Pubertal development in HIV-infected African children on first-line antiretroviral therapy

Alexander J. Szubert[a],*, Victor Musiime[b,c],*,
Mutsawashe Bwakura-Dangarembizi[d],*, Patricia Nahirya-Ntege[e],
Adeodata Kekitiinwa[f], Diana M. Gibb[a], Kusum Nathoo[d],*,
Andrew J. Prendergast[g],*, A. Sarah Walker[a],*,
the ARROW Trial Team

Objectives: To estimate age at attaining Tanner stages in Ugandan/Zimbabwean HIV-infected children initiating antiretroviral therapy (ART) in older childhood and investigate predictors of delayed puberty, particularly age at ART initiation.

Design: Observational analysis within a randomized trial.

Methods: Tanner staging was assessed every 24 weeks from 10 years of age, menarche every 12 weeks and height every 4–6 weeks. Age at attaining different Tanner stages was estimated using normal interval regression, considering predictors using multivariable regression. Growth was estimated using multilevel models with child-specific intercepts and trajectories.

Results: Median age at ART initiation was 9.4 years (inter-quartile range 7.8, 11.3) ($n = 582$). At the first assessment, the majority (80.2%) were in Tanner stage 1; median follow-up with staging was 2.8 years. There was a strong delaying effect of older age at ART initiation on age at attaining all Tanner stages ($P < 0.05$) and menarche ($P = 0.02$); in boys the delaying effect generally weakened with older age. There were additional significant delays associated with greater impairments in pre-ART height-for-age Z-score ($P < 0.05$) in both sexes and pre-ART BMI-for-age in girls ($P < 0.05$). There was no evidence that pre-ART immuno-suppression independently delayed puberty or menarche. However, older children/adolescents had significant growth spurts in intermediate Tanner stages, and were still significantly increasing their height when in Tanner stage 5 ($P < 0.01$).

Conclusion: Delaying ART initiation until older childhood substantially delays pubertal development and menarche, independently of immuno-suppression. This highlights that factors other than CD4$^+$, such as pubertal development, need consideration when making decisions about timing of ART initiation in older children.

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

---

[a]MRC Clinical Trials Unit at University College London, London, UK, [b]Joint Clinical Research Centre, [c]Makerere University College of Health Sciences, Kampala, Uganda, [d]University of Zimbabwe, Harare, Zimbabwe, [e]Medical Research Council/Uganda Virus Research Institute Uganda Research Unit on AIDS, Entebbe, [f]Paediatric Infectious Diseases Clinic/Baylor – Uganda, Mulago Hospital, Kampala, Uganda, and [g]Queen Mary University of London, London, UK.

Correspondence to Alexander J. Szubert, MRC Clinical Trials Unit at UCL, Institute of Clinical Trials and Methodology, 125 Kingsway, London WC2B 6NH, UK.

Tel: +44 20 7670 4746; fax: +44 20 7670 4814; e-mail: a.szubert@ucl.ac.uk

*All the authors contributed equally to the writing of this article.

Received: 16 September 2014; revised: 9 January 2015; accepted: 12 January 2015.

DOI:10.1097/QAD.0000000000000590

Introduction

Globally, more than 3 million children/adolescents are living with HIV infection, and more than 90% live in sub-Saharan Africa [1]. Of the 1.6 million HIV-infected children/adolescents needing antiretroviral therapy (ART) in 2012, only 34% were receiving it [2]. WHO guidelines have increasingly promoted earlier ART, recommending universal treatment for all children aged below 2 years in 2010 [3] and below 5 years in 2013 [4]. However, historically substantial numbers of children in sub-Saharan Africa have initiated ART at older ages. Whilst studying long-term outcomes in those initiating ART in infancy will be important over the coming decades, survivors initiating ART in late childhood are entering adolescence now, having experienced long durations of untreated infection. They face considerable challenges [5] that make long-term follow-up equally important in this cohort.

One concern is puberty – both pubertal delay (and associated impact including stigma) and the possibility of major body shape changes. HIV has been associated with delays in puberty, but relatively few studies have considered the impact of age at ART initiation. Furthermore, studies have predominantly been in high-income settings [6–8], with one study in Uganda [9].

Here, we use data from the AntiRetroviral Research fOr Watoto (ARROW) trial, in which pubertal development was assessed every 24 weeks in older children/adolescents initiating ART in Uganda/Zimbabwe, to estimate the impact of age at ART initiation in later childhood on pubertal milestones, menarche, height and body shape changes.

Methods

In the ARROW trial (ISRCTN24791884), previously untreated Ugandan/Zimbabwean children/adolescents aged 3 months to 17 years, eligible for ART using WHO 2006 criteria [10], were randomized 1:1 to clinically driven monitoring vs. laboratory and clinical monitoring for toxicity and efficacy ($CD4^+$) [11]. No viral loads were assayed in real time. Children were also randomized 1:1:1 in a factorial design to open-label lamivudine + abacavir + non-nucleoside reverse transcriptase inhibitor (NNRTI) continuously (arm A, no zidovudine) versus induction-maintenance with 4-drug regimen lamivudine + abacavir + NNRTI + zidovudine for 36 weeks, followed by lamivudine + abacavir + NNRTI (arm B; short-term zidovudine) or lamivudine + abacavir + zidovudine (arm C; long-term zidovudine). The NNRTI (nevirapine/efavirenz) was chosen by clinicians according to local availability (varying by

country) and age. Caregivers gave written consent; older children (8–17 years) aware of their HIV status also gave assent/consent following national guidelines. Children were recruited from three centres in Uganda and one in Zimbabwe. The ARROW trial was approved by Research Ethics Committees in Uganda, Zimbabwe and the United Kingdom.

