Abstract

Objective—To examine short term power output during growth and maturation using a multilevel modelling approach.

Methods—Body mass, stature, and triceps and subscapular skinfold thicknesses of boys and girls, aged 12.2 (0.4) years (mean (SD)) at the onset of the study, were measured at age 12, 13, and 17 years. Sexual maturation, classified according to Tanner’s stage of pubic hair development, was assessed on the first two occasions and assumed to be stage 5 at 17 years. Peak power (PP) and mean power (MP) were assessed on each occasion using the Wingate anaerobic test.

Results—Initial models, founded on 417 determinations of short term power output, identified body mass, stature, and age as significant explanatory variables of both PP and MP. The values for girls were significantly lower than those for boys, and a significant age by sex interaction described a progressive divergence in the MP of boys and girls. The introduction of sum of two skinfold thicknesses produced a model with an improvement in fit as indicated by a significant change in log likelihood. The stature term was negated and the body mass term increased. The age and sex terms were reduced but remained significant. The age by sex interaction term remained a significant explanatory variable for MP. Maturity effects were non-significant additional explanatory variables in all models of power output.

Conclusion—The values of PP and MP for boys are higher than those for girls, and, for MP, sex differences increase with age. Body mass and skinfold thicknesses are significant influences on both PP and MP, but age exerts a positive but non-linear effect on power output independent of body size and fatness.

Keywords: children; mean power; multilevel modelling; peak power

The aerobic performance of children and adolescents has been well documented but, despite recognition that the physical activity patterns of young people are characterised by short burst, predominantly anaerobic activities, data estimating anaerobic performance during growth and maturation are sparse. Direct measurement of the anaerobic performance of young people is not currently possible, and research has focused on the assessment of short term power output. The Wingate anaerobic test (WAnT), which allows the determination of cycling peak power (PP) usually over a one second or five second period and mean power (MP) over the 30 second test period, has emerged as the most widely used test of the short term power output of children and adolescents.

Cross sectional data on the short term power output of young people are conflicting and generally focused within a very narrow age range of about 10–13 years. Studies have reported a higher MP in boys, no sex differences in MP or PP, and higher MP and PP in girls. To account for differences in body size, PP and MP have been conventionally expressed in ratio with body mass (W/kg), and inconsistent data have indicated that PP and MP are significantly higher and lower in boys than in girls and that there is no significant sex difference in MP. It has been suggested that correlations observed between salivary testosterone and mass related PP and MP are indicative of the significance of boys’ sexual maturation in the development of anaerobic performance, but inadequate control of age, body mass, and stature in the analysis questions the causality of the relation.

Compelling arguments have been presented to question the validity of using ratio scaling to remove the influence of body size from size dependent performance measures, and inappropriate scaling may have clouded our understanding of the development of short term power during growth and maturation. Recognising the limitations of ratio scaling, two studies have employed linear regression to examine sex differences in PP and MP. Nindl et al compared the short term power output of 20 male and 20 female 16 year old athletes, and reported higher MP and PP in boys even with the influence of body mass partialed out using analysis of covariance. Van Praagh et al compared 15 boys and 10 girls, aged 12–13 years, using analysis of covariance to account for body mass, and observed higher PP and MP in the boys. However, the limitations of using linear regression to analyse often heteroscedastic exercise performance data have been well documented.

Armstrong et al recruited a representative sample of 200 12 year olds and controlled for body mass using an allometric analysis (log linear analysis of covariance). They reported no sex differences in PP or MP. Using the indices of pubic hair described by Tanner to determine the sexual maturity of the children, they showed a main effect for maturation.
independent of body mass. However, salivary testosterone did not increase the explained variance in boys’ PP or MP above that accounted for by body mass alone.

Longitudinal studies provide greater insights into the development of performance variables during growth, but to date only three studies of young people’s short term power have been published,12–14 and only one study included female subjects.12 Falk and Bar-Or14 reported the PP and MP of 27 boys measured four times over an 18 month period. The boys were classified for maturity according to the indices for pubic hair described by Tanner11 and, because of the small sample size, categorised into three groups: prepubertal (stage 1; n = 11), midpubertal (stages 2, 3, and 4; n = 11), and late pubertal (stage 5; n = 5). Mass related PP (W/kg) appeared to increase with each stage of maturity, but no change in mass related MP with maturation was observed. Unfortunately, the authors did not control for age effects and, as the mean ages of boys at the onset of the study varied from 10.9 years in the prepubertal group, through 13.6 years in the midpubertal group, to 16.2 years in the late pubertal group, they are likely to have confounded the analysis. The authors indicated that the experimental design precluded statistical analysis.

