Secular trends in physical fitness in Danish adolescents

Andersen LB[1,2], Froberg K[2], Lund Kristensen P[2], Moller NC[2], Resaland GK[1,3],

Anderssen SA[1]

[1]Norwegian school of Sport Sciences, Norway

[2]University of Southern Denmark, Denmark

[3]Sogn and Fjordane University College, Norway

Abstract

It is important to monitor changes over time in aerobic fitness and obesity in order to target interventions when health deteriorates. We analyzed data from three population studies of adolescents carried out in 1983, 1997 and 2003. Participants were 1050 adolescents from the whole country of Denmark in 1983, 410 in 1997 and 418 in 2003, the two latter cohorts from the City of Odense. VO_{2max} was estimated from maximal power output (MPO) in a cycle test with progressively increasing workload. Estimated VO2max for boys and girls was 52 ml·min^{-1}·kg^{-1} and 41 ml·min^{-1}·kg^{-1}, respectively, both in 1983, 1997 and 2003 with no difference between the three cohorts. However, body mass index (BMI) increased 10% in the upper decentile of the distribution since 1983. MPO decreased over time, but validation studies showed that this was not due to decreased VO_{2max}. The cohort from 1983 was tested twice by school teachers and experienced scientists, and the scientists found higher MPO than school teachers in the same subjects, which emphasize the importance of good validation studies.

Key words: maximal power output, adolescence, maximal oxygen uptake, secular trends

Introduction

Cardiorespiratory fitness and obesity are strongly related to clustering of cardiovascular (CVD) risk factors. We recently found a 13 fold increased risk of clustered cardiovascular risk in the least fit quartile compared to the most fit quartile of children and adolescents. We also found that low fitness and overweight independently were associated to clustered CVD risk (Andersen et al. 2008). Fitness is a modifiable CVD risk factor, and it is therefore important to monitor changes in fitness in the population in order to be able to intervene if levels decline. Direct measurement of maximal oxygen uptake is the golden standard for assessment of cardiorespiratory fitness, but very few population based studies using this method are available (Andersen et al. 1987; Knuttgen 1968), which makes it difficult to evaluate changes since reference data rarely exists. Some studies have been published where changes have been reported from indirect estimation of fitness, and Tomkinson and Olds (Tomkinson and Olds 2007) recently reviewed data from all over the world and stated: 'Over the 45-year period, there has been a global decline in aerobic performance of -0.36% per annum. They also found that secular changes have been very consistent across age, sex, and geographical groups. The pattern of change however, was not consistent over time, with improvements from the late 1950s until about 1970, and declines of increasing magnitude every decade thereafter. A decline in fitness is supported by a large Norwegian study, where fitness decreased 8% in subjects examined prior to military service from 1980 to 2002 (Dyrstad et al. 2005). This study used a submaximal cycle protocol.

However, methodological problems make it difficult to evaluate changes. Submaximal tests have low reproducibility (Andersen et al. 1987) and probably suffers from systematic bias (Andersen and van Mechelen 2005). Furthermore, maximal tests have systematic bias related to differences in ergometers and are highly dependent on the test leader.

The aims of the present study were to analyze secular trends in cardiorespiratory fitness, and elucidate problems related to assessment of fitness from indirect estimation of VO_{2max} during a maximal cycle ergometer test. We analyzed data from three different population studies performed in 1983, 1997 and 2003, where a cycle protocol with progressively increasing work load was used. The older data included direct measurements of VO_{2max} and repeated tests making it possible to analyze inter-tester and intra-tester reliability, and the later studies included a thorough validation of MPO against direct measurement of VO_{2max}.

Methods

Danish Youth and Sport study (DYSS)

All 1260 subjects, 710 girls and 550 boys, 15-19 years of age, attending randomly selected classes from 18 high schools and 18 vocational and trade schools in Denmark participated in the study in 1983. Students attending these types of school comprised 80% of the population. Two years later the sample was compared to a random sample selected from the National Personal Register (n=107). No difference was found in VO_{2max} measured directly with the Douglas bag method between the two samples. A few students left school before the data collection started, but among the students still attending the 36 schools, when the study started, all participated in the study. However, only 1050 (83%, 466 boys and 584 girls) reached a maximal HR of 185 bpm and were included in the analysis. This criterion for an approved max test was chosen in order to use the same criterion in all studies, and 185 bpm was used in the EYHS studies where measurement of oxygen uptake was not performed. Another criterion used in all cohorts was the subjective judgment from the test leader that subjects reached exhaustion.

