Published in final edited form as:

Int J Dev Sci. 2013 ; 7(2): 105–116. doi:10.3233/DEV-1312120.

Height-based Indices of Pubertal Timing in Male Adolescents

Ammar Khairullah^a, Margaret T. May^b, Kate Tilling^b, Laura D. Howe^b, Gabriel Leonard^c, Michel Perron^{d,e}, Louis Richer^d, Suzanne Veillette^{d,e}, Zdenka Pausova^f, and Tomáš Paus^{a,*}

^aRotman Research Institute and Institute of Medical Science, University of Toronto, Toronto, Canada

^bSchool of Social and Community Medicine, University of Bristol, Bristol, UK

^cMontreal Neurological Institute, McGill University, Canada

^dUniversity of Quebec in Chicoutimi, Chicoutimi, Quebec, Canada

^eECOBES, Research and transfert, CEGEP Jonquière, Jonquière, Quebec, Canada

^fHospital for Sick Children University of Toronto, Toronto, Canada

Abstract

It is important to account for timing of puberty when studying the adolescent brain and cognition. The use of classical methods for assessing pubertal status may not be feasible in some studies, especially in male adolescents. Using data from a sample of 478 males from a longitudinal birth cohort, we describe the calculations of three independent height-based markers of pubertal timing: Age at Peak Height Velocity (APHV), Height Difference in Standard Deviations (HDSDS), and Percent Achieved of Adult Stature (PAAS). These markers correlate well with each other. In a separate cross-sectional study, we show that the PAAS marker correlates well with testosterone levels and self-reported pubertal-stage scores. We conclude by discussing key considerations for investigators when drawing upon these methods of assessing pubertal timing.

Keywords

Males; height; pubertal timing; ALSPAC; Saguenay Youth Study; testosterone

Introduction

Puberty is a period of gonadal maturation and, in turn, the development of secondary sex characteristics (Sizonenko, 1972). Activation of the hypothalamic-pituitary-gonadal axis at the onset of puberty initiates a secretory cascade of gonadal hormones that target numerous organs, including the brain. The age at onset of puberty is not constant across the population. The variation in pubertal timing has been studied in relation to different phenomena occurring during adolescence, including cognition and the emergence of psychopathology.

For example, among college students, late maturers tended to do better than early maturers on mental rotation, a test of spatial ability (Sanders & Soares, 1986). Another study found that early puberty in females, and both very early and late puberty in males, was associated with depression (Kaltiala-Heino, Kosunen, & Rimpelä, 2003). These findings demonstrate that timing of pubertal development is an important variable that should be taken into consideration when studying the adolescent brain.

Pubertal onset and timing of puberty can be studied by assessing pubertal stages, and an individual's progression through them. The assessment of pubertal stages is traditionally achieved through a physical examination, conducted by a trained clinician, through which an individual is classified into one of the five stages of pubertal development often referred to as 'Tanner stages' (Marshall & Tanner, 1970). This mode of assessment is considered to be the gold standard (Dorn, Dahl, Woodward, & Biro, 2006). In large population-based studies, however, measuring pubertal status in this fashion is not feasible, mainly due to the potential participant discomfort associated with the intimate nature of the assessments. This is particularly relevant for male participants as palpation of the testes is an important part of the examination (Dorn et al., 2006). In a study where physical examinations were conducted by a nurse practitioner, 17% of 82 male participants refused to undergo the assessment (Shirtcliff, Dahl, & Pollak, 2009). In lieu of physical examinations, an often-used strategy is the administration of questionnaires answered by the participant or his/her parent.

There are two common ways in which pubertal staging is achieved through such questionnaires. The first involves the use of picture-based questions employing drawings or photographs of children at different Tanner stages. Alternatively, participants are asked a series of questions (without pictures) about their physical growth (height), body hair, skin changes, breast changes, and male-specific questions regarding facial hair and voice changes (Dorn et al., 2006; the Pubertal Development Scale [PDS] by Petersen, Crockett, Richards, & Boxer, 1988). According to Shirtcliff and colleagues (2009), concordance between genital stage assessed through physical exam with either the picture-based questionnaire ($\kappa = 0.36$) or with PDS ($\kappa = 0.36$) appears to be "fair" (as defined by Landis and Koch, 1977). On both types of questionnaires, adolescents overestimated pubertal maturation when they were at lower stages of development compared with their peers, and underestimated development when they were at higher stages relative to their peers (Shirtcliff et al., 2009). The literature suggests that self-reported measures should be used only as crude estimates of pubertal development (Wu, Schreiber, Klementowicz, Biro, & Wright, 2001).

Since growth hormone (GH) plays a key role in regulating height and its production is linked closely with sex steroids, particularly testosterone, the adolescent growth spurt occurs along with pubertal maturation (Rose et al., 1991; Veldhuis, Roemmich, Richmond, & Bowers, 2006). Therefore, in lieu of observing progression through pubertal stages, increases in height can be used to assess pubertal timing in an objective manner. Height measurements have the advantage of being simple, non-invasive, and inexpensive to incorporate into a study design. Furthermore, height can already be found in many datasets, allowing for height-based measures of pubertal timing to be derived retrospectively. Using data from two population-based studies, here we show the utility of height measurements for estimating pubertal timing in males through three different height-based indices; we explore

how these measures compare with each other and with self-reported pubertal stage. Furthermore, we estimate the associations of a height-based measure of pubertal status with questionnaire-based pubertal stage assessment and levels of total testosterone in a separate sample.