Duché et al7 tested 13 boys at ages 9, 10.5 and 11.5 years and 11 of the boys at 12 and 14 years, using a force-velocity test2 to determine PP and the WAnT to assess MP. The results are difficult to interpret, as data were expressed exclusively in W/kg and the method(s) of data analysis were not reported, but they indicated that both PP and MP increased from 9 to 10.5 years and from 12 to 14 years with no change observed between 10.5 and 12 years. All boys were classified as maturity stage 1 at age 12 years but ranged from maturity stage 3 to stage 5 at 14 years. The authors commented that the most important increases in PP and MP were observed when “puberty occurred”. However, the largest percentage changes in both PP and MP occurred between 9 and 10.5 years when all subjects were defined as prepubertal.

The application of allometry to longitudinal data is complex, but the recent emergence of multilevel regression modelling12 has allowed a flexible and sensitive interpretation of longitudinal data in which, for example, body size, age, maturation, and sex effects can be partitioned concurrently within an allometric framework. Armstrong et al13 retested the young people involved in their cross sectional study13 one year later and used multilevel regression modelling based on 372 determinations of short term power output to explore sex, growth, and maturational influences on PP and MP. They showed that body mass, skinfold thicknesses, chronological age, and sex were significant covariates of both PP and MP. In addition, a negative age by sex interaction term showed a smaller increase in MP for girls over the period of study. No significant maturity effects on either PP or MP were detected over this rather narrow age range.

Five years after their initial visit, at a mean age 17.0 (0.3) years, 45 of the young people in the study of Armstrong et al12 revisited the laboratory and were retested using the same procedures and equipment as in the original investigation. To extend further our understanding of the development of short term power during growth and maturation, this study uses these data to examine PP and MP at 12, 13, and 17 years. The purpose of the study is to use a multilevel modelling approach to investigate changes in MP and PP associated with age, body size, fatness, sex, and maturity.

Subjects

The subjects for the study were drawn from a longitudinal study of physical activity patterns, physiological responses to exercise, and body composition during growth. The study received ethical approval from the Exeter District Health Authority ethical committee. To ensure that the subject sample was as representative as possible, all of the children in school year 6 (aged 10–11 years) in the 15 state schools in the city of Exeter (United Kingdom) were invited to participate in the study. From this total cohort, 747 children (70% of those eligible) provided written informed consent signed by themselves and their parent or guardian. In an attempt to detect sample bias, the stature and body mass of the volunteers were compared with the stature and body mass of those who declined to participate. No significant differences (p>0.05) were detected in either sex. From the volunteers, 25% of the eligible children in each school were randomly selected.

As part of the laboratory investigations, a random sample of the young people completed a 30 second WAnT for the determination of PP and MP on two occasions at about yearly intervals. In previous publications we have reported their WAnT performance in relation to sex and maturity15 at the age of 12 years15 and examined longitudinally over the initial two years of the study.12 At age 17 years, the young people were invited back to the research centre to repeat the laboratory tests. Forty five of the subjects who had carried out the WAnT in the initial phase of the study revisited the laboratory, and, although it is impossible to ascertain whether this subsample was representative of the whole sample at age 17 years, there was no significant difference (p>0.05) between their anthropometric measures and short term power output and the rest of the sample at the first observation.

The data reported in this paper reflect the development of WAnT determined short term power output at ages 12, 13, and 17 years, and are drawn from those subjects who completed the WAnT and associated measures on the first two test occasions plus those who returned for the third test occasion. Subject numbers are therefore: test 1, 97 boys, 100 girls; test 2, 95 boys, 80 girls; test 3, 28 boys, 17 girls.

Methods

Age was computed from date of birth and date of examination. Stature, body mass, and skinfold thickness over the triceps and subscapular regions were measured by techniques
described by Weiner and Lourie. At test occasions 1 and 2, sexual maturity was visually assessed by the same experienced research nurse using the indices for pubic hair development described by Tanner. As the mean or median age at which young people reach stage 5 for pubic hair is 13.9–15.2 years for girls and 14.9–16.1 years for boys, it was assumed that at the third test occasion, age 17.0 years, the subjects had attained stage 5.

The participants had visited the laboratory on several occasions and were habituated to both the general environment and the experimental procedures. The WAnT was conducted on a friction loaded cycle ergometer (Monark 814E; Monark-Crescent AB, Varberg, Sweden) interfaced with a microcomputer. The same ergometer was used for all tests throughout the study. The seat height and handlebars were adjusted appropriately for each subject, and the resistance was set at 0.075 kg (0.74 N) body mass.