Weight and height were measured by nurses at enrolment (ART initiation/baseline) and every 4–6 weeks subsequently. At baseline and every 24 weeks, physicians assessed Tanner staging [12] of pubic hair (P1–P5) and genitalia (G1 = pre-pubertal to G5 = adult, males) or breasts (B1–B5, females) in those aged at least 10 years. Attainment of menarche (year and month) was assessed in females every 12 weeks. In a separate metabolic sub-study at the Joint Clinical Research Center and Harare sites, consent was obtained to measure body circumferences (waist/hip/mid-thigh/mid-upper-arm) and skinfold thicknesses (triceps/sub-scapular/supra-iliac) every 24 weeks in those aged at least 10 years of age using standardized procedures with pre-specified positions for three repeated measurements [13]. Analysis used the mean of these measurements. Height and BMI Z-scores were calculated using WHO references [14,15] and anthropometry Z-scores using Dutch references [16,17] (no African data available).

Statistical analysis

Age at attaining Tanner stages 2, 3, 4 and 5 was analysed for males (separately for genitalia and pubic hair) and females (separately for breasts and pubic hair). For individual children, the age when each transition between stages occurred was known only to be between the last observation of the previous stage and the first observation of the relevant stage (i.e. between assessments). To account for this interval-censoring, normal interval regression was used to estimate mean transition age, assumed normally distributed (Supplementary Methods, http://links.lww.com/QAD/A640) [18]. Age at menarche was analysed similarly. Sensitivity analyses used a time-to-event Weibull model allowing asymmetry in the age distribution (results similar unless shown).

As G2 in boys and B2 in girls generally precedes P2 [19], subsequent analyses focused on time to G2/B2 (separately in boys/girls), G3/B3, G4/B4, G5/B5 and menarche. Factors considered as predictors were centre; ART-strategy randomization; monitoring randomization; age, $CD4^+$, WHO stage, height/BMI-for-age at ART initiation and change in $CD4^+$ and height/BMI-for-age during the first 6 months on ART (to reflect initial response). Weight-for-age was not considered because WHO standards only exist for those aged below 10 years, and it was highly associated with height-for-age (Spearman correlation 0.73). As the ARROW trial recruited over 18 months, birth cohort was co-linear with age at ART initiation (Spearman correlation 0.98) and was

therefore not considered. For G2/B2, backward elimination (exit $P = 0.05$) was used to identify important factors: all factors significant in either model were included in final sex-specific models for G2/B2. Models for G3/B3, G4/B4, G5/B5 and menarche included the same factors, allowing trends in effects over Tanner stages to be assessed. Non-linearity in effects of continuous variables was explored using natural cubic splines [20].

For the 539 (93.4%) children/adolescents in Tanner stage 1/2 at first assessment, all height measurements aged above 5 years (after which growth is approximately linear [14]) were modelled separately for males and females using multilevel models with a random intercept and individual random slopes for each genitalia (male) or breast (female) Tanner stage (unstructured correlation matrix). Age was the underlying time-scale (Supplementary Methods, http://links.lww.com/QAD/A640).

Anthropometric Z-scores were modelled using multilevel models with a random intercept and random slopes before and after G2/B2 (unstructured correlation matrix), using age as the underlying time-scale. Children already in G2/B2 or higher at first assessment were included using a fixed-effect indicator variable (missing transition age). Randomized ART strategy was included as a fixed effect on intercept (Z-score at G2/B2 transition) and prior/subsequent slopes to estimate associations with zidovudine exposure given concerns about potential lipoatrophy with thymidine analogues. A few (2.4%) outliers at least 1.25 Z-score different from both the previous and following measurements were excluded (as was the first/

last measurement if ≥ 1.25 Z-score different from the following/previous measurement).

Analyses were performed using Stata 13.1 (StataCorp, College Station, Texas, USA). P values were two-sided. No adjustment was made *a priori* for multiple testing.

Results

Six hundred and twenty children reached their 10th birthday during follow-up or were aged at least 10 years at enrolment. Five hundred and eighty-two (93.9%) had at least one assessment of Tanner staging or menarche; in these children, Tanner staging and menarche were assessed at 87.9 and 87.4% of scheduled visits after 10 years, respectively. Median age at ART initiation was 9.4 years [inter-quartile range (IQR) 7.8, 11.3; range 5.2–17.6) (Supplementary Table 1, http://links.lww.com/QAD/A640). At first Tanner assessment, most (80.2%) were in stage 1 (median age 10.4 years), 13.2% were in stage 2 (10.9 years) and 6.6% in stage 3/4/5 (13.9 years). Three hundred and forty-nine (60.5%) children had already initiated ART at their first Tanner assessment. Median follow-up between first and last staging was 2.8 years (IQR 1.4, 3.7).

Tanner staging and menarche
In males, mean age at reaching G2 was 12.8 (3rd–97th percentiles 10.0–15.5), P2 13.0 and adulthood (G5) 16.9 years. In females, mean age at B2 was 11.7 (3rd–97th percentiles 9.4–14.0), P2 12.3, menarche 14.3 and

Table 1. Age on reaching each Tanner stage and menarche.