Height and weight were measured to the nearest 1 cm and 0.5 kg with a Harpenden stadiometer and a Seca beam scale, and body mass index (BMI) calculated as weight·height^{-2}

Aerobic power was measured in a progressive cycle ergometer test. Subjects warmed up for 7 minutes at 69 watts and 103 watts for girls and boys, respectively, on a mechanically braked cycle ergometer at 70 rpm (Monarch, Sweden). Mean HR during the

last minute was recorded. Thereafter, the intensity was increased 35 watts (0.5 kp) every 2 minutes until exhaustion. Maximal power output (MPO) was calculated as the power output before the last intensity increase, plus 35 watts/the percentage of the 2 minutes cycled during the final stage. In 1983, 293 subjects were randomly selected from the cohort of 1260 subjects and tested a second time 1-4 weeks after the first test. The dropout at the second test was 0.7% (two girls), and 152 girls and 120 boys fulfilled the HR criteria of 185 bpm for a valid test. We chose to use the same inclusion criteria as for the other cohorts where direct measurement of VO_{2max} was not performed. At the second test expired air was collected in Douglas bags and O_2 content measured by an Applied Electrochemistry oxygen analyzer (s-3a)(AEI Technologies, Illinoise). Volumes were measured with a dry gas meter. VO_2 was calculated assuming a respiratory exchange ratio (r) of 1.12, which was the r-value found in the first 60 tests where CO_2 was measured (for details reference (Andersen et al. 1987)). Heart rate was measured with a Cardioline during these tests (Cardioline, Italy). From this test where maximal oxygen uptake (VO_{2max}) was measured directly, a calibration curve was calculated by linear regression between maximal power output measured by the gym teachers and VO_{2max} (r=0.90, p<0.001). The following equation was used to estimate VO_{2max} ($l \cdot min^{-1}$) from maximal power output in DYSS:

VO_{2max} ($ml \cdot min^{-1}$)= (6.36*watts) + (607*sex) + 1129 ; girls=0 and boys=1

EYHS1 population

Schools were stratified according to school type, location and the socio-economic character of its uptake area. From each stratum, a proportional, two-stage cluster sample

of children was selected. The primary units (clusters) were the schools. The sampling frame for schools was a complete list of public schools in Odense, from which schools were selected using probability proportional to school size. Each school on the list was allocated a weighting equivalent to the number of children enrolled who were eligible for selection into the study.

The secondary units were the children within the schools, and equal numbers of children were selected from each school. Children within the appropriate age bands (8-10 yr and 14-16 yr) were allocated code numbers and randomly selected using random number tables. The 14-16 year old children from EYHS-I constituted the population tested in 1997-98, and the 8-10 year old children from EYHS-I were tested again in EYHS-II 6 years later and constituted the population from 2003-4. Twenty-eight out of 35 schools were randomly sampled, and 25 agreed to participate. Of the three non participating schools, one was rural, one was urban from middleclass area and one was urban from a low income area. All three schools gave interference with the educational process as reason for not participating.

Two hundred twenty four female and 206 male adolescents (36.1% of the ninth grade population in Odense) participated, but 20 of these were excluded with an invalid test (8 boys and 12 girls had HRmax below 185 bpm). Mean age of the adolescents were 15.5 years.

EYHS2 population

Seven hundr ed seventy o ne third gra de childre n from 25 different schools were invited to participate in EYHS-I (Wedderkopp et al. 20 04). A total number of 58 9 children (310 girls and 279 boys) pa rticipated in the stud y, cor responding to 76. 4 % of the invited c hildren. Six years later 438 of t hese children were re- examined in EYHS-II as ninth grade stude nts, corresponding to 74 % of the total number of participants in EYHS-I. One of the reasons for non-participation in the follow-up study was reluctance to travel to the city of Odense for the examinations. Several of the participants in EYHS1 no longer lived in the Community of Odense at the time EY HS2 was initiated. This dropout is probably random. Among these 20 subjects had not a valid fitness test, which left 234 girls and 184 boys with valid test.

The sample o f 589 particip ants in EYHS-I have b een shown to be a representative subset of the total sam ple of 771 i nvited children with respect to the param eters: phy sical activit y level and body composition (Wedderkopp et al. 2001). Other parameters have not been investigated. Possible dropout effects in EYHS2 were examined by com paring baseline values between the group of children who participated in both studies and the gr oup who only participated in EYHS-I. No significant differences in mean BMI or fitness was found between dropouts and participants in EYHS2.

EYHS1 and EYHS2 cycle test

The cycle ergometer was a Monark 839 Ergomedic (Sweden) in 1997-98 and in 2003. The bikes were pre-programmed to increase the workload every third minute with 40 or 50 watts for girls and boys, respectively. The workload was increased until exhaustion, and the time and heart rate (HR) were registered (Polar, Finland). Heart rate was measured using Polar heart rate monitors, Polar 3000 in the first study and Polar Vantage NV in the second study. The Polar 3000 used integration of the heart rate for every 5

seconds for calculating the reported heart rate per minute, where as the Polar Vantage NV used beat by beat integration for the heart rate per minute measurement. Criteria for exhaustion were 1) heart rate above 185 bpm, 2) that the child could not keep a pedalling frequency of 30 rpm per minute or more and 3) a subjective judgement of the observer that the child could no longer keep up, even after vocal encouragement. All tests were performed by only two researchers, one in each cross-sectional study. The maximal power output (Wattmax) was calculated as the watts in the last completed workload (W_l), plus the increment in watts (W_i) of the last step divided by 180 seconds multiplied by the number of seconds completed of the last step (t_{ls}).

