ceen, R.F. and Buschang, P.H. (2005) Development and testing of multilevel models for longitudinal craniofacial growth prediction. *American Journal of Orthodontics and Dentofacial Orthopedics*, 128, 45-56.
14. van der Beek, M.C.J., Hoeksma, J.B. and Prahl-Andersen, B. (1996) Vertical facial growth and statural growth in girls: A longitudinal comparison. *European Journal of Orthodontics*, 18, 549-555.

15. van der Beek, M.C.J., Hoeksma, J.B. and Prahl-Andersen, B. (1991) Vertical facial growth: a longitudinal study from 7 to 14 years of age. *European Journal of Orthodontics*, 13, 202-208.

16. Smith, S.L. and Buschang, P.H. (2002) Midsagittal facial soft-tissue growth of French Canadian adolescents. *American Journal of Human Biology*, 14, 457-67.

17. van Diepenbeek, A.F., Buschang, P.H. and Prahl-Andersen, B. (2009) Age-dependant cephalometric standards as determined by multilevel modeling. *American Journal of Orthodontics and Dentofacial Orthopedics*, 135, 79-87.

18. Wolfe, S.M., Araujo, E., Behrents, R.G. and Buschang, P.H. (2011) Craniofacial growth of Class III subjects six to sixteen years of age. *The Angle Orthodontist*, 81, 211-6.

19. Patcas, R., Wiedemeier, D.B., Markic, G., Beit, P. and Keller, H. (2017) Evidence of secular trend in mandibular pubertal growth. *European Journal of Orthodontics*, 39, 680-685.

20. Mellion, Z.J., Behrents, R.G. and Johnston, L.E., Jr. (2013) The pattern of facial skeletal growth and its relationship to various common indexes of maturation. *American Journal of Orthodontics and Dentofacial Orthopedics*, 143, 845-54.

21. Patcas, R., Herzog, G., Peltomaki, T. and Markic, G. (2016) New perspectives on the relationship between mandibular and statural growth. *European Journal of Orthodontics*, 38, 13-21.

22. Cole, T.J., Donaldson, M.D. and Ben-Shlomo, Y. (2010) SITAR--a useful instrument for growth curve analysis. *International Journal of Epidemiology*, 39, 1558-66.

23. Cole, T.J. (2020) Tanner's tempo of growth in adolescence: recent SITAR insights with the Harpenden Growth Study and ALSPAC. *Annals of Human Biology*, 47, 181-198.

24. Greulich, W.W. and Pyle, S.I., *Radiographic atlas of skeletal development of the hand and wrist*1959: Stanford university press.

25. Sherwood, R.J., Oh, H.S., Valiathan, M., McNulty, K.P., Duren, D.L., Knigge, R.P., et al. (2021) Bayesian approach to longitudinal craniofacial growth: The Craniofacial Growth Consortium Study. *Anatomical record*, 304, 991-1019.

26. Simpkin, A.J., Sayers, A., Gilthorpe, M.S., Heron, J. and Tilling, K. (2017) Modelling height in adolescence: a comparison of methods for estimating the age at peak height velocity. *Annals of human biology*, 44, 715-722.

27. Al-Jewair, T., Stellrecht, E., Lewandowski, L. and Chakaki, R. (2018) American Association of Orthodontists Foundation Craniofacial Growth Legacy Collection in the orthodontic literature-use and trends: A systematic review. *American Journal of Orthodontics and Dentofacial Orthopedics*, 153, 15-25 e10.

28. Cole, T.J. (2020) Estimating peak height velocity in individuals. *Annals of Human Biology*, 47, 584-584.

29. Gomes, A.S. and Lima, E.M. (2006) Mandibular growth during adolescence. *The Angle Orthodontist*, 76, 786-90.

30. Hinton, R.J. (2014) Genes that regulate morphogenesis and growth of the temporomandibular joint: a review. *Developmental Dynamics*, 243, 864-74.

31. Bookstein, F.L. (2016) Reconsidering "The inappropriateness of conventional cephalometrics". *American Journal of Orthodontics and Dentofacial Orthopedics*, 149, 784-97.

32. Moyers, R.E. and Bookstein, F.L. (1979) The inappropriateness of conventional cephalometrics. *American Journal of Orthodontics*, 75, 599-617.