The first index, Age at Peak Height Velocity (APHV), identifies the age at which a characteristic feature of puberty, the growth spurt, takes place (Tanner&Whitehouse, 1976). The second index, Height Difference in Standard Deviations (HDSDS), allows us to ascertain pubertal timing relative to the rest of the study sample based on the shift of their height standard-scores from height at 14 years to final height; the mechanism by which this index operates is discussed in the respective methods section (Wehkalampi et al., 2008). The third index incorporates the Khamis-Roche method of predicting adult stature to derive Percent Achieved of Adult Stature (PAAS), which informs us of relative pubertal timing based on how close participants are to their final height predicted by their parental stature and anthropometric measurements (Khamis&Roche, 1994). Table 1 provides an overview of the advantages and disadvantages of some of the approaches used to assess pubertal status and timing.

Methods

Data from two population-based cohorts, the Avon Longitudinal Study of Parents and Children (ALSPAC) and the Saguenay Youth Study (SYS), were used for our analyses. Detailed descriptions of the ALSPAC and SYS datasets have been reported previously (Boyd et al., 2012; Pausova et al., 2007). Ethical approval for ALSPAC was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. The SYS was approved by the local ethics committee. Consent and assent forms were completed by parents/guardians and participants, respectively.

The Avon Longitudinal Study of Parents And Children

ALSPAC is a birth cohort that recruited 14,541 pregnant women resident in Avon, United Kingdom with an expected delivery date between April 1991 and December 1992. Since then the children and their parents have been studied extensively, with longitudinal data available from numerous self-administered questionnaires sent to their homes, linkages to medical records, and through examinations carried out during study clinics.

The children were invited to participate in nine clinic examinations over the course of the study, roughly at the following years of age: 7, 8, 9, 10, 11, 12, 13, 15, and 17. Clinic staff measured height to the last complete mm using a Harpenden Stadiometer, and weight to the nearest 50 g using a Tanita Body Fat Analyser (TBF 305) at each visit. Maternal and paternal heights are available from questionnaires completed by the parents. Outside of the clinics, nine pubertal questionnaires were sent over time by mail to participants' homes inviting them to answer questions about growth of genitalia and pubic hair with the help of pictures corresponding to different Tanner stages. Our sample of 478 males is composed of a subset of the ALSPAC cohort that had been recruited for further study using magnetic resonance imaging. In the next section, we describe the process of deriving three independent height-based indices of pubertal timing from the ALSPAC height data.

Age at Peak Height Velocity (APHV)

Closely spaced longitudinal measurements of height are required for the calculation of Age at Peak Height Velocity (APHV). We have developed a script using MATLAB (Mathworks, Natick, MA USA) to automate the process of APHV estimation. First, the longitudinal measurements of height and the respective age in months are used to plot the participant's height over time. The cubic spline-interpolation function from the MATLAB Basic Fitting toolbox (Mathworks, Natick, MA USA) is used to generate a curve through the data-points (see Supplementary Section; Supplementary Figure 1) as reported previously (Ramsay, Altman, & Bock, 1994; Sherar et al., 2011). Plotting the derivative of this height curve produces the participant's growth curve, and indicates height velocity at each month (Supplementary Figure 2). The maximum height velocity and corresponding age are identified, providing an estimate of the PHV (cm/month) and the APHV (months) respectively. Due to the variations in age at attendance for the clinics, the spacing between ages of observation (in months) varies across participants, with a mean of 15.51, SD of 6.62, and range between 1 and 61 (months). A biologically driven restriction is placed on the script to ensure that the APHV is recalculated if it falls earlier than an age of 10 years, which is under three standard deviations from established norms (Marshall & Tanner, 1970).

Height Difference in Standard Deviations (HSDSDS)

Unlike APHV, calculation of the Height Difference in Standard Deviations (HSDSDS) requires only two height measurements: Height at the average age at Peak Height Velocity for the population and final height, 14 and 17.5 years respectively (Wehkalampi et al., 2008). We interpolated height at exactly 14 years and 17.5 years for all participants using the fitted spline described in the Section on APHV. This step is not necessary if heights were measured exactly at ages 14 and 17.5 years, but this was not the case in our sample. Participants for whom we were missing data from the closest clinic to 14 and 17.5 years were excluded from the analysis. Internally derived standardized scores of height at 14 years (z_{14y}) were calculated for each participant, and the same was done for height at 17.5 years ($z_{17.5y}$). The following formula is then used: $HSDSDS = z_{14y} - z_{17.5y}$.

Percent Achieved of Adult Stature (PAAS)

We used the Khamis-Roche method to predict final stature for each participant based on their age, height, weight and parental stature (Khamis & Roche, 1994). To use this as an index of pubertal timing, we calculated the percentage of adult stature achieved by 14 years of age. The average height of the two parents (mid-parental stature), and both height and weight at age 14 (acquired by using a fitted spline as described in the Section on APHV) were obtained for all participants; height was measured in inches and weight in pounds. These values, along with coefficients provided by Khamis and Roche for males at age 14, were used in the following equation:

$$\text{Predicted adult stature} = -6.4299 + 0.59151 * (\text{height}) - 0.09776 * (\text{weight}) + 0.58762 * (\text{mid-parental stature})$$

We then calculated PAAS using the following equation:

$$PAAS = (\text{Height at 14} / \text{Predicted adult stature}) * 100$$

Interpretation of the Indices of Pubertal Timing

We calculated standardized-scores of APHV for all participants and multiplied these values by -1 . Furthermore, the percentages calculated for all the participants through PAAS were converted into standardized scores. This way, across each index, participants have a score that is on a similar scale and centered around zero, where a score of zero denotes an exactly average maturer. A positive score on an index indicates the participant is a relatively early maturer, whereas a negative score shows that a participant is a late maturer when compared with the rest of the sample.