Wattmax=W_l+(W_i x t_{ls}/180sec.).

VO$_{2max}$ was estimated from MPO from an equation derived after using the same protocol an d the same cycle ergometer in 2 79 girls and 312 boy s from another r epresentative sample of the same age, where di rect measurement of VO$_2$ was assessed (Metamax 3X,, Cortex, G ermany) together with MPO. The following equation was calculated by linear regression:

VO_{2max} (ml·min^{-1})= (10.5*watts) + (225*sex) + 560; girls=0 and boys=1

Ethics

The European Youth Heart Study 1 and 2 was approved by the local ethics committee and performed by the rules stipulated by the Helsinki declaration. All children gave verbal consent and their parents gave written consent. The Danish Youth and Sport Study was conducted before the Ethical Committee was formed.

Statistics

The Statistical Package for Social Sciences version 15 was used. VO_{2max} was related to age in DYSS where the age span was 15-19 years of age. Subjects in this study was 15-19 years of age (mean 17.1 years), and the age of EYHS1 and EYHS2 were 15.5 years and 15.8 years, respectively. We therefore calculated the linear regression with age for MPO, MPO/kg body weight, VO_{2max}, and BMI in DYSS, and adjusted all cohorts to the age of 16 years in these variables. Regression equations used for this adjustment were: Regressions with age in DYSS where data from the whole cohort of 1050 subjects was used:

Boys: MPO=151.4 watts + (7.0*age);   Girls: MPO=173.1 watts + (1.36*age)

Boys: Watts/kg=3.74 W/kg + (0.020*age);   Girls: Watts/kg=4.14 W/kg - (0.044*age)

Boys: BMI=16.0 kg/m2 + (0.276*age);   Girls: BMI=14.9 kg/m2 + (0.335*age)

The whole distribution of fitness and BMI were compared by ranking the variables into decentiles. This procedure enabled us to see if only part of the population had changed, and at the same time to some extend control for bias. Indirect estimation of fitness can introduce systematic bias, because the tests were performed by different researchers and on different ergometers.

Results

Descriptives of the three cohorts tested in 1983 (DYSS), 1997 (EYHS1) and 2003 (EYHS2) are presented in Table 1. Also, the randomly selected subgroup from DYSS, 129 boys and 163 girls, who were tested a second time by experienced physiologists is described in Table 2.

Insert table1 and table 2.

Fitness:

Figure 1 and Figure 2 shows the MPO in deciles for each cohort. Of note is the difference between the main group tested by the gym teachers in 1983 and the subgroup randomly selected from the main group tested by scientists. Further, test results were compared between first and second test in 1983 (paired t-test between selected group tested by gym-teachers and by physiologists). As the same subjects were included, the differences in results were attributed to the test leader or the ergometer, and this analysis illustrates the bias we would expect in fitness in studies where direct measurements of oxygen uptake is not available. In boys, the HR_{max} did not differ (p=0.976), maximal power output was 9 watts higher (SD 32 watts) (p<0.01), and weight was 0.6 kg (SD 1.7 kg) higher (p<0.001). In girls, HRmax was 2.6 bpm (SD 11.8 bpm) lower (p<0.05), maximal power output 4 watts (SD 30 watts) higher (p=0.011), and weight 1.0 kg (SD 2.6 kg) higher at the second test (p<0.001). VO_2 will level off in many subjects in the end of a maximal test with increasing work load, but work load will continue to increase until the end of the test. The difference between experienced and non-experienced test leaders may therefore be bigger in MPO than in VO_{2max}. Because results of MPO in the maximal cycle test from the test performed by experienced physiologists are higher than results obtained by the teachers we decided to present data on absolute workload both from the main group in 1983 and the selected group (fig. 1 and 2).

Insert Figures 1 and 2

MPO was used to estimate VO_{2max} from the equations calculated in each of the validation studies. Results presented in percentiles of fitness are shown for boys and girls in Figures 3 and 4, respectively. Comparisons of cohorts in key variables are shown in table 3.

The changes in BMI were not just changes in mean values and BMI is therefore presented in decentiles, which shows a polarization over time. The upper part of the distribution increased BMI over time, while no change was seen in the low end of the distribution.

Insert figures 3 and 4

Insert table 3

Discussion

Three cohorts tested in 1983, 1997 and 2003 were analyzed. We found no change in cardiorespiratory fitness over time, and quite high levels were found in these representative cohorts. This study shows how difficult it is to analyze trends in physical fitness, when direct measurement of VO_{2max} is not available. We found substantial differences in MPO, but no differences when VO_{2max} was estimated from the equations derived in the validation studies. This could suggest a difference in the ergometers even if these were calibrated repeatedly. Different protocols could also be an explanation, but the mean increase in watts per minute was very similar even if work load was increased every second minute in DYSS and every third minute in EYHS. This type of bias is not possible to detect unless good validation studies are available for each specific cohort. Also, a difference in MPO was found between the same subjects tested by school teachers and scientists. Different cycle ergometers were used in these tests, but this difference could also be explained by the fact that Saltin and Andersen did all the tests in the subgroup, and that they might have been more experienced than the school teachers. However, when VO_{2max} was estimated from the equation derived from the regression between direct measurement of VO_{2max} from the scientists' test and MPO from the teachers' test, this difference also disappeared. No studies have yet analyzed secular trends in cardiorespiratory fitness using direct measurement of VO_{2max}, and it will not be done in the near future in this age group, because reference data are only available in two studies, where population representative samples have been tested with direct measurements (Andersen et al. 1987; Knuttgen 1968).