33. Cole, T.J. (2018) Optimal design for longitudinal studies to estimate pubertal height growth in individuals. *Annals of Human Biology*, 45, 314-320.

Legends of Figures and Tables

Figures

Figure 1: Cephalometric landmarks and traced distances (purple and blue). Ar: Articulare; Co: Condylion; FH: Frankfort horizontal plane; Pg: Pogonion; cp (purple): distance condylion-pogonion; ap (blue): distance articulare-pogonion.

Figure 2: Plots with individual, colour-coded growth charts. Raw data presented for each measurement and sex individually. AP: Articulare-Pogonion; CP: Condylion-Pogonion.

Figure 3: SITAR models with the individual unadjusted curves in gray; individual adjusted curves colour-coded and the mean curve (heavy black line). Dashed line: mean velocity curve. Age at peak velocity: vertical dotted line.

Supplementary figure 1: Plots with individual, colour-coded growth curves versus Greulich-Pyle bone age. Raw data presented for each measurement and sex individually. AP: Articulare-Pogonion; CP: Condylion-Pogonion.

Supplementary figure 2: SITAR models versus Greulich-Pyle bone age with the individual unadjusted curves in grey; individual adjusted curves colour-coded and the mean curve (heavy black line). Dashed line: mean velocity curve. Age at peak velocity: vertical dotted line.

Tables

Table 1: Summary statistics by sex.

Table 2: Summary statistics for the four SITAR models versus chronological age.

Supplementary table 1: Summary statistics for the four SITAR models versus Greulich-Pyle bone age.