Using Quintiles to Investigate Concordance of Height-Based Pubertal Timing Indices

To investigate concordance across the three indices beyond correlation coefficients, we ranked and divided values obtained for each index into quintiles. The quintiles were assigned scores from one through five: the first quintile (score of one) contains the lowest scores and indicates the latest maturers while the last quintile (score of five) contains the highest scores and indicates the earliest maturers. These quintiles should not be confused with the five Tanner Stages of pubertal status.

We calculated the percentage of participants with the exact quintile score on the HSDS or PAAS indices when compared with the APHV index. Furthermore, we calculated the percentage of participants who were classified into either the same quintile or a higher or lower adjacent quintile in the HSDS or PAAS indices given their APHV quintile score. Additionally, unweighted kappa and kappa with linear weighting were calculated between APHV and the other two indices.

The Saguenay Youth Study

The SYS is a cross-sectional study of 1,024 adolescents (12 to 18 years of age) of French Canadian origin living in the Saguenay-Lac-Saint-Jean region of Quebec, Canada. In brief, the primary focus of SYS is to evaluate the long-term consequences of prenatal exposure to maternal cigarette smoking. Both exposed and non-exposed adolescents, matched on maternal education and school attended were recruited through local high schools. All participants were assessed using an extensive phenotyping protocol, including magnetic resonance imaging (MRI) of the brain and body, to characterize a broad range of phenotypes relevant for their mental, cardiovascular, and metabolic health (see Pausova et al., 2007 for details). Importantly for our analyses, anthropometric measurements, scores on Pubertal Development Scale, and plasma levels of sex hormones were also obtained. Data from the 496 male participants were used in our analyses.

Anthropometry and PAAS Scores

As the SYS data are cross-sectional, only the PAAS can be derived. Height, weight, and mid-parental stature along with the appropriate age-based coefficients, available for each half year from 12 years to 17.5 years, for males were used to estimate adult stature for each participant using the Khamis-Roche method as described above. PAAS scores were calculated and transformed into standardized z-scores.

Pubertal Development Scale (PDS)

All participants filled out the Pubertal Development Scale (Petersen et al., 1988), an eight-item self-reported measure of physical development based on the Tanner stages. The male participants answer questions about their growth in stature, pubic hair, and voice changes which are used for classification into one of five categories of pubertal status: (1) prepubertal, (2) beginning pubertal, (3) midpubertal, (4) advanced pubertal, and (5) postpubertal.

Serum Testosterone

Fasting blood samples were taken in the morning (between 8 : 00 and 9 : 00 A.M.) and analyzed via radioimmunoassay (Testosterone RIA DSL- 4000; Diagnostic Systems Laboratory) to measure serum levels of total testosterone (ng/ml).

Statistical Analyses

Correlations were used to examine the relationships between different indices of height-based pubertal timing obtained in ALSPAC. Fixed-effects ANOVA models were used to examine the relationship between data from the ALSPAC pubertal questionnaire and the three indices of pubertal timing. In the Saguenay Youth Study, linear regression was used to relate PAAS to total testosterone. Fixed-effects ANOVA models were used to examine PAAS scores and testosterone levels across the five PDS-based stages and the responses to the pubic-hair question on the PDS questionnaire. The pubic-hair question was selected from the PDS for comparison with an analogous question present in the ALSPAC questionnaires. All statistical analyses were done using JMP 9 for Macintosh (SAS Institute, Inc., Cary, NC).

Results

Correlation Between Three Height-Based Pubertal Timing Indices in ALSPAC

Our subset from the Avon Longitudinal Study of Parents and Children consisted of 478 male participants. Of these, a total of 18 were excluded: 8 for missing four or more height measurements, 1 for having an anomalous height measurement, and 9 for very high Age at Peak Height Velocity estimates that are imprecise due to spline interpolation at the far extreme of the height measurements. This left a remaining sample of 460 participants. Comparison of means with two-tailed *t*-tests between excluded participants and the remaining sample show that the 18 excluded participants had shorter stature in the four clinics between 12 and 17 years of age ($r^2 = 0.02, 0.01, 0.04, 0.01$; all $p < 0.05$). This effect is present because the excluded group contains the nine late maturers with imprecise APHV estimates.

Note that from the remaining 460 participants, we recalculated APHV for the 53 participants whose APHV estimates lay below 10 years. We used a lower limit of 10 years, as a growth spurt before this age is unlikely (Marshall&Tanner, 1970). The mean recalculated APHV for these participants was 161.42 months, which matches closely the mean APHV for the full sample of 160.26 months.

For the calculation of Height Difference in Standard Deviations score, 34 participants who were missing important data from the clinics adjacent to ages 14 and 17.5 were excluded from this particular analysis (see section on *HDSDS*), leaving 426 participants in whom we could calculate *HDSDS* scores. Similarly, after excluding participants with missing data required for the Khamis-Roche calculations (mainly due to missing maternal and/or paternal height), we were able to calculate Percent Achieved of Adult Stature scores in 354 participants.

Table 2 presents statistics describing the distribution of scores from each index of pubertal timing, and pairwise correlations between the three indices; APHV and PAAS were first standardized (z scored), and APHV scores multiplied by -1. All three indices are strongly correlated; the strongest correlation is found between the APHV and *HDSDS* scores ($n = 426$; $r = 0.76$, $p < 0.0001$, 95% Confidence Interval [CI] 0.72 – 0.80). The correlation between APHV and PAAS is also good ($n = 354$; $r = 0.59$, $p < 0.0001$, 95% CI 0.52 – 0.66) as is the correlation between *HDSDS* and PAAS ($n = 325$; $r = 0.63$, $p < 0.0001$, 95% CI 0.56 – 0.69).