Sex	Stage[a]	Number of children contributing measurement	Mean age on reaching stage (years)	SE	95% CI	Percentiles 3%	Percentiles 25%	Percentiles 75%	Percentiles 97%	Mean/median age Black American US children [24,25]	Percentiles in uninfected Italian controls [6] 3%	Percentiles in uninfected Italian controls [6] 97%
Male	G2	266	12.8	0.1	12.6, 13.0	10.0	11.8	13.8	15.5	9.3	9.4	13.3
	G3	171	14.3	0.1	14.0, 14.6	10.9	13.1	15.5	17.6	11.8	11.0	14.0
	G4	107	15.3	0.2	14.9, 15.7	11.7	14.0	16.7	19.0	13.4	12.3	14.8
	G5	71	16.9	0.2	16.5, 17.3	14.3	16.0	17.8	19.5	14.9	–	–
	P2	270	13.0	0.1	12.8, 13.2	10.5	12.1	13.9	15.5	–	9.4	13.5
	P3	161	14.3	0.1	14.1, 14.6	11.1	13.2	15.5	17.6	–	11.0	13.1
	P4	109	15.3	0.2	14.9, 15.7	11.8	14.1	16.6	18.8	–	12.3	14.8
	P5	69	16.8	0.2	16.4, 17.2	14.6	16.0	17.7	19.1	–	–	–
Female	B2	273	11.7	0.1	11.6, 11.9	9.4	10.9	12.5	14.0	9.5	8.5	12.6
	B3	236	13.1	0.1	12.9, 13.3	10.4	12.1	14.0	15.7	–	9.8	13.2
	B4	179	14.4	0.1	14.1, 14.7	11.4	13.3	15.5	17.4	–	11.2	14.4
	B5	107	16.1	0.2	15.7, 16.5	13.2	15.1	17.2	19.0	–	–	–
	P2	284	12.3	0.1	12.1, 12.4	9.9	11.4	13.1	14.6	–	8.3	12.7
	P3	219	13.5	0.1	13.3, 13.7	10.8	12.5	14.5	16.2	–	9.1	13.3
	P4	161	14.5	0.1	14.3, 14.8	11.7	13.5	15.5	17.3	–	11.2	14.4
	P5	107	16.5	0.2	16.1, 16.8	13.8	15.5	17.4	19.2	–	–	–
	Menarche	303	14.3	0.1	14.1, 14.5	11.6	13.3	15.3	17.0	12.1	–	–

Note: Time-to-event Weibull approach (see Methods section) resulted in similar estimates (e.g. G2 mean 12.6 years; B2 11.3; menarche 14.4). No uninfected African control data available. The symbol '–' indicates 'not available'.
[a]Tanner staging of genitalia (G1 = pre-pubertal to G5 = adult) in males or breasts (B1 to B5) in females, and pubic hair (P1 to P5) in both.
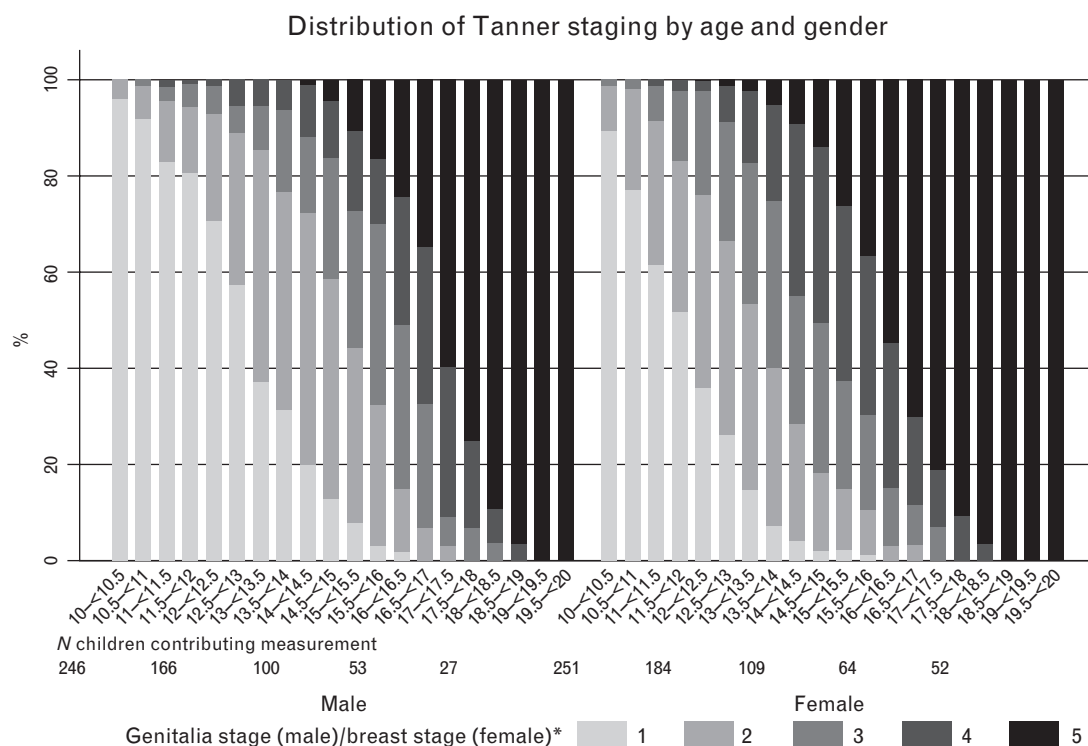
Fig. 1. Distribution of Tanner staging by age and sex. Including closest to mid-point measurement per child in each age range. (*) 1 = pre-pubertal to 5 = adult.

adulthood (P5) 16.5 years (Table 1). At 12 years, 36% of girls had not attained B2 (Fig. 1), 56% had not attained P2 and 58% had not attained both B2 and P2; at 13 years, 37% of boys had not attained G2, 50% had not attained P2 and 53% had not attained both G2 and P2.

Factors independently affecting age at attaining G2 or B2 were age at ART initiation; height and BMI Z-scores at ART initiation; change in $CD4^+$, height and BMI Z-scores during the first 6 months on ART and centre. To allow comparison across stages, all these factors were included in sex-specific models for G2/B2, G3/B3, G4/B4, G5/B5 and menarche. There was a strong delaying effect of initiating ART at older rather than younger ages on age at attaining each Tanner stage for boys and girls ($P < 0.05$) and for menarche ($P = 0.02$). Figure 2 shows the relationship between age at ART initiation and reaching different Tanner stages: being 1 year older at ART initiation had different impacts on pubertal delay depending on the specific age at ART initiation (i.e. non-linear effects). For example, a boy initiating ART aged 9 years would reach G2/G3/G4/G5 aged 13.3/14.0/14.4/16.8 compared to 13.2/14.8/15.6/16.3 if initiating aged 11. Similarly a girl initiating ART aged 9 would reach B2/B3/B4/B5 aged 11.8/12.7/13.7/14.6 compared to 12.3/13.6/14.5/15.7 if aged 11. In boys, the delaying effect of being older at ART initiation generally weakened with increasing age: a boy initiating ART aged 7 would reach G2 0.5 years later than one who is aged 6; a boy aged 12

would reach G2 0.2 years later than one aged 11 (Supplementary Table 2, http://links.lww.com/QAD/A640). In contrast, in girls, the delaying effect on age at B2 in particular was stronger in those initiating at older ages: a girl initiating ART aged 7 would reach B2 0.1 years later than one aged 6; a girl aged 12 would reach B2 0.5 years later than one aged 11. Sensitivity analyses with Weibull-distributed transition ages gave broadly similar results (Supplementary Figure 1, http://links.lww.com/QAD/A640).