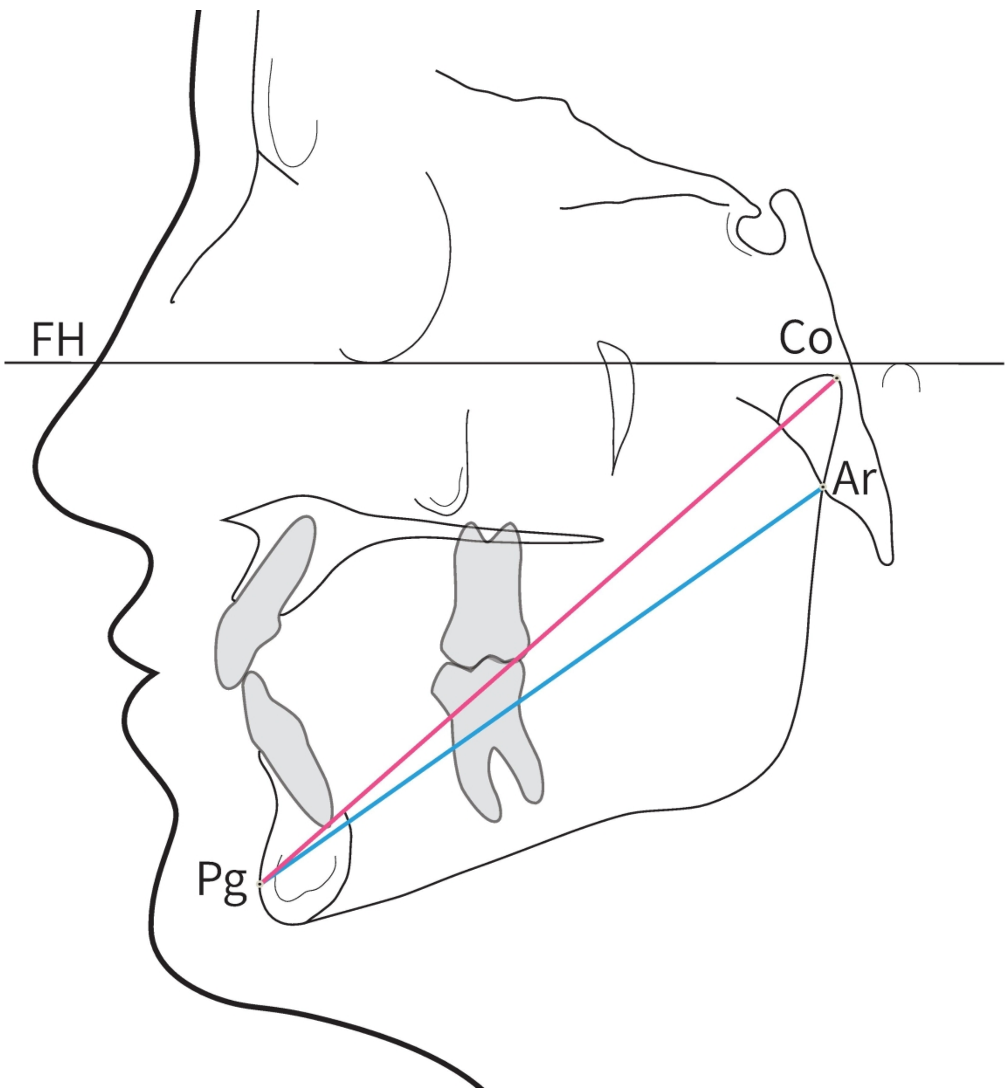


Figure 1

139x149mm (600 x 600 DPI)

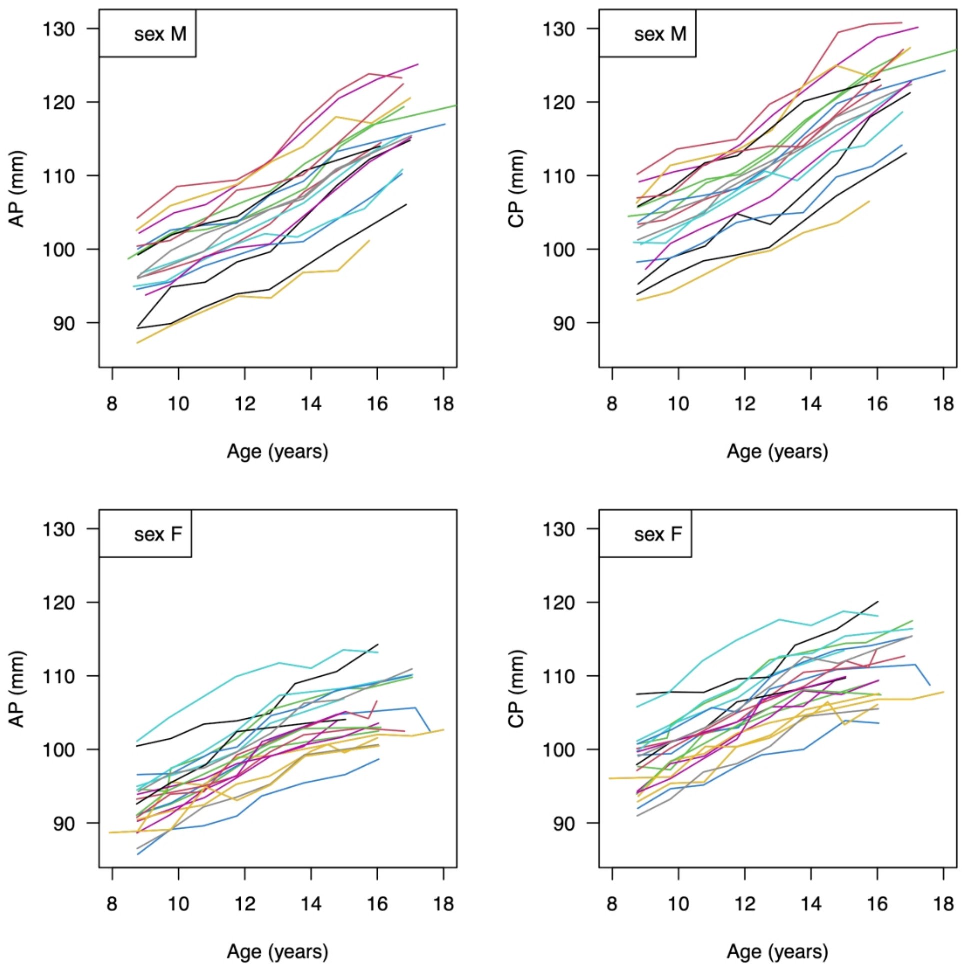

sex M
AP (mm)
Age (years)
sex M
CP (mm)
Age (years)
sex F
AP (mm)
Age (years)
sex F
CP (mm)
Age (years)