Concordance of Indices using Quintiles

Table 3 presents the ranges contained within the quintiles described in the respective section on all three indices.

As can be seen in Fig. 1, correct classification into the corresponding or neighboring APHV quintile occurred approximately 90% across each quintile for *HDSDS*, and approximately 80% for PAAS. Kappa was calculated for each comparison: for APHV and *HDSDS*, the unweighted ($\kappa = 0.41$; $p < 0.0001$) and linear-weighted kappas ($\kappa = 0.64$; $p < 0.0001$) indicate, respectively, moderate and substantial agreement (Landis & Koch, 1977). For APHV and PAAS, the unweighted ($\kappa = 0.21$; $p < 0.0001$) and linear-weighted kappas ($\kappa = 0.43$; $p < 0.0001$) indicate, respectively, fair and moderate agreement.

Association of the Percent Achieved of Adult Stature Index with Measures of Pubertal Status

Data from the Saguenay Youth Study were used for these analyses. Of the 496 participants, a total of 70 were excluded: 38 participants were over the age of 17.5 years and could not be included because the Khamis-Roche coefficients were not estimated beyond this age, 30 participants were missing at least one parental height, 1 participant did not have height and weight recorded, and 1 participant had an anomalous maternal height. Furthermore, two participants did not have PDS scores recorded and thus were not included in the analyses requiring PDS data. Comparison of excluded participants with the remaining sample shows that the 70 excluded participants have higher weight, height, pubertal status, and paternal height ($r^2 = 0.02, 0.02, 0.04, 0.01$; all $p < 0.05$), but not maternal height; this is consistent with the 17.5 + age of 38/70 of the excluded participants. Only 216 participants had testosterone data; the remaining samples are yet to be analyzed. Table 4 presents the sample mean and standard deviations of important variables used in calculation of PAAS and further analyses.

Figure 2A shows that standardized scores from the PAAS index correlate strongly with testosterone levels in male adolescents ($n = 216$; $r = 0.73$; $p < 0.0001$; 95% CI 0.66 – 0.79), thus explaining over half of the variance ($r^2 = 0.53$). Fig. 2B depicts the ANOVA comparing PAAS scores across PDS-based stages ($n = 424$; $r^2 = 0.45$; $p < 0.0001$). We see a similar association between pubic-hair stage and PAAS ($n = 424$; $r^2 = 0.44$; $p < 0.0001$). The ANOVA comparing testosterone levels across PDS-based stages ($n = 214$; $r^2 = 0.44$; $p < 0.0001$) shows a strong relationship, as does testosterone across pubic-hair stages ($n = 214$; $r^2 = 0.38$; $p < 0.0001$).

Comparing Performance of the Pubertal Timing Indices in the Two Studies

We used ANOVA models to assess performance of the three indices of pubertal timing in ALSPAC in relation to the Genitalia and Pubic-Hair stages ascertained with the fifth mailed pubertal questionnaire, which was completed at approximately 13.5 years of age (range: 13.1–13.8 years). Pubic-Hair stage consistently outperformed Genitalia stage in explaining variance in height-based measurements, often by more than 10%: APHV by Pubic-Hair stage ($n = 338$, $r^2 = 0.38$; $p < 0.0001$); APHV by Genitalia stage ($n = 312$, $r^2 = 0.22$; $p < 0.0001$); HSDSD by Pubic-Hair stage ($n = 314$, $r^2 = 0.43$; $p < 0.0001$); HSDSD by Genitalia stage ($n = 291$, $r^2 = 0.27$; $p < 0.0001$); PAAS by Pubic-Hair stage ($n = 268$, $r^2 = 0.31$; $p < 0.0001$); PAAS by Genitalia stage ($n = 247$, $r^2 = 0.18$; $p < 0.0001$).

To obtain an analogous comparison between questionnaire-based stages and pubertal timing index in SYS within a similar narrow age range of the respondents, as with the ALSPAC questionnaire, we restricted our analysis shown in Fig. 2B to participants between the ages of 13 and 14 years. This ANOVA between PAAS and PDS scores using the narrowed SYS age group ($n = 118$; $r^2 = 0.22$; $p < 0.0001$) showed that the PDS scores in SYS performed slightly better than the ALSPAC questionnaire Genitalia Stage but not as well as staging based on Pubic Hair, which were shown above to explain 18% and 31% of variance, respectively. The ANOVA between pubic-hair stage in SYS and PAAS using the narrowed age group ($n = 118$; $r^2 = 0.25$; $p < 0.0001$) explains slightly less variance than the analogous question in ALSPAC, which explains 31% of variance.

Discussion

As shown with the ALSPAC data, the three height-based indices of pubertal timing correlate well with each other, indicating that they capture similar information about the timing of puberty from height data. When we categorized index scores into quintiles, we found moderate to substantial concordance with linear-weighted kappa between the three different methods. For an assessment of relative pubertal timing between participants, exact concordance between the indices may not be essential; we are able to distinguish early from late maturers if the score from one index falls into an equal or adjacent quintile in another index. For instance, we are able to infer that a participant is a relatively early maturer if he receives scores of 4, 4, and 5 on the three indices, but not if he scores – for example – 1, 3, and 5.

