DATA NOTE

Using SITAR (SuperImposition by Translation and Rotation) to estimate age at peak height velocity in Avon Longitudinal Study of Parents and Children [version 1; referees: 1 approved]

Monika Frysz\textsuperscript{1,2}, Laura D. Howe\textsuperscript{1,2}, Jonathan H. Tobias\textsuperscript{1,3}, Lavinia Paternoster\textsuperscript{1,2}

\textsuperscript{1}MRC Integrative Epidemiology Unit at the University of Bristol, University of Bristol, Bristol, UK
\textsuperscript{2}Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
\textsuperscript{3}Musculoskeletal Research Unit, University of Bristol, Bristol, UK

Abstract
Puberty is a time of substantial biological and psychological changes. One of the hallmarks of puberty is a rapid growth spurt, however its timing varies between individuals. The impact of pubertal timing on later health outcomes has been of interest in life course epidemiology, however its measurement can be challenging. Age at peak height velocity (aPHV) offers an objective measure of pubertal timing without having to rely on physical examination or self-report. We describe the derivation of aPHV estimates in Avon Longitudinal Study of Parents and Children (ALSPAC) offspring, using Superimposition by Translation And Rotation (SITAR) mixed effects growth curve analysis. ALSPAC is a rich source of phenotypic and genotypic data and given the importance of pubertal timing for later health outcomes, these data offer an opportunity to explore the determinants and consequences of aPHV.

Keywords
ALSPAC, pubertal timing, growth, age at peak height velocity

This article is included in the Avon Longitudinal Study of Parents and Children (ALSPAC) gateway.
Corresponding author: Monika Frysz (monika.frysz@bristol.ac.uk)

Author roles: Frysz M: Data Curation, Formal Analysis, Writing – Original Draft Preparation, Writing – Review & Editing; Howe LD: Conceptualization, Data Curation, Writing – Review & Editing; Tobias JH: Conceptualization, Writing – Review & Editing; Paternoster L: Conceptualization, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the Wellcome Trust through a PhD Studentship to MF [105504] and the ALSPAC core programme grant [102215]. The UK Medical Research Council and Wellcome [102215] and the University of Bristol provide core support for ALSPAC. LDH is supported by a Career Development Award fellowship from the UK Medical Research Council [MR/M020894/1]. MF, LDH and LP work in a unit that receives support from the UK Medical Research Council and the University of Bristol [MC_UU_12013/4 & MC_UU_12013/5]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2018 Frysz M et al. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Frysz M, Howe LD, Tobias JH and Paternoster L. Using SITAR (Superimposition by Translation and Rotation) to estimate age at peak height velocity in Avon Longitudinal Study of Parents and Children [version 1; referees: 1 approved] Wellcome Open Research 2018, 3:90 (doi: 10.12688/wellcomeopenres.14708.1)

Introduction

Puberty is a period of significant biological and psychological changes in human development. One of its hallmarks includes a rapid growth spurt, the timing and speed of which varies between individuals, with marked sexual dimorphism.

The relationship of pubertal timing with adverse health outcomes in later life has been investigated previously. For example, previous studies reported an association between late menarche and increased risk of osteoporosis, whereas early pubertal timing has been found to be related to higher risk of obesity and cardiovascular disease in both men and women. Thus, investigating the influences of pubertal timing and understanding its relationship with later health outcomes is of great public health importance. While clinical assessment remains the gold standard for the assessment of pubertal status, this is difficult to achieve in large-scale studies. Self-reported puberty measures lack reliability and may be unpopular with study participants, potentially leading to large amounts of missing data. Another measure, age at peak height velocity (aPHV) provides an objective and non-invasive assessment of pubertal status.

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a longitudinal birth cohort, which was established in the 1990s in the South West of England. ALSPAC is a rich source of data, including phenotypic and genetic data collected for the mothers, fathers and children.

This paper describes the application of Superimposition by Translation And Rotation (SITAR) mixed effects growth curve model (described previously by Cole et al.7) for analysis of height in puberty and estimating aPHV in ALSPAC offspring which resulted in a new dataset being generated. Previous studies have shown this to be a suitable method for estimating aPHV9, even with fairly sparse data.

Methods

ALSPAC recruited a total of 14,541 pregnant women with expected delivery date between 1st April 1991 and 31st December 1992. Of these pregnancies, 68 have no known birth outcome, 195 were twin, 3 were triplet and 1 was quadruplet. In total, these pregnancies resulted in 14,062 live births of which 13,988 children were alive at 1 year of age.

In order to increase sample size additional recruitment took place when the children were on average 7 years old (Phase II) and again between age 8 – 18 years (Phase II) resulting in a total of 15,247 enrolled pregnancies, of which 14,701 were alive at 1 year of age. For more details regarding eligibility and recruitment please refer to Boyd et al.8.