Figure 2

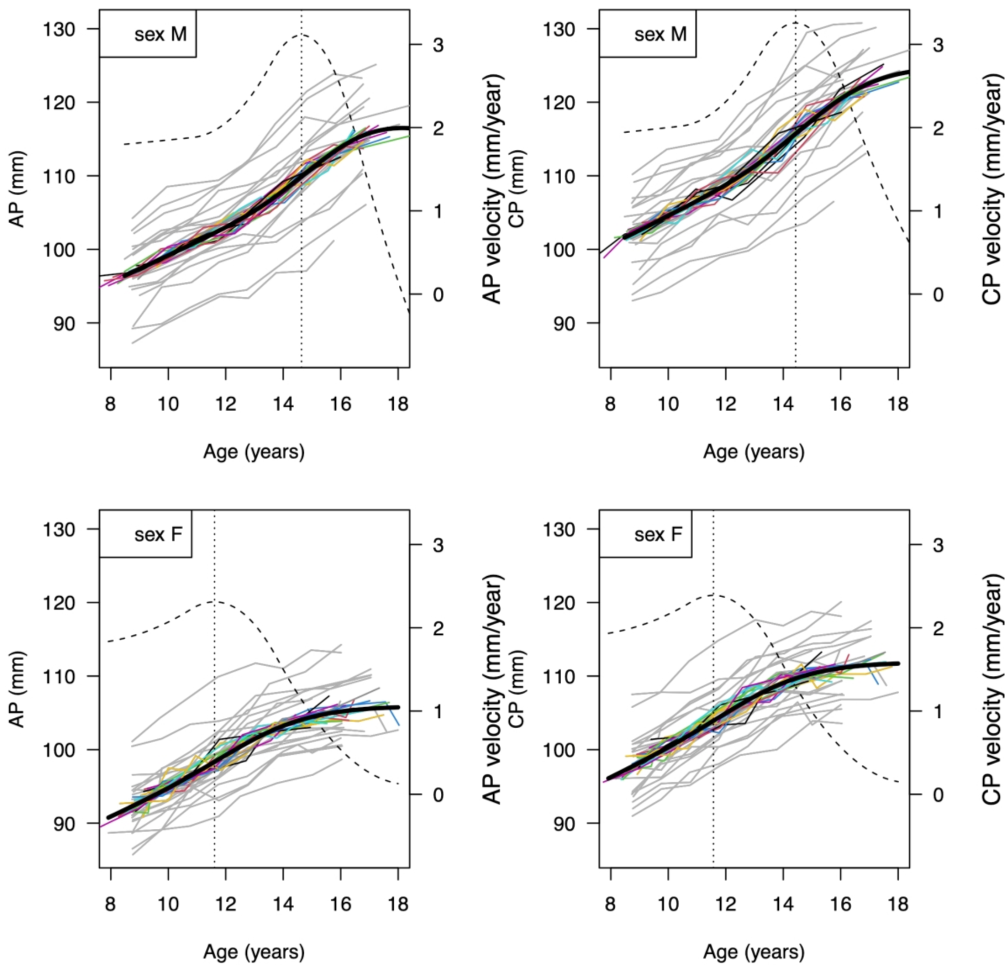

sex M
AP (mm)
AP velocity (mm/year)
Age (years)
sex M
CP (mm)
CP velocity (mm/year)
Age (years)
sex F
AP (mm)
AP velocity (mm/year)
Age (years)
sex F
CP (mm)
CP velocity (mm/year)
Age (years)

Figure 3

Table 1: Summary statistics by sex

	Males (N = 18)		Females (N = 21)	
Measurement	n	Mean (SD)	n	Mean (SD)
Chronological age years	137	12.3 (2.6)	154	12.6 (2.7)
ap mm	137	105.3 (8.4)	154	99.4 (6.2)
cp mm	137	111.2 (8.9)	154	105.1 (6.4)
Bone age years	112	11.8 (2.8)	128	11.9 (2.7)
Interpolated bone age less chronological age years	137	-0.9 (0.9)	145	-0.1 (1.0)

ap: distance articulare-pogonion; cp: distance condylion-pogonion; SD: standard deviation