We compared quintile scores from Height Difference in Standard Deviations and Percent Achieved of Adult Stature against the Age at Peak Height Velocity quintile scores, which

were used as the standard. APHV scores are likely the most reliable given that its derivation takes into account all the longitudinal height data available from the participant. The HDSDS takes into account just two height data-points from the individual, and PAAS uses only one. The average percentage of participants' HDSDS and PAAS scores correctly classified within one quintile on APHV is 91.75% and 79.59%, respectively. Thus we are confident that a participant's pubertal timing as indicated by the HDSDS or PAAS indices will likely match what is obtained from APHV.

Association of the PAAS Index with Measures of Pubertal Status

It would be interesting to compare the three indices of pubertal timing with pubertal stage assessed through clinical examinations in a future study containing the appropriate data to see how well these height-based measures perform when compared to the gold standard. This would also answer the question as to which of the three indices most closely matches physical characteristics of puberty, including genitalia and pubic hair growth, assessed through the clinical examination.

Until we are able to answer these questions it is reassuring to know that the height-based indices correlate with each other as shown in ALSPAC; between 35% and 58% of variance in each index is explained by the other indices. Furthermore, the SYS data illustrate that PAAS is indicative of pubertal timing: it explains 53% of the variance in testosterone levels and 45% of variance in the PDS scores. By extension, the APHV and HDSDS indices likely also reflect changes in testosterone and PDS levels given their strong correlation with the PAAS; this will be verified in the future. It is important to note that the strength of the association between the various height-based indices of pubertal timing and characteristics of pubertal development (e.g., genitalia, pubic hair) would vary with age, likely showing an inverted U-shaped relationship. The indices based on physical growth will not distinguish signs of pubertal development at pre-pubertal and post-pubertal ages when there is little to no variance in pubertal stages. They will match pubertal development closest between the ages of 13 to 15 years, when there is most variance in the development of secondary sexual characteristics in the male population. We have described the process used to calculate each index, and will now discuss the background, interpretations, and key considerations in the use of the three pubertal timing indices.

Age at Peak Height Velocity

In 1976, Tanner published norms for APHV in males (Tanner&Whitehouse, 1976). Since then, many studies have used a number of different methods to calculate APHV, ranging from simple mid-year estimations (e.g., Lindgren, 1976) to spline interpolants (e.g., Sherar et al., 2011). Our script is automated and efficient, allowing for robust and consistent estimates in large datasets. In order to obtain a good estimate of APHV, it is important to obtain numerous closely spaced measurements of height between the ages of 10 and 17 years. At least five height measurements in this time period spaced no more than two years apart should be used. Naturally, additional measurements spaced closer together will yield more precise estimates of APHV.

A noteworthy observation about our APHV estimates is that the mean lies at approximately 13.3 years. This contrasts with the reported average estimate of 14 years from Tanner (Tanner & Whitehouse, 1976). Secular trends towards earlier female maturation have been studied (Euling et al., 2008) and are generally accepted. An analogous pattern towards earlier pubertal onset in males has been recently identified (Herman-Giddens et al., 2012). Our APHV estimates are congruent with claims of a secular trend towards earlier maturation in males.

Height Difference in Standard Deviations

Calculation of HSDS only requires two height data-points, height at 14 years of age and final height (Wehkalampi et al., 2008). Since adult stature in males is not achieved until past age 20 and these data are not yet available, we used height at 17.5 years at which time most males have slowed their annual growth to less than 1 cm (Roche & Davila, 1972). This index informs us of pubertal timing in a participant by describing the shift of their rank in the population based on height from an age when only half of the population has experienced a growth spurt compared with an age when growth has ceased. The logic is as follows: if an individual is an early maturer, he would have experienced a growth spurt already at 14 years, the average age for peak height velocity in the population (Marshall & Tanner, 1970). When comparing his rank in the population based on height at age 14 to final height, it should be apparent that his rank has decreased since the rest of the population caught up to him in growth. A late maturer, on the other hand, will have a lower rank at age 14 but a relatively higher rank in final height due to the fact that he will experience a growth spurt in the interim. This change in relative position within the population is characterized by the HSDS: a positive score indicates early maturation and a negative score indicates late maturation. Although the average APHV in our sample is slightly lower than that of 14 years reported elsewhere (Marshall & Tanner, 1970), the strong correlation we find between APHV and HSDS indicates that the use of height at 14 years does not adversely affect this index.

Percent Achieved of Adult Stature

The Khamis-Roche method allows the prediction of an individual's adult stature by using age and sex dependant coefficients along with the individual's current height, weight, and mid-parental stature (Khamis & Roche, 1994). Through dividing current height by the Khamis-Roche prediction, we calculate the individual's Percent Achieved of Adult Stature. Participants who mature early would have attained a higher PAAS at 14 years of age when compared with their late maturing counterparts. This method will be most useful between the ages of 13 to 15 where the variation of pubertal stages across the population is at a maximum.

It is important to note that the original paper from Khamis and Roche contained incorrectly recorded coefficients; therefore the revised paper containing the corrected coefficients should be used (Erratum in: *Pediatrics*. 1995; 95 : 457). There are a number of similarities and differences between the sample used to derive the coefficients and the ALSPAC and SYS studies. The sample used by Khamis and Roche, from the Fels Longitudinal Study, comprises of Caucasians from a developed country, under represents low socioeconomic

(SES) groups, and has recruited participants since 1929. These participants may have experienced puberty as early as the 1930 s up until the 1990 s. Both the ALSPAC and SYS studies are also composed of Caucasian participants (>96% in ALSPAC, 100% in SYS) from a developed country, with the participants going through puberty in the first decade of the 2000 s. The ALSPAC mothers have, on average, indicators of higher SES than the rest of the UK, whereas the SYS sample over-represents lower SES groups (due to its focus on smoking during pregnancy).