These children have been followed up from birth (those recruited during Phases II and III were followed up from the time they joined the study) and data collection included questionnaire and clinical assessments. For the purpose of estimating age at PHV, only height measurements obtained by trained fieldworkers during assessment clinics were used. These clinics encompassed ‘children in focus clinics’ (CIF) to which a random 10% subsample of children were invited between the age 2 and 5 years, and six assessment clinics in late childhood (between age 7 and 13 years), three further assessment clinics in adolescence (ages 13, 15 and 17 years). These data were restricted to include height measurements collected when the children were between 5 to 20 years of age which resulted in a total of 61,290 height measurements available for 10,236 participants of whom 5,099 were female and 5,137 were male. To maximise accuracy of estimate of aPHV, these data were further restricted to include individuals with at least one height measurement for the following time periods: 5 to <10 years, 10 to < 15 years and 15 to 20 years. A total of 46,246 height measurements for 5,707 individuals (3,019 females, 2,688 males) were available for SITAR analysis.

SITAR is a mixed effects shape-invariant growth curve model, consisting of a mean growth curve along with three transformations (size, tempo and velocity), used to describe how each individual differs from the mean curve. The three SITAR parameters are size, reflecting up/down shift from the mean curve; tempo, reflecting left/right shift (on the age scale) which corresponds to the relative timing of puberty based on aPHV, and velocity reflecting stretching/shrinking of the age scale and hence describing differences in the rate at which individuals pass through puberty. For a detailed description of the method please refer to Cole et al.7.

Height data were uploaded into R and following outlier removal of people with velocity exceeding 4 SDs (using velout function; for more details see sitar package documentation in R10), an initial SITAR model was fitted, for males and females separately, in R version 3.4.1. Following initial model fitting, standardized residuals exceeding 4 in absolute value were removed leaving a total of 40,037 height measurements for 5,707 individuals available for analysis with an average of 8 measurements (range 1–10 measurements) available per participant (see Figure 1 for details regarding participant recruitment).

Dataset validation

The final SITAR model was fitted with 5 degrees of freedom, for males and females separately, and it explained 98% and 98.3% of variance in the dataset for males and females, respectively. Mean aPHV (SD) was 13.6 (0.9) for males (Figure 2) and 11.7 (0.8) for females (Figure 3) and these estimates are similar to those reported in the literature.

Table 1 shows mean and SD of additional variables that were estimated, which include: size (cm) (negative values indicating smaller whereas positive taller children and zero represents the mean), tempo (years) (negative values indicating early puberty, positive late puberty and zero represents the mean), velocity (years) (measure of intensity, positive values indicating short growth spurt, zero corresponding to mean velocity) along with peak velocity (cm).

Ethical approval and consent

Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committee, and Ethical approval and consent was obtained from the children and parents.
Enrolled foetuses
N = 14,676

Alive at 1 year
N = 13,988

Additional recruitment at Phase II/III
N = 713

Total eligible
N = 14,701

Total height measurements collected between age 5 – 20 years
N = 61,290
for 5,137 males (30,076 measurements) and
5,099 females (31,214 measurements)

Total height measurements for individuals with at least one
measure between age 5-10, 10-15 and 15-20 years
N = 46,246
for 2,688 males (21,840 measurements) and
3,019 females (24,406 measurements)

Final height measurements available for SITAR modelling (outliers removed)
N = 40,037
for 2,688 males (20,849 measurements) and
3,019 females (24,216 measurements)

**Figure 1.** Avon Longitudinal Study of Parents and Children (ALSPAC) participant recruitment and height measurements available for analysis.

![Figure 2](image)

**Figure 2.** Mean growth curve (solid line) and velocity (dashed line) plots estimated by Superimposition by Translation and Rotation (SITAR) for males. Vertical dotted line represents mean age at peak height velocity.
Table 1. Means and standard deviations of estimated SITAR growth parameters in Avon Longitudinal Study of Parents and Children (ALSPAC) offspring.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Combined (N= 5,707)</th>
<th>Males (N=2,688)</th>
<th>Females (N=3019)</th>
</tr>
</thead>
<tbody>
<tr>
<td>aPHV (years)</td>
<td>12.6 (1.3)</td>
<td>13.6 (0.9)</td>
<td>11.7 (0.8)</td>
</tr>
<tr>
<td>Size (cm)</td>
<td>0.0 (6.3)</td>
<td>0.0 (6.5)</td>
<td>0.0 (6.0)</td>
</tr>
<tr>
<td>Tempo (years)</td>
<td>0.0 (0.9)</td>
<td>0.0 (0.9)</td>
<td>0.0 (0.9)</td>
</tr>
<tr>
<td>Velocity (years)</td>
<td>0.0 (0.1)</td>
<td>0.0 (0.1)</td>
<td>0.0 (0.1)</td>
</tr>
<tr>
<td>Peak velocity (cm)</td>
<td>8.8 (1.5)</td>
<td>10.0 (1.1)</td>
<td>7.7 (0.8)</td>
</tr>
</tbody>
</table>

Abbreviations: aPHV (age at peak height velocity), SD (standard deviation)

Figure 3. Mean growth curve (solid line) and velocity (dashed line) plots estimated by SuperImposition by Translation and Rotation (SITAR) for females. Vertical dotted line represents mean age at peak height velocity.