Independently of the strong association between age at ART initiation and age at G2–G5 and B2–B5, pubertal stages were consistently delayed even further in children with lower pre-ART height-for-age (Fig. 3; Supplementary Table 3, http://links.lww.com/QAD/A640). For example, having one Z-score lower height-for-age at ART initiation delayed G2 in boys by 0.4 years and B2 in girls by 0.3 years. There was weak evidence of further delays with relatively small increases in height-for-age during the first 6 months on ART. In girls, but not boys, there were also independent consistent delays with lower pre-ART BMI-for-age and relatively small initial increases in BMI-for-age on ART. In boys, earlier G2 was independently associated with greater initial increases in $CD4^+$ on ART, but for G3–G5 and in girls there were no statistically significant effects. Transition ages also varied on the basis of centre, with G2–G5 and B2–B4 reached earlier in boys and girls in the only non-urban
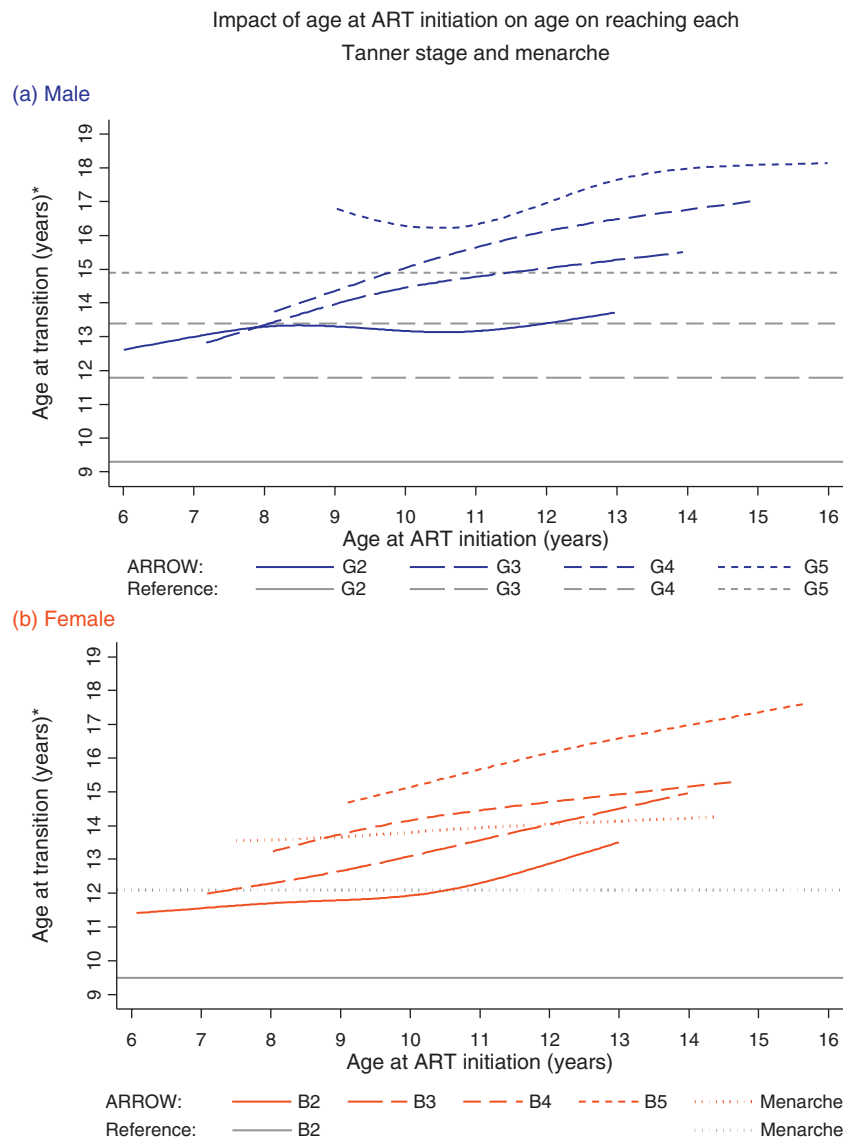
Impact of age at ART initiation on age on reaching each
Tanner stage and menarche

Fig. 2. Impact of age at antiretroviral therapy initiation on age on reaching each Tanner stage and menarche (interval regression). (*) Absolute age at reaching different Tanner stages is presented at the reference category (mode or approximate median) for other significant predictors (see Supplementary Table 3, http://links.lww.com/QAD/A640), specifically for a Zimbabwean child with height and BMI-for-age −2 and −1 at ART initiation, and −2 and −0.7 six months after ART initiation; and 250 cells/μl increase in $CD4^+$ during the first 6 months on ART. The effect of age at ART initiation (slope of line) does not depend on these factors, but the absolute age does. Tanner staging of genitalia (G1 = pre-pubertal to G5 = adult) in males or breasts (B1 to B5) in females. Grey lines indicate overall mean/median age on reaching each Tanner stage and menarche in Black American US children when available (no data for G3–G5 in girls). ART, antiretroviral therapy.

centre (Entebbe, Uganda) (Supplementary Table 3, http://links.lww.com/QAD/A640). Centre effects were strongest in males and at G2/B2. Other differences between centres were smaller and inconsistent across stages. There were no statistically significant effects of centre on age at menarche, and the impacts of the other predictors described above were generally stronger for menarche than for Tanner staging, suggesting less measurement error in menarche. Fitting the final sex/stage–specific models separately in each centre resulted in broadly consistent effects.