After a standardised three minute warm up involving pedalling at 60 rpm interspersed with three all out sprints lasting two to three seconds, the subject rested on the ergometer. The WAnT started from a rolling start, at 60 rpm against minimal resistance (weight basket supported). When a constant pedal rate of 60 rpm was achieved, a countdown of “3, 2, 1, go” was given, and on “go” the test resistance was applied abruptly, and simultaneously the computer was activated. Subjects remained seated and were verbally encouraged throughout the test. The power output was calculated each second for the duration of the test. The PP over one second and the MP over the 30 second period were recorded. Immediately after the test, subjects completed a cool down consisting of continuous light pedalling against minimal resistance for three minutes.

Duplicate fingertip blood samples were taken three minutes after the WAnT and immediately assayed for lactate concentration using a YSI 2300 Stat Plus whole blood analyser (Clandon Scientific, Farnborough, Hampshire, UK). The analyser self calibrated with a known concentration of lactate every five minutes, and the calibration was checked regularly against commercially prepared standards of verified concentration.

Data analysis
Descriptive statistics (means (SD)) for anthropometric variables, PP, MP, and blood lactate concentration after exercise were computed by sex for each test occasion.

Factors associated with the development of WAnT performance (age, sex, and maturity) adjusted for differences in anthropometric measures were investigated using the multilevel modelling programme MlwiN. Multilevel modelling is essentially an extension of multiple regression and is appropriate for analysing hierarchically structured or nested data. In longitudinal data sets, the hierarchy can be defined as level 1 units—that is, the repeated measurement occasions—grouped within the level 2 unit—that is, the individual subject.

Multilevel modelling is preferable to traditional analytical approaches for longitudinal data as, in addition to describing the population mean response, variation around the mean may be modelled at both levels of the analysis. For example, at level 2, children have their own growth rates which vary randomly around the underlying population response and, at level 1, each child’s observed measurements may vary around their own growth trajectory. Furthermore, unlike traditional methods based on repeated measures analysis of variance, which require a complete longitudinal data set, this method is able to handle unbalanced data—for example, where one or more measurement occasion has been missed. Similarly, as individual growth trajectories can be modelled, differing intervals between measurement occasions can be accommodated.

In this study a multiplicative allometric approach was adopted based on the model proposed by Nevill and associates as follows:

\[
y = \text{mass}^{k_1} \times \text{stature}^{k_2} \times \exp(u_1 + (b_1 \times \text{age}) + (c_1 \times \text{age}^{2})) + \varepsilon_i
\]

where all parameters were fixed except for the constant (intercept term) and age parameters, which were allowed to vary randomly at level 2 (between subjects), and the multiplicative error ratio \(\varepsilon_i\), which varied randomly at level 1 describing the error variance between occasions. The subscripts \(i\) and \(j\) denote random variation at levels 1 and 2 respectively. The variable “age” was centred on the mean age of 13.1 years.

Linearisation of the model by logarithmic transformation allows the unknown parameters to be solved using multilevel regression analysis on log \(y\). Once transformed, the equation above becomes:

\[
\log y = (k_1 \times \log \text{mass}) + (k_2 \times \log \text{stature}) + u_1 + (b_1 \times \text{age}) + (c_1 \times \text{age}^{2}) + \log \varepsilon_i
\]

From this baseline model, additional explanatory variables were investigated including sum of triceps and subscapular skinfold thicknesses, sex, and stage of maturity (stages 2–5 for pubic hair development). The latter two variables were incorporated into the model as indicator variables—for example, for sex, boys = 0, girls = 1. This sets the boys’ constant as the baseline from which the girls’ parameter is allowed to deviate. Similarly, maturity stage 1 (prepuberty) is set as the baseline from which effects due to later maturation can be assessed. The interaction term age \(\times\) sex was also constructed to investigate differential development of short term power in boys and girls.

The statistical significance of a parameter estimate is judged by dividing the value of the parameter estimate by its standard error. If this ratio exceeds \(\pm 2.0\), the estimate may be considered significantly different from zero at \(p<0.05\).