A crucial assumption inherent in the Khamis-Roche method used for derivation of the PAAS is that height is a highly heritable trait. The heritability of height in males ranges by country from .87 to .93 in the large twin samples from the GenomEUtwin study (Silventoinen et al., 2003). Nonetheless, it is important to exercise caution when applying the Khamis-Roche coefficients across a sample where some children and parents may have grown up in different environments. For example, immigration may lead to children being reared in different nutritional environments from their parents, and as a consequence parental height might be a weaker predictor of final height of the offspring. As such, the Khamis-Roche method and consequently the PAAS method are best suited for populations where both parents and children were exposed to similar non-genetic effects, as is seen in the Saguenay Youth Study sample.

Comparison of PAAS to Pubertal Questionnaires

The questionnaires used to assess pubertal timing in ALSPAC and SYS are notably different in the use of pictures. The ALSPAC questionnaires contained four questions that asked about voice changes, armpit hair, pubic hair, and genital growth; the latter two questions were accompanied by drawings of the five Tanner stages to aid participants in answering accurately. Conversely, the PDS questionnaire used by the SYS asked questions about growth in height, presence of body and facial hair, acne, and voice changes. There were no pictures, nor any questions pertaining to genitalia. Answers from this set of questions are combined into a composite score that allows estimation of Tanner stage.

In analyses with questionnaire-derived measures of pubertal stage and the pubertal timing indices, the Pubic Hair question from ALSPAC performed the best. The PDS score from SYS also performed well. The genitalia staging from ALSPAC came third. One issue that the ALSPAC questionnaires suffered from was missing data. The prevalence of missing data ranged from 15% to 36% over the various pubertal questionnaires sent out to participants. Participants may potentially be uncomfortable with answering questions about genitalia as evidenced by patterns of missing data in the five questionnaires sent after age 13. The genitalia questions are left unanswered 5%–7% more often than the pubic hair questions. The PDS questionnaire presents a viable alternative for use in studies, as it does not contain a question asking about genitalia. Missing data did not play a role in our subset of the SYS data, where less than half a percent of participants failed to complete the PDS questionnaire.

Other Considerations

It may not be the case that height data from exactly age 14 or 17.5 is available for all participants. Despite having a planned target age, each ALSPAC clinic measured

participants in a wide age range. For example, the minimum, average, and maximum ages at the 9-year clinic were 8.75, 9.9, and 11.67 years, respectively. Since ALSPAC contained longitudinal data, we were able to use spline interpolation to estimate height at the exact ages of 14 and 17.5 years for subsequent HSDS and PAAS calculations.

Although we use population-based cohorts to demonstrate the derivation and correspondence between these indices, they can be used in smaller samples. This is because the scores of each index can be standardized, and thus provide information on a participant's relative timing of maturation compared with the rest of the sample. The data provided in Tables 2 and 3 may prove useful as a reference. We hope that other investigators will find our descriptions helpful in order to account for pubertal timing when studying the adolescent brain.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors would like to thank the following for their assistance: Drs. Mallar Chakravarty, Erin Dickie, Kate Northstone, Mark Palmert, and Melissa Pangelinan. AK received student financial support from the Ontario Graduate Scholarship and the Max and Ruth Wiseman Graduate Student Fellowship. TP is the Tanenbaum Chair in Population Neuroscience at the Rotman Research Institute, University of Toronto.

The Saguenay Youth Study project is funded by the Canadian Institutes of Health Research (TP, ZP), Heart and Stroke Foundation of Quebec (ZP), and the Canadian Foundation for Innovation (ZP). We thank all families who took part in the Saguenay Youth Study and the following individuals for their contributions in acquiring data: Manon Bernard (database architect, The Hospital for Sick Children), Jacynthe Tremblay and her team of research nurses (Saguenay Hospital), Helene Simard and her team of research assistants (Cègep de Jonquière) and Rosanne Aleong (program manager, Rotman Research Institute).

The UK Medical Research Council and the Wellcome Trust (Grant Ref: 092731) and the University of Bristol provide core support for ALSPAC. We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. LDH is funded by a UK Medical Research Council Population Health Scientist fellowship (G1002375). This publication is the work of the authors and Dr. Margaret May will serve as guarantor for the contents of this paper. This research was funded by NIH, 5R01MH085772 (TP).

References

- Boyd A, Golding J, Macleod J, Lawlor D, Fraser A, Henderson J, et al. Cohort Profile: The “Children of the 90s”. The index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*. 2012
- Dorn LD, Dahl RE, Woodward HR, Biro F. Defining the boundaries of early adolescence: A user's guide to assessing pubertal status and pubertal timing in research with adolescents. *Applied Developmental Science*. 2006; 10:30–56.
- Euling SY, Herman-Giddens ME, Lee PA, Selevan SG, Juul A, Sørensen TIA, et al. Examination of US puberty-timing data from 1940 to 1994 for secular trends: Panel findings. *Pediatrics*. 2008; (121 Suppl):S172–S191. [PubMed: 18245511]
- Herman-Giddens ME, Steffes J, Harris D, Slora E, Hussey M, Dowshen SA, et al. Secondary sexual characteristics in boys: Data from the pediatric research in office settings network. *Pediatrics*. 2012; 130:1058–1068.
- Huang B, Hillman J, Biro FM, Ding L, Dorn LD, Susman EJ. Correspondence between gonadal steroid hormone concentrations and secondary sexual characteristics assessed by clinicians,