Pre-ART $CD4^+$ was not an independent predictor of G2 ($P = 0.97$ adjusting for factors above including $CD4^+$ change during first 6 months on ART) or B2 ($P = 0.15$), and had no significant independent effect on G3, B3, B5 and menarche ($P > 0.1$; inverse associations with G4/B4 suggested over–fitting; Supplementary Table 4, http://links.lww.com/QAD/A640).

Previous studies have found household structure can influence pubertal development. Considering additional effects of the number of boys, girls, men (≥15 years) and
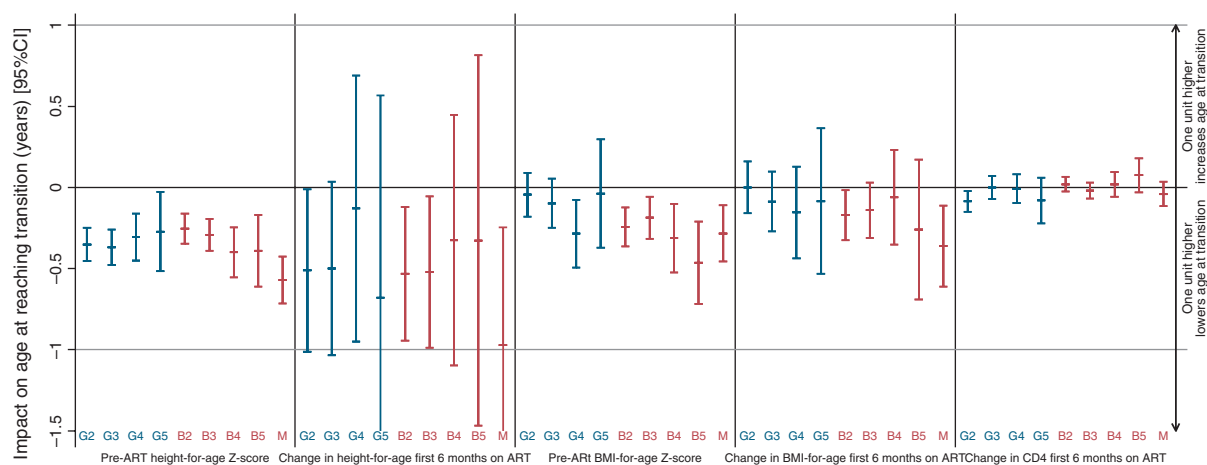
Fig. 3. Impact of height/BMI-for-age Z-scores at ART initiation and change in height/BMI-for-age during the first 6 months on ART on age on reaching each Tanner stage and menarche (interval regression). Tanner staging of genitalia (G1 = pre-pubertal to G5 = adult) in males or breasts (B1 to B5) in females; M, menarche. Z-score effects are per unit higher; CD4$^+$ effects per 100 cells/μl higher. Confidence intervals (CIs) for impact of change in height-for-age on age at reaching G5 and menarche truncated at −1.5. See Figure 2 for the effect of age at ART initiation. For point estimates and values within a 95% CI below the line at $y = 0$, one unit higher value lowers the age at which the Tanner stage is attained. For point estimates and values within a 95% CI above the line at $y = 0$, one unit higher value raises the age at which the Tanner stage is attained. ART, antiretroviral therapy.

women in the household at ART initiation, and adjusting for factors above, we found that in girls, B5 was significantly earlier with more boys and/or men in the household at ART initiation (both $P = 0.03$), and B4 was marginally faster with more men in the household ($P = 0.07$) (Supplementary Table 5, http://links.lww.com/QAD/A640). There was no evidence that more boys or men affected age at B2, B3 or menarche ($P > 0.1$). In boys, G2 was significantly later with more boys in the household ($P = 0.05$), with similar but weaker effects on G3/G4 ($0.05 < P \le 0.1$).

Growth

In boys, overall mean height growth varied across Tanner stages: from 5 years of age to G2, it was 4.9 cm/year [95% confidence interval (CI) 4.7, 5.1], accelerating slightly during G2 (5.5 cm/year, 95% CI 5.0, 6.0) and G3 (6.1 cm/year, 95% CI 5.3, 6.9), before decreasing during G4 (4.9 cm/year, 95% CI 4.0, 5.8) and G5 (2.5 cm/year, 95% CI 0.9, 4.0) (all $P < 0.01$; heterogeneity $P < 0.001$). In girls, these effects were earlier and more pronounced (G1−G5: 5.4 cm/year, 95% CI 5.2, 5.7; 7.1 cm/year, 95% CI 6.7, 7.5; 5.3 cm/year, 95% CI 4.8, 5.9; 2.8 cm/year, 95% CI 2.3, 3.3 and 2.4 cm/year, 95% CI 1.7, 3.2, respectively; all $P < 0.001$; heterogeneity $P < 0.001$). Children initiating ART at an older age grew slightly more slowly than those initiating it at a younger age and, as above, also experienced pubertal delay, which itself impacts growth. Figure 4 displays the combined effect on growth of initiating ART at three representative ages (8, 10 or 12 years), incorporating its effects of Tanner staging. For example, a boy initiating ART at 8 years would expect to be 143 cm tall at G2 (13.3 years), compared to 139 cm (13.4 years) if ART was initiated at 12 years of age.

Interestingly, whilst girls normalized height on ART (e.g. crossing WHO reference percentiles upwards), boys remained at a considerable height disadvantage. However, we were unable to reliably estimate whether this would be detrimental to the final height because follow-up during G5/B5 was relatively short (median 0.5 years, IQR 0.2, 0.9).