The multiplicative allometric modelling approach has been shown to be theoretically and statistically superior for longitudinal analyses to the alternative additive polynomial model as it accommodates the skewness and
Results

Table 1 presents descriptive anthropometric data, WAnT performance, and blood lactate concentrations measured after exercise. At test occasion 1, the girls were significantly (p<0.01) heavier than the boys, and, at test occasion 3, the boys were significantly (p<0.01) taller than the girls. The girls had significantly (p<0.01) greater skinfold thicknesses at each test occasion. Boys generated significantly (p<0.01) higher PP at test occasions 2 and 3 and higher MP at test occasions 2 and 3. Sex differences in blood lactate levels after exercise were not significant at each test occasion.

Table 2 summarises the results of the multilevel regression analysis for PP. In model 1, mass and stature proved to be significant explanatory variables with exponents of 0.878 (0.054) and 0.946 (0.187) respectively. There was an additional positive effect of age, but a significant negative age\(^2\) term was also observed. Girls’ PP was lower than that of boys (denoted by the parameter for sex, −0.068 (0.016), which is deducted from predicted scores for girls). All maturity effects investigated were non-significant. These fixed parameter estimates describe the subject population mean response and the random parameters reflect individual departures from this mean response at both levels of the analysis—that is, between subjects (level 2) and within subjects (level 1). The random structure of model 1 was simple, with significant variation in the constant (intercept) parameter at levels 1 and 2 evident, reflecting departures from the mean intercept term both within and between subjects. Once data were modelled for anthropometric effects, age, and sex, the level 2 random variation associated with age became non-significant.

From the baseline model 1, sum of skinfold thicknesses was investigated as an additional explanatory variable with the results summarised in table 2, model 2. The addition of skinfolds yielded a significant negative coefficient (−0.159 (0.024)) and rendered the stature term non-significant. Mass remained a significant covariate, but the value of the coefficient increased considerably above that in model 1 to 1.232 (0.050). The addition of skinfold thickness improved the fit of the model as reflected by the significant change (−513.683 to −530.243) in the deviance statistic (−2 × log likelihood) for one fewer fitted parameters. The model’s random structure remained essentially unchanged.

Table 3 summarises the multilevel models for MP. As for PP, mass and stature were initially entered as anthropometric explanatory variables (see model 1) and yielded significant exponents of 0.607 (0.055) and 1.335 (0.196) respectively. A positive term for age was obtained with negative age\(^2\) and age by sex interaction terms observed. The MP of girls was significantly lower than that of boys, as evidenced by the negative term for sex (−0.087 (0.016)). Again, all maturity effects were non-significant. The random structure showed simple random variance around the intercept at both levels of the analysis.

Tables 1 and 2 were summarised in the text to show the results of the multilevel regression analysis for peak and mean power, respectively.
The addition of skinfold thicknesses to this baseline model (model 2) rendered the stature coefficient non-significant as for PP. Parameter estimates for age and age squared remained essentially unchanged and the age by sex interaction term remained significant although reduced in magnitude. Minimal changes were observed to the random structure but the fit of this second model was considerably improved over the first, as shown by the reduction in the size of the deviance statistic from $-565.784$ to $-598.498$ for the same number of fitted parameters.

**Discussion**

There are several methodological variants of the WAnT, and this plus the analytical limitations inherent in several studies confound direct interstudy comparison. The power output and blood lactate concentrations after exercise are, however, generally consistent with the cross sectional literature at each point of observation. Armstrong and JR Welsman, unpublished data). This finding supports the cross sectional data that indicate that the anaerobic to aerobic power ratio increases through the teen years. Blood lactate concentration after exercise increased over time by 23% in girls and 31% in boys but, in accord with the extant literature, there were no significant sex differences at any point of observation.

Except for the previous report on the present sample, longitudinal data on young people's short term power output are restricted to two published studies involving boys only. These studies controlled for body mass using the ratio standard (mass exponent of one) and exponentially expressed their data in W/kg. Conflicting findings emerged, with Duché et al reporting an increase in MP and PP with age, and Falk and Bar-Or reporting no significant change in either MP or PP with age. Both studies alluded to a positive effect of maturity on MP and PP but an inappropriate means of controlling for body mass and not controlling concurrently for other covariates threaten the security of the results and confound further interpretation.

The multilevel regression analyses used in the current study allow simultaneous investigation of covariates likely to influence PP and MP and show their interdependence. For example, model 1, in tables 2 and 3, shows that stature is a significant independent predictor of MP and PP when included alongside mass in an allometric analysis. However, as shown in model 2 in tables 2 and 3, the addition of sum of triceps and subscapular skinfold thicknesses to the initial model rendered the stature term redundant and increased the size of the mass exponent. This finding is consistent with recent analyses of peak oxygen uptake and emphasises the importance of concurrent analyses of covariates when investigating developmental changes in exercise performance variables.