- adolescents, and parents. *Journal of Research on Adolescence*. 2012; 22:381–391. [PubMed: 23204809]
- Hundley, D. *Cubic Splines in Matlab*. 2009.
- Kaltiala-Heino R, Kosunen E, Rimpelä M. Pubertal timing, sexual behaviour and self-reported depression in middle adolescence. *Journal of Adolescence*. 2003; 26:531–545. [PubMed: 12972267]
- Khamis H, Roche A. Predicting adult stature without using skeletal age: The Khamis-Roche method. *Pediatrics*. 1994; 94:504–507. [PubMed: 7936860]
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159–174. [PubMed: 843571]
- Lindgren G. Height, weight and menarche in Swedish urban school children in relation to socio-economic and regional factors. *Annals of Human Biology*. 1976; 3:501–528. [PubMed: 999228]
- Marshall WA, Tanner JM. Variations in the pattern of pubertal changes in boys. *Archives of Disease in Childhood*. 1970; 45:13–23. [PubMed: 5440182]
- Pausova Z, Paus T, Abrahamowicz M, Almerigi J, Arbour N, Bernard M, et al. Genes, maternal smoking, and the offspring brain and body during adolescence: Design of the Saguenay Youth Study. *Human Brain Mapping*. 2007; 28:502–518. [PubMed: 17469173]
- Petersen A, Crockett L, Richards M, Boxer A. A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*. 1988; 17:117–133. [PubMed: 24277579]
- Ramsay JO, Altman N, Bock RD. Variation in height acceleration in the Fels growth data. *Canadian Journal of Statistics*. 1994; 22:89–102.
- Roche AF, Davila GH. Late adolescent growth in stature. *Pediatrics*. 1972; 50:874–880. [PubMed: 4673900]
- Rose SR, Municchi G, Barnes KM, Kamp GA, Uriarte MM, Ross JL, et al. Spontaneous growth hormone secretion increases during puberty in normal girls and boys. *The Journal of Clinical Endocrinology & Metabolism*. 1991; 73:428–435. [PubMed: 1856268]
- Sanders B, Soares MP. Sexual maturation and spatial ability in college students. *Developmental Psychology*. 1986; 22:199–203.
- Schmitz KE, Hovell M, Nichols J, Irvin V, Keating K, Simon G, et al. A validation study of early adolescents' pubertal self-assessments. *The Journal of Early Adolescence*. 2004; 24:357–384.
- Sherar LB, Eisenmann JC, Chilibeck PD, Muhajarine N, Martin S, Bailey DA, Baxter-Jones ADG. Relationship between trajectories of trunk fat mass development in adolescence and cardiometabolic risk in young adulthood. *Obesity*. 2011; 19:1699–1706. [PubMed: 21273991]
- Shirtcliff E, Dahl R, Pollak S. Pubertal development: Correspondence between hormonal and physical development. *Child Development*. 2009; 80:327–337. [PubMed: 19466995]
- Silventoinen K, Sammalisto S, Perola M, Boomsma DI, Cornes BK, Davis C, et al. Heritability of adult body height: A comparative study of twin cohorts in eight countries. *Twin Research*. 2003; 6:399–408. [PubMed: 14624724]
- Sizonenko P. Endocrinology in preadolescents and adolescents. *American Journal of Diseases in Children*. 1972; 132:704–712.
- Tanner JM, Whitehouse RH. Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. *Archives of Disease in Childhood*. 1976; 51:170–179. [PubMed: 952550]
- Veldhuis JD, Roemmich JN, Richmond EJ, Bowers CY. Somatotrophic and gonadotropic axes linkages in infancy, childhood, and the puberty-adult transition. *Endocrine Reviews*. 2006; 27:101–140. [PubMed: 16434512]
- Wehkalampi K, Silventoinen K, Kaprio J, Dick DM, Rose RJ, Pulkkinen L, Dunkel L. Genetic and environmental influences on pubertal timing assessed by height growth. *American Journal of Human Biology*. 2008; 20:417–423. [PubMed: 18293372]
- Wu Y, Schreiber GB, Klementowicz V, Biro F, Wright D. Racial differences in accuracy of self-assessment of sexual maturation among young black and white girls. *The Journal of Adolescent Health*. 2001; 28:197–203. [PubMed: 11226842]

Biographies

Ammar Khairullah, MSc student, Institute of Medical Science at the University of Toronto. Research into the effects of testosterone levels and pubertal timing on white matter properties in the male adolescent brain.

Margaret T. May, PhD, Reader in Medical Statistics, School of Social and Community Medicine, University of Bristol, UK. Research interests in statistical methodology including prognostic modelling, longitudinal data analysis, and methods for missing data and causal inference from observational studies. Health services research includes clinical epidemiology of HIV, chronic fatigue syndrome, renal disease and biomarkers for cancer severity.

Kate Tilling, PhD, Professor of Medical Statistics since 2011 School of Social and Community Medicine, University of Bristol. Research on modelling of growth and its relation to subsequent outcomes, and the statistical modelling of repeated measures over the lifecourse.

Laura D. Howe, PhD, MRC Population Health Scientist Research Fellow since September 2011, MRC Centre for Causal Analyses of Translational Epidemiology, School of Social and Community Medicine, University of Bristol. Research on life course epidemiology and child growth.

Gabriel Leonard is a clinical scientist at the Montreal Neurological Institute. Interests include examining long-term consequences of prenatal exposure to cigarette smoking on function and mental health in adolescence, MR imaging of multiple sclerosis and associated cognitive profiles, and the development of a new computerized device to measure hand and arm movements.