Body circumferences and skin-fold thicknesses

In the metabolic sub-study, 326 children [156 (48%) male] had anthropometric measurements available at 1110 visits aged at least 10 years (including 102 children only observed in G2/B2, and 44 only in G2/B2 or higher). Before G2/B2, anthropometric Z-scores were decreasing very slightly (Supplementary Table 6, http://links.lww.com/QAD/A640), significantly so for waist/hip/mid-thigh/mid-upper arm body circumferences and supra-iliac skin-fold thickness. After G2/B2, all body circumference Z-scores increased over time (by 0.1−0.2 Z-scores higher/year; $P < 0.0001$ for all); skin-fold thicknesses also increased, but changes were only significant for sub-scapular skin-folds ($P = 0.03$). There were no significant differences in changes in body circumference or skin-fold thickness Z-scores before or after G2/B2 between children receiving lamivudine/abacavir/NNRTI throughout and either zidovudine for a longer term, or just for 36 weeks (all $P > 0.1$).

Discussion

Here we present longitudinal data from a large cohort of older children in sub-Saharan Africa, followed for
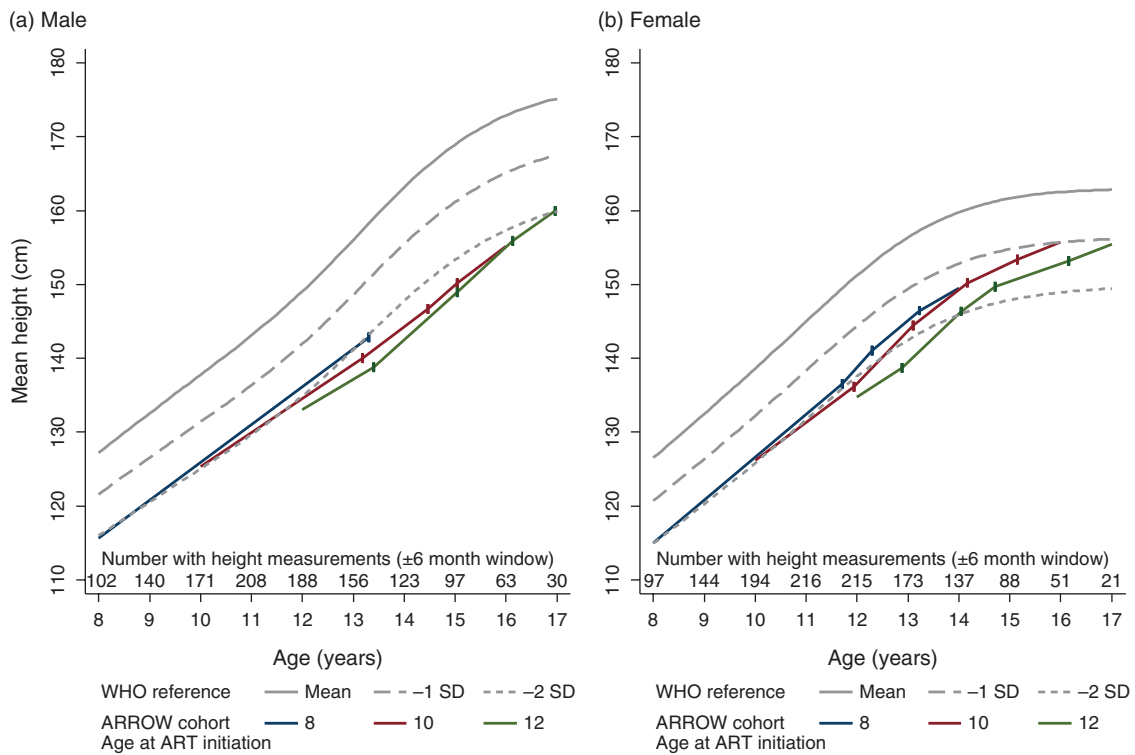
Fig. 4. Dependence of growth on Tanner staging and age at antiretroviral therapy initiation in comparison to WHO growth standards. Each solid coloured line indicates the estimated height at ART initiation, and subsequent growth for a child aged 8, 10 or 12 years at ART initiation, in comparison with WHO standards (grey lines). Vertical lines indicate the expected age at each Tanner transitions for these ages at ART initiation (estimated at reference categories for other significant predictors as per Figure 2, i.e. for a Zimbabwean child with height and BMI-for-age −2 and −1 at ART initiation, and −2 and −0.7 six months after ART initiation; and 250 cells/μl increase in $CD4^+$ during the first 6 months on ART). Growth predictions assume linear growth in between these transitions (as per multilevel models; see Methods section). ART, antiretroviral therapy.

approximately 3 years on ART, having initiated treatment mostly aged approximately 7–12 years. Our main finding is a substantial delay in pubertal stages and menarche associated with initiating ART at older ages (Fig. 2), unrelated to pre-ART immune-suppression. This highlights the importance of considering factors other than current $CD4^+$ in decisions to initiate ART in older children, such as the potential for pubertal delay and long-term immune reconstitution [21]. In particular, it suggests that, regardless of immune parameters, vertically infected children aged over 8 years may need to start ART to reduce substantial pubertal delay. For example, recent Ugandan guidelines recommend universal ART in all vertically infected children aged below 15 years [22].

Our results contrast a recent US meta-analysis [8], which found an effect of birth cohort rather than age at ART initiation on pubertal delay. We were unable to assess independent effects of birth cohort and age as recruitment was over only 18 months; however, US children had a wide range of birth years, whereas most ARROW children included in this analysis were born between 1996 and 2000 [8]. Importantly, the US study included many children with mono/dual ART exposure, so the effect of

age at initiation of combination ART in previously untreated children could not be assessed; these children also received combination ART with protease inhibitors or NNRTIs. The fact that we definitively identified large and consistent associations between later age at ART initiation and pubertal delays in which birth cohort covered a relatively narrow range and WHO-recommended first-line regimens were used, suggests that previously identified birth cohort effects could be due to complex confounding (by indication) between age at initiation of combination ART, previous mono/dual regimens and year of birth, as noted [8].