The negative parameter estimate for sex indicates the girls' lower PP and MP, even with age and anthropometric variables controlled for. The negative age by sex interaction term which is significant in both models for MP shows a smaller increase in MP with age for girls over the period studied. The faster oxygen uptake kinetics in children during high intensity exercise may represent less dependence on anaerobic responses early in exercise in children than in adults, and, during the 30 second test, children's MP is supported by a significant contribution from aerobic metabolism, which has been estimated to vary between 18 and 44%. The interplay between aerobic and anaerobic metabolism makes the age by sex interaction for MP difficult to interpret, but no published data show sex differences in either oxygen uptake kinetics or aerobic contribution to 30 second high intensity exercise during childhood and adolescence, despite the well documented sex difference in peak oxygen uptake.

The addition of the sum of two skinfold thicknesses to the baseline models not only increased the size of the mass exponent but also reduced the magnitude of the sex term. This reflects the greater increase in skeletal muscle mass in boys over the period 12–17 years. Skeletal muscle is responsible for the generation of power, and dramatic sex changes in muscle mass occur during the adolescent period. Boys' muscle mass reaches about 54% of body mass at age 17 years, whereas, in girls, muscle mass is about 45% of body mass at 13 years and then, in relative terms, declines because of fat accumulation at adolescence. Power output in the WAnT develops from the legs, and, by the end of adolescence, boys have only about 70% of the lower limb muscle mass of boys. Furthermore, changes in muscle pennation with increased muscle size may influence force output, and Blimkie and Sale have reported substantial increases in muscle pennation angle with growth in both sexes for the vastus lateralis and gastrocnemius medialis. In females, however, fascicle angles appear to plateau during late adolescence, although in males they continue to increase until early to mid-adulthood. The WAnT methodology itself may therefore accentuate sex differences in power output, particularly for MP. The recommended resistance for use with a Monark cycle ergometer is calculated as a simple per body mass ratio and does not accommodate differences in body fatness or muscularity.

Power is the product of force and velocity. The maximum force of a muscle is determined by the number of sarcomeres arranged in parallel and is consequently proportional to the cross sectional area of the muscle or the square of the linear dimensions ($l^2$). The velocity of muscle shortening depends on the number of sarcomeres in series and is directly proportional to the length of the muscle. Power is accordingly related to $l^2$ and best reflected by measurement of muscle volume. PP and MP
generated during the WanT are therefore likely to be influenced by thigh muscle volume. Even with young children, girls have been shown to be exercising against a higher resistance than boys when the WanT resistance is expressed relative to their magnetic resonance imaging (MRI) determined thigh muscle volume.6 Initial indications from a continuing longitudinal study in our centre suggest that MRI determined thigh muscle volume is a significant additional explanatory variable for PP and MP between the ages of 10 and 12 years, even with body size, skinfold thicknesses, and age controlled for.

Muscle action is not just a function of size and power output, it is dependent on muscle fibre type. The percentage distribution of type II fibres is age dependent,26 and an adult profile is not reached until late adolescence.27 Some28 but not all29 studies have indicated a greater prevalence of type IIA than type IIX fibres during childhood and adolescence. There is apparently no clear sex difference in either fibre distribution or size during early childhood but, from childhood to adolescence, fibre size increases 3.5-fold in girls and 4.5-fold in boys.28 Fibre diameter peaks in adolescence in girls whereas boys achieve peak fibre diameter in young adulthood.30 Boys have been reported to increase the size of their type IIX fibres more than girls.31

Larson and Moss32 have shown that, in the quadriceps of adults, type I and type IIX fibres have respectively three and 10 times greater maximum shortening velocity than type I fibres. If these relative characteristics are similar during childhood and adolescence, they will contribute to both age and sex related influences on short term power output.

In accord with our earlier report of these children at 12 and 13 years, a significant effect of age over and above that due to anthropometric covariates was observed for both power measures. However, in this wider age range, a sex covariate was observed for both power of age over and above that due to anthropomet-

We gratefully acknowledge the technical assistance of Alison Husband and Dr Craig Williams. The work was supported by the British Heart Foundation, the Healthy Heart Research Trust, and the Norman Family Trust.

Take home message

Body mass and skinfold thicknesses are significant influences on short term power output in 12–17 year olds. The values for boys are higher than those for girls even with body mass and fatness controlled for. Independently of body mass and fatness, age exerts a positive but non-linear effect on power output, but maturity does not have a significant independent effect on either peak or mean power as determined using the Wingate anaerobic test.