The most reliable type of indirect assessment of VO_{2max} is from maximal testing, and cycling is a good alternative in Denmark where two third of adolescents cycle to school every day. However, even if great care was taken to ensure standardized conditions, our data shows that most population studies may have systematic bias if changes are evaluated from maximal power output. We were able to take inter-tester differences into account and had good data on intra-individual variation. We also had good validation studies for each time point, and when we analyzed the estimated VO_{2max}, we found no change in fitness since 1983. This was despite a polarization in BMI had occurred. More fat adolescents were found in the later cohorts similar to what many other studies have found (WHO 2007). However, when a whole BMI distribution shift to the right this may reflect two things. First, there are more obese children where the extreme part of the distribution has very low fitness level. This is still a minority. Second, the right shift of the major part of the distribution, which has occurred, may reflect a better nutritional status and better health, which is supported by that fact that height has increased substantially since the early 1900. It may therefore not be surprising that fitness has not decreased in the present study.

Other studies have analyzed secular trends in cardiorespiratory fitness. Rowland looked critically at the literature and concluded that the evidence was scant (Rowland 2002). There was some evidence of deteriorating endurance performance evaluated from different field tests, but studies had methodological problems and the changes in different fitness measures were not consistent. Rowland also plotted values of maximal aerobic power from studies over a 30 year period and found no evidence of change. However,

studies have used different methodologies and mainly reflect the values of subjects

volunteering for exercise studies. Since Rowlands study, Westerstahl et al. have found a

3% decrease in aerobic fitness in Sweden from 1974-1995 estimated from a 9-min

run/walk test (Westerstahl et al. 2003). They measured a number of different types of

fitness, and some increased while most decreased. The run/walk test, which is the test

estimating cardiorespiratory fitness, was carried out on a 400 m outdoor track. It is not

possible from the report to see whether the surface was the same type in 1974 and 1995.

Another Swedish study by Ekblom et al studied changes in aerobic fitness in 20-65 year

old adults from 1990 to 2000 (Ekblom et al. 2007). They used a submaximal cycle test

and estimated fitness level was extremely low compared to Danish data. They found no

change in 20-29 year old women, but a 10% decrease in men of this age. It is not possible

to see how many subjects the analysis included from this age group. Dyrstad et al.

analyzed data from Norwegian military conscripts from 1980-2002, and found an 8%

decrease in estimated VO_{2max} (Dyrstad et al. 2005). This study used submaximal cycle

ergometer test, which has a low reproducibility and may have severe bias related to the

ergometers and manual measurement of heart rate (Andersen and van Mechelen 2005).

The latest comprehensive study was published by Tomkinson et al. who collected data

from 55 different studies from 11 countries (Tomkinson et al. 2003). They found a

0.43% decline per year in running performance in the 20m shuttle run test, and in the

literature they reviewed, where other aerobic fitness field tests were used, they found a

similar picture. However, the change differed quite much between countries. Both

Rowland and Tomkinson et al. pointed against the fact that fatness have increased in the

same period of time, and that this may explain the change in running performance,

because subjects carry their own weight. We found an increase in fatness in the upper percentiles of BMI, but no difference among the lean part of the population. This increase is reflected in the extreme low part of the fitness distribution, but because the number of fat adolescents is rather low, it is difficult to detect on population level. In other countries with a more severe obesity problem, it may be reflected in population values in fitness.

We have mentioned that comparison of fitness levels between studies is difficult with indirect estimation of VO_{2max}. Even with direct measurements there can be problems related to test protocol, choice of ergometer and type of procedure for VO_2 analysis, i.e. on-line systems and the way they select peak values may give different results than Douglas bags. However, differences between populations do exist. When data are compared from the studies of Knuttgen and Andersen et al. it is striking that they find the exact same values for VO_{2max} in boys using the same type of protocol and Douglas bags for assessment of VO_2 (Andersen et al. 1987; Knuttgen 1968). However, in girls where they also used the same methods, Danes had more than 20% higher VO_{2max}. This type of difference may reflect a different culture in relation to girls' sports participation. It can be concluded that there are differences in aerobic fitness between adolescent populations in different countries, and aerobic fitness may have decreased over the last decades, because most studies point in that direction. At the same time most studies have used methods, which are problematic, and it cannot be ruled out that the conflicting results may be related to methodological problems.

Strengths and limitations

The main strengths of this study are large representative cohorts, comprehensive validation studies against direct measurement of VO_{2max}, and quite detailed data on parameters that could cause bias. There was no dropout in the reference population, and subjects were sampled from the whole country. The lack of dropout was caused by the fact that the fitness testing was performed during school time, and the teachers considered it as a part of the curriculum. This study was performed before the Ethical Committee existed. We may question the ethics today, but data do provide a good reference.