Independent of this association between age at ART initiation and pubertal delay, we found additional significant delays associated with greater pre-ART impairments in height Z-score and smaller initial height responses on ART. These most likely reflect a child-specific impact of untreated HIV on growth: the more impaired a child's growth has been by HIV for their age, and the smaller any early ART-associated improvements, the longer the subsequent pubertal delay. Similar effects of pre-ART BMI and initial BMI response in girls, but not boys, are plausibly related to direct HIV-specific effects on

sexual maturation. HIV may impair growth and puberty in several common ways, including nutritional deficiencies, recurrent infections and chronic immune activation, which may all skew endocrine axes. In other childhood diseases characterized by chronic inflammation, such as Crohn's disease [23], the combination of under-nutrition and elevated pro-inflammatory cytokines causes reductions in insulin-like growth factor 1 and alterations in gonadotrophin-releasing hormone-secretion patterns, delaying growth and puberty. It is likely that gonadotrophin and growth hormone resistance similarly occur in HIV-infected children, with greater delays in ART initiation exposing children to longer periods of uncontrolled viral replication and chronic inflammation, exacerbating impairments in linear growth and puberty.

In contrast to the previous studies, we did not find consistent effects of more men in the household hastening female development, other than for B5 (and marginally for B4). This could reflect an overwhelming influence of HIV on development. Alternatively, this could be because African household structures may be more variable and this was assessed at ART initiation, not longitudinally; men/boys might be variably present due to absence for work/boarding school, and/or because HIV-infected children commonly move between carers/households. We did find a small but consistent effect of more boys in the household delaying male development (G2−G4), which could reflect restriction of access to assets (including food).

Of note, we found no evidence for independent effects of immune-suppression at ART initiation, although boys with greater initial $CD4^+$ increases on ART reached G2 slightly earlier than those with poorer immune responses. In contrast, a study of 983 HIV-infected US older children/adolescents found consistent effects of immune-suppression [7], which was confirmed in the US meta-analysis [8]. However, these studies included children initiating combination ART with and without previous mono/dual ART exposure, so were unable to investigate the impact of height/BMI-for-age at ART initiation in those previously untreated. Pre-ART growth could be a more sensitive marker of underlying developmental problems than immuno-suppression, as it integrates various insults, including infection (related and unrelated to pre-ART $CD4^+$, e.g. tuberculosis), growth hormone resistance, micronutrient deficiency, inflammation, poverty and food insecurity. In high-income countries, immune-suppression may capture most of these effects.

Overall, our estimates of mean age at different Tanner stages are broadly similar to HIV-infected girls in an Italian study in 2001 [6], but boys showed a delay of approximately 0.5 years compared to these children, and both sexes were delayed by approximately 1 year compared with the large US meta-analysis [8].

Correspondingly, much higher proportions had delayed pubertal onset (56 vs. 4% in [8]). Both sexes were delayed by approximately 1−2 years compared to uninfected Italian controls [6], and boys by approximately 2−3 years and girls by approximately 2 years compared to uninfected Black American children [24,25]. Interestingly, there appear to be few published data on the variability of individual children around these means: compared to the uninfected Italians, 3rd−97th percentiles of ages at various Tanner stages were approximately 1 year lower and 2 years higher. Before analysis, we hypothesized that this lack of symmetry in the age distribution at different pubertal stages [19] could be important. We therefore also considered Weibull models, allowing more individuals to be above or below the mean. Results were broadly similar, suggesting the asymmetry is not sufficiently severe to substantially affect models.

Clinically, adolescents would be expected to stop growing in Tanner stage 5, and WHO (and other) references indicate little growth after 16 years in girls and 17 years in boys [14]. ARROW children, however, were still growing significantly in G4/B4 and G5/B5, consistent with delay in calcification of the growth plates. The normal sequence of puberty in HIV-infected children is not well characterized, and it is plausible that discordance in pubertal stages occurs in chronic HIV infection. The fact that growth continued in late adolescence may reflect the 'advantage' delayed puberty affords for linear growth, meaning final height may be rescued to some extent because of the greater potential for catch-up growth than that anticipated on the basis of age and/or pubertal stage alone. There is emerging interest in the potential window of opportunity that adolescence provides [26] for catch-up linear growth in low-income countries where early childhood stunting is common, even outside HIV infection. However, our results show the importance of early ART, because the greater the growth impairment at treatment initiation, the greater the requirement for catch-up growth, which is eventually limited by growth plate senescence.

There have been case reports of lipodystrophy developing very quickly at puberty [27,28]. With long-term abacavir and zidovudine-based regimens, we found no evidence of decreasing body circumferences/skin-fold thicknesses-for-age associated with transition into puberty. Whilst absolute Z-scores may be influenced by the reference population, changes are more likely to be preserved. Our results reassuringly suggest that significant lipodystrophy in adolescence is unusual with these currently recommended antiretrovirals, at least in the medium term.

Our study has several limitations. A few children were already in Tanner stage 2 at their first assessment at approximately 10 years: ideally Tanner staging should have started earlier (∼9 years). Although effects of centre on age at various pubertal stages highlight the potential

for inter–observer variation to affect staging [29], centre-specific models produced broadly similar estimates, suggesting overall associations with different factors are relatively robust. Furthermore, in girls, predictors of Tanner staging were similar to menarche, which is much less subject to observer (and recall) bias. Viral load – a predictor of pubertal delay in the US meta-analysis [8] – was only measured in 49% of the ARROW children at ART initiation (those recruited post-June 2008), so could not be considered; in the US study, associations with viral load were independent of BMI and height-for-age Z-scores.

In summary, untreated HIV infection in later childhood leads to substantial delays in puberty and menarche, independent of its effect on $CD4^+$, with greater impairments in height-for-age pre-ART magnifying these effects. However, once on ART, these children appear to benefit from a prolonged window of opportunity during which they continue to grow throughout older adolescence. These results highlight the importance of considering factors other than current $CD4^+$ to guide ART initiation in older children. Future studies should explore the impact of under-nutrition, viral replication and chronic inflammation on endocrine axes to better understand the factors influencing linear growth and puberty in HIV-infected older children and adolescents.