Written informed consent was obtained from parents, and children were invited to give consent where appropriate. Study members have the right to withdraw their consent for elements of the study or from the study entirely at any time.

Data availability
ALSPAC data access is through a system of managed open access. The steps below highlight how to apply for access to the data included in this data note and all other ALSPAC data. The dataset generated in this data note has been deposited within the ALSPAC data resource and is linked to ALSPAC project number B2325. Please quote this number to request required variables which have been described in this dataset (size, tempo, velocity, aPHV and peak velocity).

1. Please read the ALSPAC access policy (PDF, 627kB) which describes the process of accessing the data and samples in detail, and outlines the costs associated with doing so.
2. You may also find it useful to browse our fully searchable research proposals database, which lists all research projects that have been approved since April 2011.
3. Please submit your research proposal for consideration by the ALSPAC Executive Committee using the online process. You will receive a response within 10 working days to advise you whether your proposal has been approved.

If you have any questions about accessing data, please email alspac-data@bristol.ac.uk.

The ALSPAC data management plan describes in detail the policy regarding data sharing, which is through a system of managed open access.

Competing interests
No competing interests were disclosed.

Grant information
This work was supported by the Wellcome Trust through a PhD Studentship to MF [105504] and the ALSPAC core programme grant [102215]. The UK Medical Research Council and Wellcome [102215] and the University of Bristol provide core support for ALSPAC. LDH is supported by a Career Development Award fellowship from the UK Medical Research Council [MR/M020894/1]. MF, LDH and LP work in a unit that receives support from the UK Medical Research Council and the University of Bristol [MC_UU_12013/4 & MC_UU_12013/5].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements
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.

References

Frysz and colleagues from the University of Bristol describe the use of the statistical programme, Superimposition by Translation and Rotation (SITAR), to determine the age of peak height velocity (aPHV) in a cohort of children participating in the Avon Longitudinal Study of Parents and Children (ALSPAC). The results they obtained for both boys and girls (13.6 years and 11.7 years respectively) are similar to those obtained by other researchers in well-resourced countries.

SITAR, a recently developed modelling approach by Tim Cole, has been used increasingly in longitudinal studies to describe secular and population trends in childhood growth and pubertal development using skeletal development (Tanner RUS score) and aPHV in both poorly- and well-resourced communities. Its advantages are that it assesses each individual's curve against the mean curve and is able to determine the size (amplitude of change), tempo (the timing of the onset of change) and velocity (rate of change) of each individual curve in relation to the mean curve. It is also robust with missing data. Cole in a paper published this year assessed the interval required between height measurements in individual adolescent children to optimize the determination of aPHV. Cole concluded that height measurements need not be taken more frequently than biennially to obtain an accurate assessment of the average aPHV.

The reviewers have a few minor comments, which might help to provide further clarity on the participants included in the cohort. The authors describe the recruitment of participants from the ALSPAC cohort and the inclusion criteria for height measurements to be included into the database used for this particular aspect of the study. It is however unclear if black or Asian children were included in the analytical cohort sample. The reason for this comment is that data are available to suggest the age of pubertal development and aPHV in black children may differ from that of white children living in both well- and poorly-resourced countries. Thus their inclusion might have influenced the mean values of aPHV obtained and the subsequent comparisons with other cohorts.

The inclusion of an individual into the SITAR study required that at least one height measurement was available at each of three different time periods, 5-<10 years, 10-<15 years and 15-20 years. Yet in their analysis of the average number of measurements available per individual, the average is given as 8 with a range from 1-10. It is unclear how children with only one measurement were included based on
the criteria for inclusion given above. It is possible that the authors meant that the participants included had at least one measurement in any of the age bands.

In the second paragraph of the Methods section, line 3, Phase II is referred to as being between age 8-18 years. We believe this should be Phase III.

References

Is the rationale for creating the dataset(s) clearly described?
Yes

Are the protocols appropriate and is the work technically sound?
Yes

Are sufficient details of methods and materials provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Yes

**Competing Interests:** No competing interests were disclosed.

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.