Table 2: Summary statistics for the four SITAR models versus chronological age

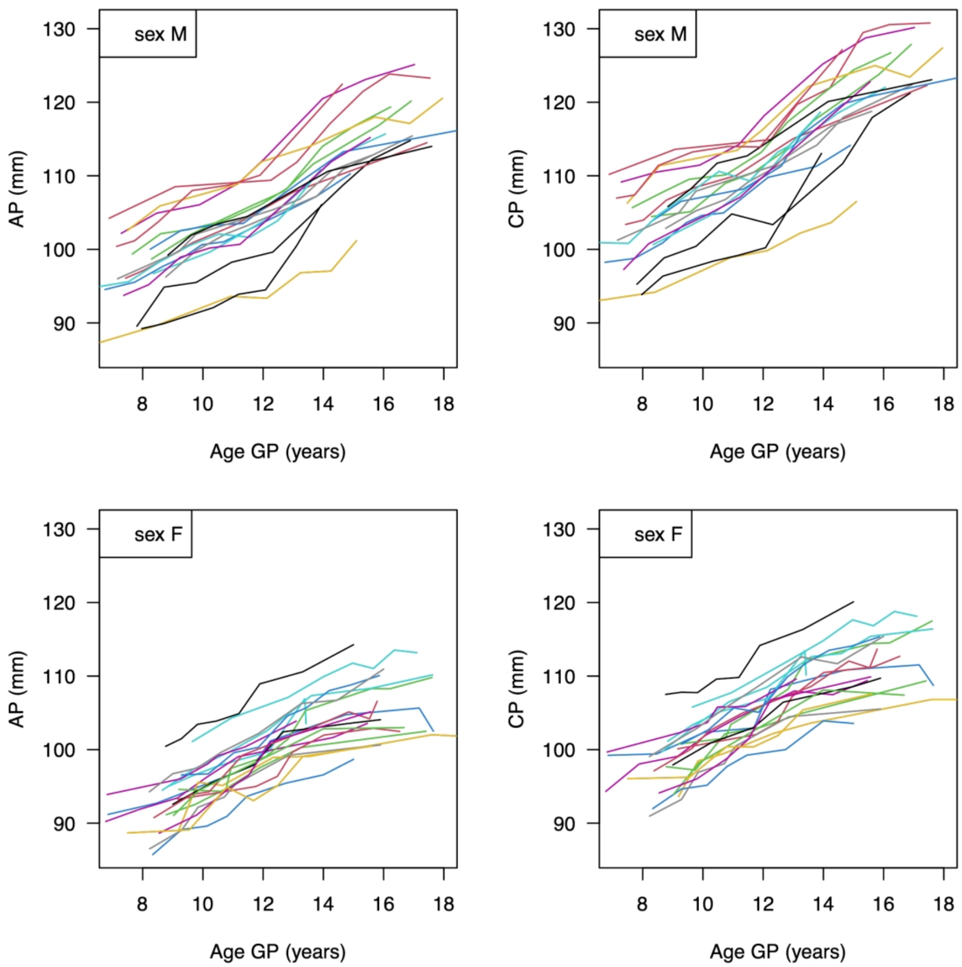
	Males		Females	
Measurement	ap	cp	ap	cp
Mean age at peak velocity years (SE)	14.6 (0.3)	14.4 (0.6)	11.6 (0.4)	11.6 (0.4)
Mean peak velocity mm per year (SE)	3.1 (0.2)	3.3 (0.2)	2.3 (0.1)	2.4 (0.1)
Mean peak velocity % per year (SE)	2.8 (0.2)	2.8 (0.2)	2.4 (0.1)	2.3 (0.1)
SD of size random effect mm	5.6	5.6	3.8	3.9
SD of size random effect %	5.1	4.9	3.7	3.6
SD of intensity random effect mm per year	0.5	0.4	0.3	0.3
SD of intensity random effect %	16	12	14	14
Variance explained by model %	96.5	95.5	92.5	92.8