The observation that MPO decreased, but no change was found in estimated VO_{2max} is interesting, because the conclusion would have been different if comprehensive validation studies had not been done. The validation in the DYSS was done in about 300 subjects and in EYHS in about 600 subjects. Smaller validation studies can easily include bias, and a less well define slope and intercept of the regression line.

The EYHS cohorts were limited in their representativeness because subjects were sampled from a smaller geographical area, and a dropout around 30% was seen as in most other epidemiological studies. The fitness level of the subjects in DYSS who came from this geographical area was not different from the rest of the country, which could suggest that there was no major selection bias. We would have liked to be able to compare direct measurements of VO_{2max}, but we believe that the good validation studies provide the second best option for a valid comparison.

Perspectives

The present study questions how much cardiorespiratory fitness has declined during the last decades. It seems clear that a smaller part of the population has increased body weight and some have lower fitness than earlier seen. This observation could lead to the suggestion that intervention strategies should be aimed against not just the obese children but also include children with very low fitness.

Reference List

1. Andersen LB, Henckel P, Saltin B. Maximal oxygen uptake in Danish adolescents 16-19 years of age. Eur J Appl Physiol 1987: 56: 74-82.

2. Andersen LB, Sardinha LB, Froberg K, Riddoch CJ, Page AS, Anderssen SA. Fitness, fatness and clustering of cardiovascular risk factors in children from Denmark, Estonia and Portugal: the European Youth Heart Study. Int J Pediatr Obes 2008: 3 Suppl 1: 58-66.

3. Andersen LB, van Mechelen W. Are children of today less active than before and is their health in danger? What can we do? Scand J Med Sci Sports 2005: 15(5): 268-270.

4. Dyrstad SM, Aandstad A, Hallen J. Aerobic fitness in young Norwegian men: a comparison between 1980 and 2002. Scand J Med Sci Sports 2005: 15(5): 298-303.

5. Ekblom B, Engstrom LM, Ekblom O. Secular trends of physical fitness in Swedish adults. Scand J Med Sci Sports 2007: 17(3): 267-273.

6. Knuttgen HG. Aerobic capacity of adolescents. J Appl Physiol 1968: 22: 655-658.

7. Rowland T. Declining Cardiorespiratory Fitness in Youth: Fact or Supposition? Pediatr Exerc Sci 2002: 14: 1-8.

8. Tomkinson GR, Leger LA, Olds TS, Cazorla G. Secular trends in the performance of children and adolescents (1980-2000): an analysis of 55 studies of the 20m shuttle run test in 11 countries. Sports Med 2003: 33(4): 285-300.

9. Tomkinson GR, Olds TS. Secular changes in pediatric aerobic fitness test performance: the global picture. Med Sport Sci 2007: 50: 46-66.

10. Wedderkopp N, Andersen LB, Hansen HS, Froberg K. [Obesity among children-- with particular reference to Danish circumstances]. Ugeskr Laeger 2001: 163(21): 2907-2912.

11. Wedderkopp N, Froberg K, Hansen HS, Andersen LB. Secular trends in physical fitness and fatness in Danish 9-year old girls and boys. Odense School child Study and Danish substudy of The European Youth Heart Study. Scand J Med Sci Sports Exerc 2004: 14: 1-6.

12. Westerstahl M, Barnekow-Bergkvist M, Hedberg G, Jansson E. Secular trends in body dimensions and physical fitness among adolescents in Sweden from 1974 to 1995. Scand J Med Sci Sports 2003: 13(2): 128-137.

13. WHO. The challenge of obesity in the WHO European Region and the strategies for response. Branca FNH, Lobstein T. Copenhagen, WHO Regional Office for Europe. 2007: 1-56.

Legend to figures

Figure 1 and 2. Maximal power output per kg body weight in boys (1) and girls (2). These results indicate a major decrease in fitness, which would have been the conclusion if no validation study or direct measurements had been available.

Figure 3 and 4. Estimated fitness using equations from the validation studies in boys (3) and girls (4). These results indicate no difference over time. Data were adjusted to 16 years of age in all cohorts.