Conflicts of interest
A.J.P. is funded by the Wellcome Trust (093768/Z/10/Z).

References

1. World Health Organisation. *Global update on HIV treatment 2013.* Geneva: World Health Organisation; 2013.
2. Joint United Nations Programme on HIV/AIDS. *Global report: UNAIDS report on the global AIDS epidemic 2013.* Geneva: UNAIDS; 2013.
3. World Health Organisation. *Antiretroviral therapy for HIV infection in infants and children: towards universal access. Recommendations for a public health approach.* Geneva: World Health Organisation; 2010.

4. World Health Organisation. *Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: Recommendations for a public health approach.* Geneva: World Health Organisation; 2013.

5. Bernays S, Jarrett P, Kranzer K, Ferrand RA. **Children growing up with HIV infection: the responsibility of success.** *Lancet* 2014; **383**:1355–1357.

6. De Martino M, Tovo P-A, Galli L, Gabiano C, Chiarelli F, Zappa M, *et al.* **Puberty in perinatal HIV-1 infection: a multicentre longitudinal study of 212 children.** *AIDS* 2001; **15**:1527–1534.

7. Buchacz K, Rogal AD, Lindsey JC, Wilson CM, Hughes MD, Seage GRI, *et al.* **Delayed onset of pubertal development in children and adolescents with perinatally acquired HIV infection.** *J Acquir Immune Defic Syndr* 2003; **33**:56–65.

8. Williams PL, Abzug MJ, Jacobson DL, Wang J, Van Dyke RB, Hazra R, *et al.* **Pubertal onset in children with perinatal HIV infection in the era of combination antiretroviral treatment.** *AIDS* 2013; **27**:1959–1970.

9. Bakeera-Kitakaa S, McKellarb M, Sniderb C, Kekitiinwaa A, Piloyaa T, Musokea P, *et al.* **Antiretroviral therapy for HIV-1 infected adolescents in Uganda: assessing the impact on growth and sexual maturation.** *J Pediatr Infect Dis* 2008; **3**:97–104.

10. World Health Organisation. Antiretroviral therapy for HIV infection in infants and children: Towards universal access. Recommendations for a public health approach. http://www.who.int/hiv/pub/guidelines/paediatric020907.pdf.

11. ARROW Trial team. **Routine versus clinically driven laboratory monitoring and first-line antiretroviral therapy strategies in African children with HIV (ARROW): a 5-year open-label randomised factorial trial.** *Lancet* 2013; **381**:1391–1403.

12. Tanner JM, Whitehouse RH. **Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty.** *Arch Dis Child* 1976; **51**:170–179.

13. Lohman TG, Roche AF, Martorell R. *Anthropometric standardization reference manual.* Champaign, IL: Human Kinetics Books; 1988.

14. De Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmanna J. **Development of a WHO growth reference for school-aged children and adolescents.** *Bull World Health Organ* 2007; **85**:660–667.

15. WHO Multicentre Growth Reference Study Group. *WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development.* Geneva: World Health Organisation; 2006.

16. Gerver WJM, de Bruin R. *Paediatric morphometrics: a reference manual.* Maastricht: Universitaire Pers Maastricht; 2001.

17. Fredriks AM, van Buuren S, Fekkes M, Verloove-Vanhorick SP, Wit JM. **Are age references for waist circumference, hip circumference and waist-hip ratio in Dutch children useful in clinical practice?** *Eur J Pediatr* 2005; **164**:216–222.

18. 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**:S172–S191.

19. Parent A-S, Teilmann G, Juul A, Skaakebaek NE, Toppari J, Bourguignon J-P. **The timing of normal puberty and the age limits of sexual precocity: variations around the world, secular trends, and changes after migration.** *Endocr Rev* 2003; **24**:668–693.

20. Hess K. **Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions.** *Stat Med* 1994; **13**:1045–1062.

21. Picat M-Q, Lewis J, Musiime V, Prendergast A, Nathoo K, Kekitiinwa A, *et al.* **Predicting patterns of long-term CD4 reconstitution in HIV-infected children starting antiretroviral therapy in sub-Saharan Africa: a cohort-based modelling study.** *PLoS Med* 2013; **10**:e1001542.

22. The Republic of Uganda Ministry of Health. Addendum to the national antiretroviral treatment guidelines. Kampala: The Republic of Uganda Ministry of Health; 2014.

23. Walters TD, Griffiths AM. **Mechanisms of growth impairment in pediatric Crohn's disease.** *Nat Rev Gastroenterol Hepatol* 2009; **6**:513–523.

24. Karpati AM, Rubin CH, Kieszak SM, Marcus M, Troiano RP. **Stature and pubertal stage assessment in American boys: the 1988-1994 Third National Health and Nutrition Examination Survey.** *J Adolesc Health* 2002; **30**:205–212.

25. Wu T, Mendola P, Buck GM. **Ethnic differences in the presence of secondary sex characteristics and menarche among US girls: the Third National Health and Nutrition Examination Survey, 1988–1994.** *Pediatrics* 2002; **110**:752–757.

26. Prentice AM, Ward KA, Goldberg GR, Jarjou LM, Moore SE, Fulford AJ, *et al.* **Critical windows for nutritional interventions against stunting.** *Am J Clin Nutr* 2013; **97**:911–918.

27. Arpadi SM, Bethel J, Horlick M, Sarr M, Bamji M, Abrams EJ, *et al.* **Longitudinal changes in regional fat content in HIV-infected children and adolescents.** *AIDS* 2009; **23**:1501–1509.

28. Alves C, Oliveira AC, Brites C. **Lipodystrophic syndrome in children and adolescents infected with the human immunodeficiency virus.** *Braz J Infect Dis* 2008; **12**:342–348.

29. Hergenroeder AC, Hill RB, Wong WW, Sangi-Haghpeykar H, Taylor W. **Validity of self-assessment of pubertal maturation in African American and European American adolescents.** *J Adolesc Health* 1999; **24**:201–205.