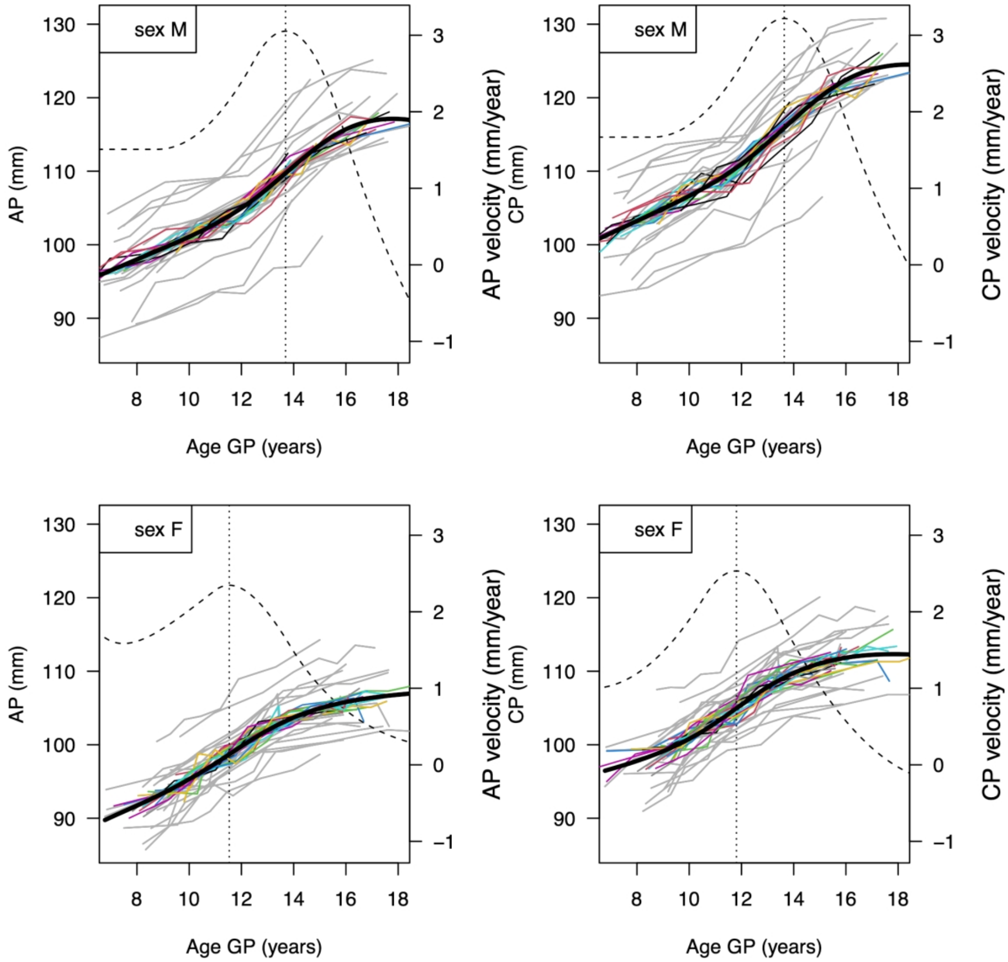
ap: distance articulare-pogonion; cp: distance condylion-pogonion; SD: standard deviation; SE: standard error

Supplementary table 1: Summary statistics for the four SITAR models versus Greulich-Pyle bone age

	Males		Females	
Measurement	ap	cp	ap	cp
Mean age at peak velocity years (SE)	13.7 (0.3)	13.6 (0.4)	11.5 (0.4)	11.8 (0.4)
Mean peak velocity mm per year (SE)	3.1 (0.3)	3.2 (0.3)	2.3 (0.1)	2.5 (0.1)
Mean peak velocity % per year (SE)	2.8 (0.2)	2.8 (0.2)	2.4 (0.1)	2.4 (0.1)
SD of size random effect mm	5.0	4.9	3.5	3.6
SD of size random effect %	4.5	4.2	3.7	3.3
SD of intensity random effect mm per year	0.7	0.6	0.5	0.2
SD of intensity random effect %	25	20	19	7
Variance explained by model %	94.8	92.6	90.4	87.3

ap: distance articulare-pogonion; cp: distance condylion-pogonion; SD: standard deviation; SE: standard error


sex M
AP (mm)
Age GP (years)
sex M
CP (mm)
Age GP (years)
sex F
AP (mm)
Age GP (years)
sex F
CP (mm)
Age GP (years)


sex M
AP (mm)
AP velocity (mm/year)
Age GP (years)
sex M
CP (mm)
CP velocity (mm/year)
Age GP (years)
sex F
AP (mm)
AP velocity (mm/year)
Age GP (years)
sex F
CP (mm)
CP velocity (mm/year)
Age GP (years)