Figures 5 and 6. Details of changes in BMI over time. Changes mainly occur in the upper part of the distribution.
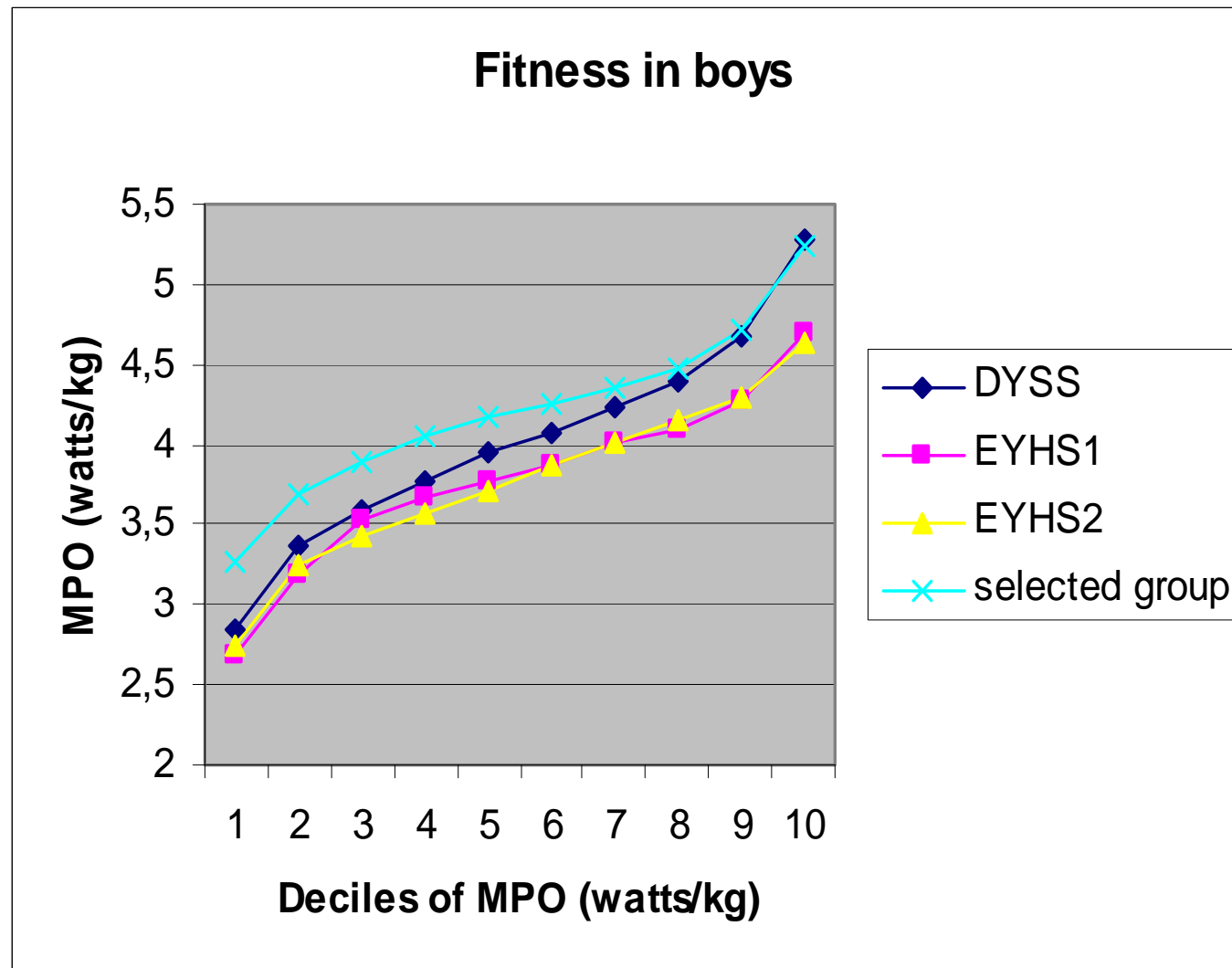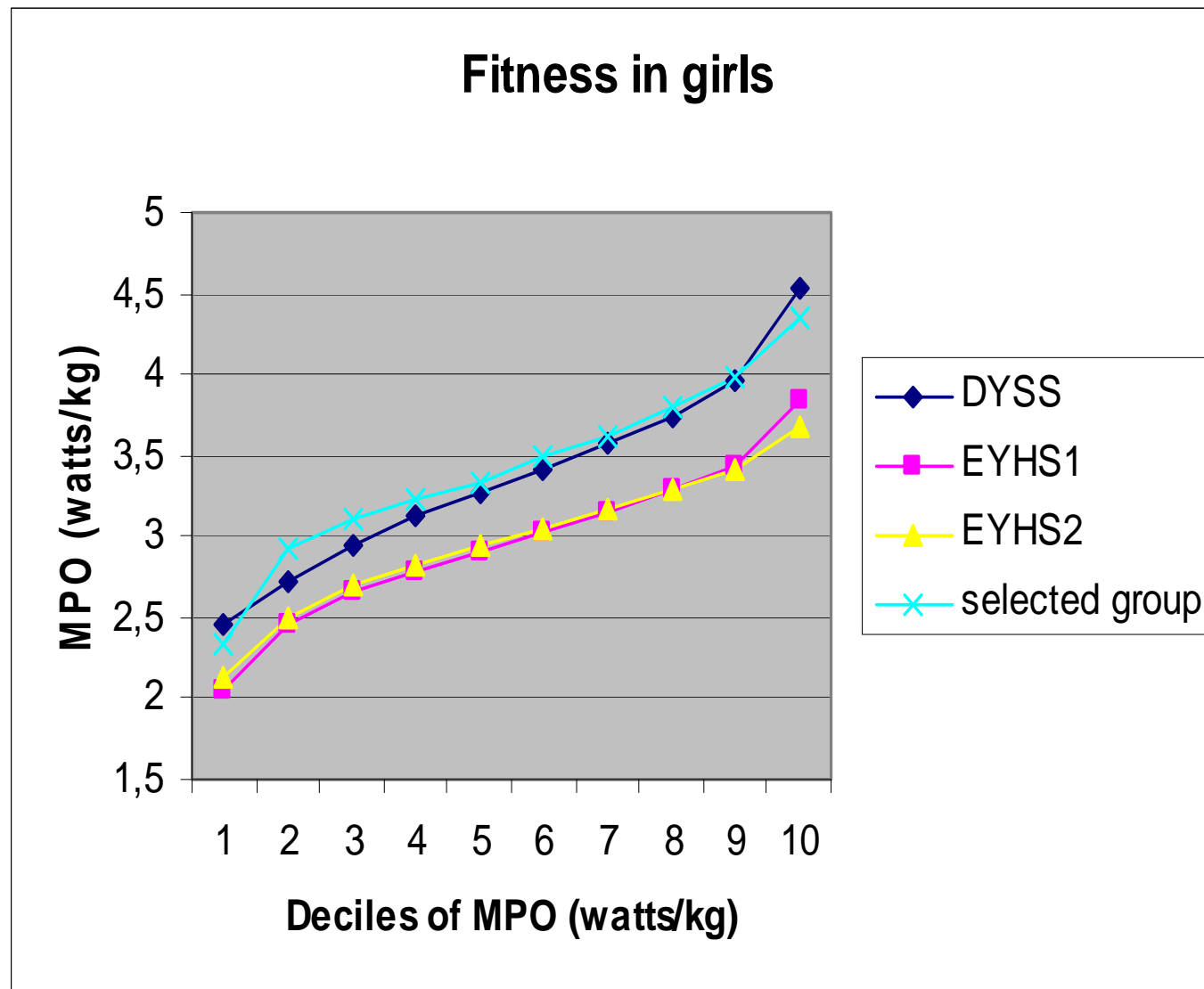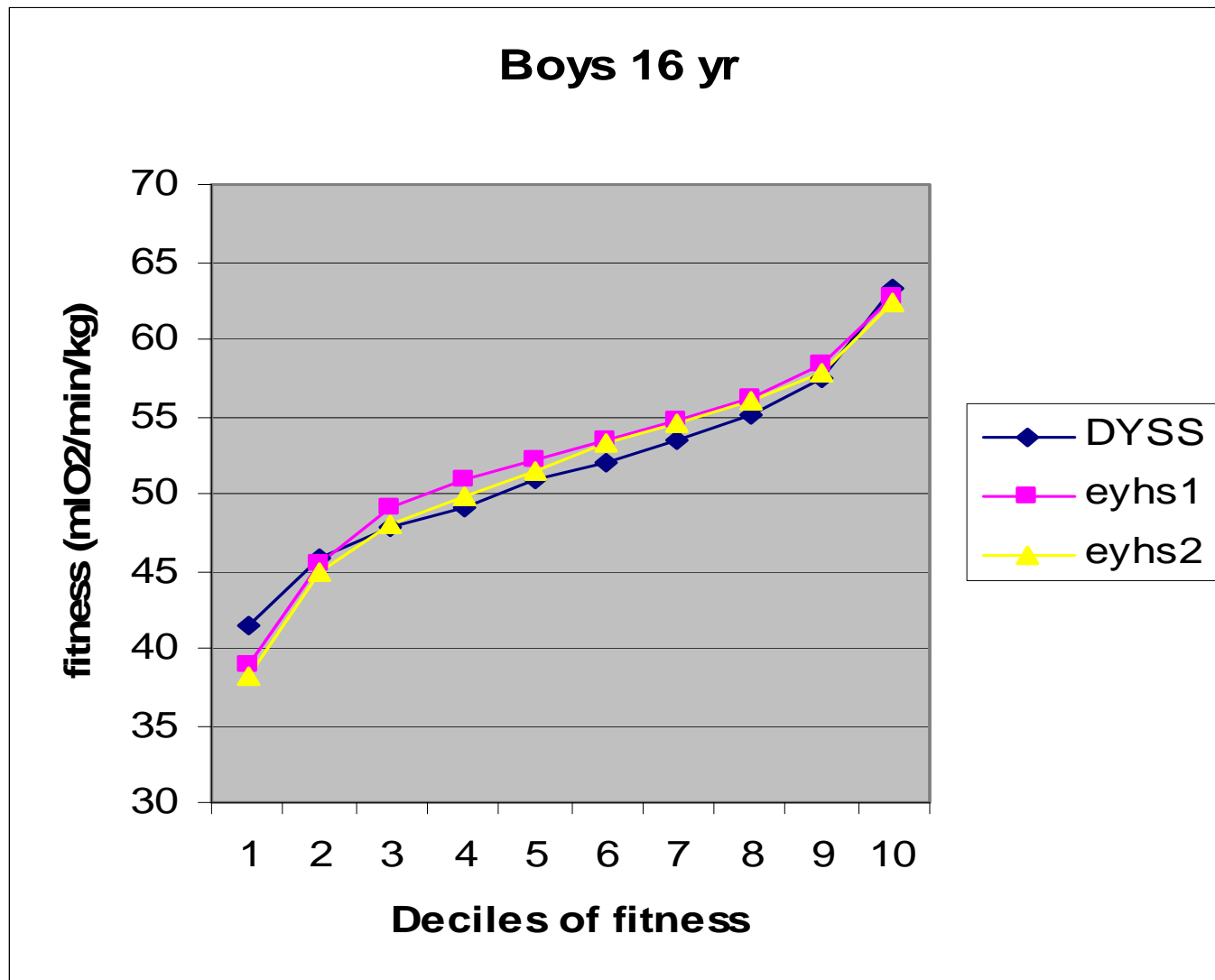
Figure 1.