Michel Perron, PhD, Professor at University of Quebec in Chicoutimi (Quebec, Canada) since 2008 and Researcher in Sociology of Education and Geography of Health since 1982. Director of the Chair UQAC-Cegep of Jonquiere on Life Conditions, Health and Aspirations of Youth. Research on Adolescents (Habits of Life, Educational Aspirations), on Inequalities of Students Pathways and on School Drop out.

Louis Richer, PhD, Professor of Psychology, Département des sciences de la santé and director of the Laboratoire sur l'adaptation personnelle, sociale et neuropsychologique (LAPERSONNE), Université du Québec à Chicoutimi, Québec, Canada. Research interests in neuropsychology concerning links between brain development, cognitive and executive functions, and adaptation of individuals and relatives exposed to environmental constraints, transition periods, and chronic/acute stressors.

Suzanne Veillette, PhD, Researcher in Sociology of Education at ECOBES, CEGEP of Jonquiere, Quebec since 1982 and, since 2010, Associate Professor, University of Quebec in Chicoutimi, Chicoutimi, Quebec, Canada. Research on Adolescents (Habits of Life,

Educational Aspirations, Transitions and Inequalities in Students Pathways: Success, Persistence).

Zdenka Pausova, M.D. is a Scientist in The Hospital for Sick Children and an Associate Professor of Physiology and Nutritional Sciences at the University of Toronto. She studies genetic and environmental modifiers of cardio-metabolic health in adolescents.

Tomáš Paus, M.D., PhD, is the Tanenbaum Chair in Population Neuroscience and Professor of Psychology and Psychiatry at the University of Toronto. He applies brain-mapping tools in population-based studies investigating brain-behaviour relationships during adolescence. In this work, he and his colleagues explore the interplay between genes and environment in shaping the adolescent brain.

Figure 1. Occurrence of correct classification into the corresponding or neighboring APHV quintile.
 +: % Participants within the same or adjacent quintile on HDSDS and APHV. X: % Participants within the same or adjacent quintile on PAAS and APHV. O: % Participants within the same HDSDS and APHV quintiles. ◇ Participants within the same PAAS and APHV quintiles.

Figure 2.

a) Linear regression shows that 53% of variance in testosterone is explained by Percent Achieved of Adult Stature; b) Analysis of variance shows that 45% of variance in Percent Achieved of Adult Stature is explained by Pubertal Development Scale scores.

Table 1

Overview of Measures Used to Assess Pubertal Status and Timing

Measure	Cost	Invasiveness	Minimum timepoints needed	Accuracy ¹	Precision ²
Physical Exam	High	High	1	High ³	High ³
Hormones	High	High (blood) Low (saliva)	1	High ³	High ³
Self-Report	Low	Low	1	Medium ³	Medium ³
Questionnaire					
APHV	Medium	Low	5	Medium	High
HDSDS	Medium	Low	2	Medium	High
PAAS	Medium	Low	1	Medium	High

¹The accuracy of a measurement is the degree of closeness of measurements of a quantity to that quantity's true value.

²The precision of a measurement is the degree to which repeated measurements under unchanged conditions show the same results.

³Huang et al., 2012; Schmitz et al., 2004; Shirtcliff et al., 2009.

Descriptive Statistics, Distribution Characteristics, and Correlations of Pubertal Timing Indices in the Avon Longitudinal Study of Parents and Children

Table 2

Index	n	Mean	(SD)	Range	Skewness	Kurtosis	Correlation with APHV	Correlation with HDSDS	Correlation with PAAS
APHV (months)	460	160.26	13.04	(120, 189)	-0.32	0.24	1	0.76	0.59
HDSDS	426	-0.015	0.82	(-2.18, 1.85)	-0.25	-0.58	0.76	1	0.63
PAAS at 14y (%)	354	93.02	2.18	(83.6, 99.9)	-0.3	0.61	0.59	0.63	1

Note that the correlations reported use standard scores of APHV and PAAS (see Results).

Table 3

Range of Quintiles for each Pubertal Timing Index

Index	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
APHV (z-score)	(-1.92, -0.89)	(-0.89, -0.07)	(-0.07, 0.13)	(0.13, 0.75)	(0.75, 2.80)
APHV (months)	(189, 174)	(174, 162)	(162, 159)	(159, 150)	(150, 120)
HSDS	(-2.18, -0.79)	(-0.76, -0.25)	(-0.24, 0.23)	(0.23, 0.70)	(0.71, 1.85)
PAAS (z-score)	(-4.26, -0.92)	(-0.92, -0.27)	(-0.26, 0.32)	(0.34, 0.78)	(0.79, 3.12)
PAAS (percent)	(83.6, 91.0)	(91.0, 92.4)	(92.4, 93.7)	(93.7, 94.7)	(94.7, 99.9)

Table 4
Descriptive Statistics and Distribution Characteristics of the Saguenay Youth Study Sample

	n	Mean	(SD)	Range	Skewness	Kurtosis
Age (months)	426	175.96	18.64	(144, 212)	0.14	-1.07
Height (cm)	426	166.24	10.65	(140.5, 188.5)	-0.35	-0.56
Weight (kg)	426	60.41	16.56	(28.8, 153.5)	1.07	2.77
Maternal height (cm)	426	162.88	6.27	(147.3, 177.8)	0.09	-0.5
Paternal height (cm)	426	175.18	6.48	(151.1, 203.2)	0.04	0.92
PAAS (%)	426	94.18	5.00	(81.9, 101.0)	-0.52	-0.84
PDS	424	3.27	0.81	(1, 5)	-0.42	-0.11
Testosterone (ng/ml)	216	5.06	2.60	(0.1, 11.3)	-0.18	-0.61