Fitness in boys

Figure 2.

Fitness in girls

MPO (watts/kg) vs Deciles of MPO (watts/kg)

Legend:
- DYSS
- EYHS1
- EYHS2
- selected group

Figure 3.

Figure 4.

Figure 5.

BMI in boys

Figure 6.

BMI in girls

Y-axis: BMI (kg/m2) — 15, 17, 19, 21, 23, 25, 27, 29

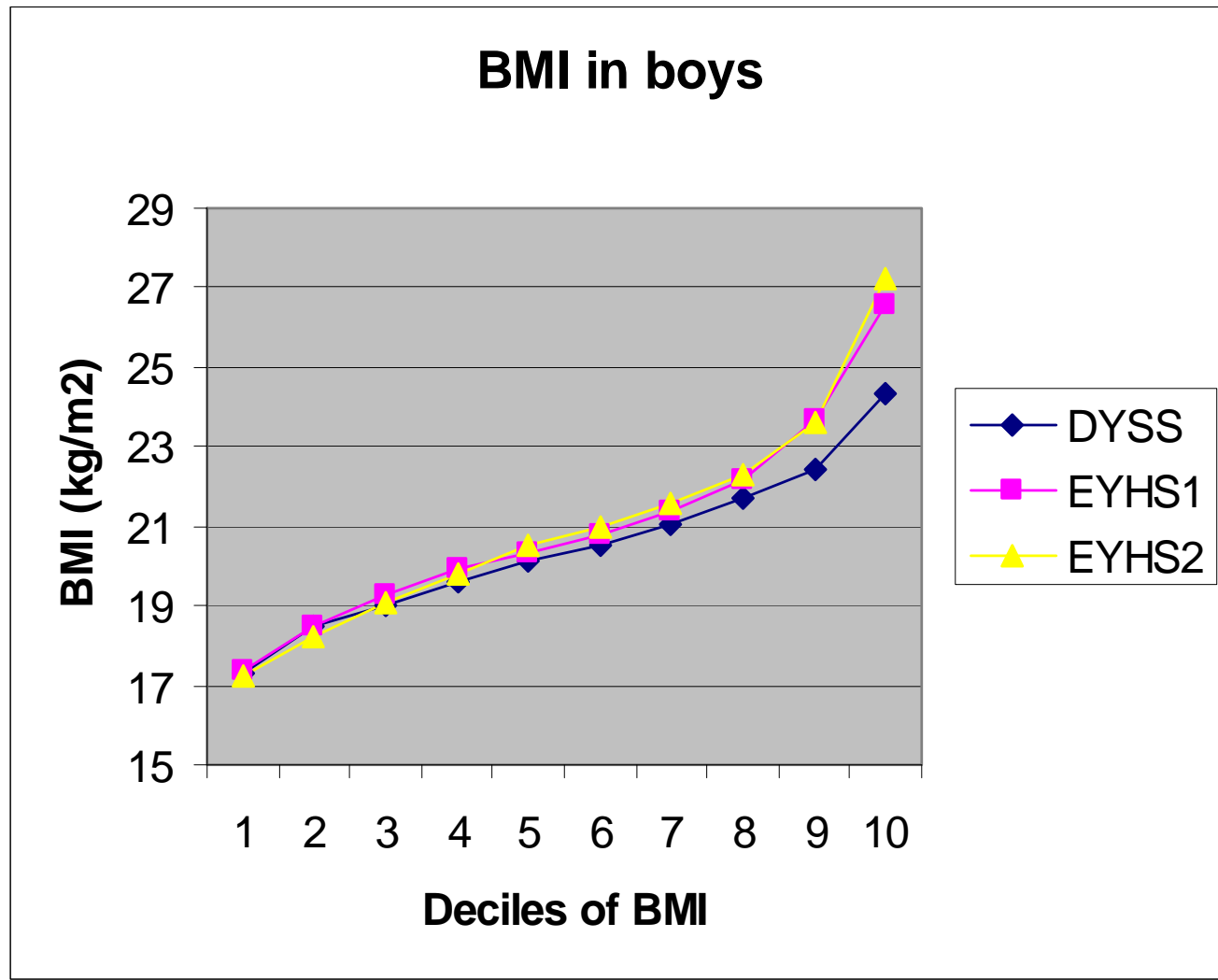_X-axis:_ Deciles of BMI — 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
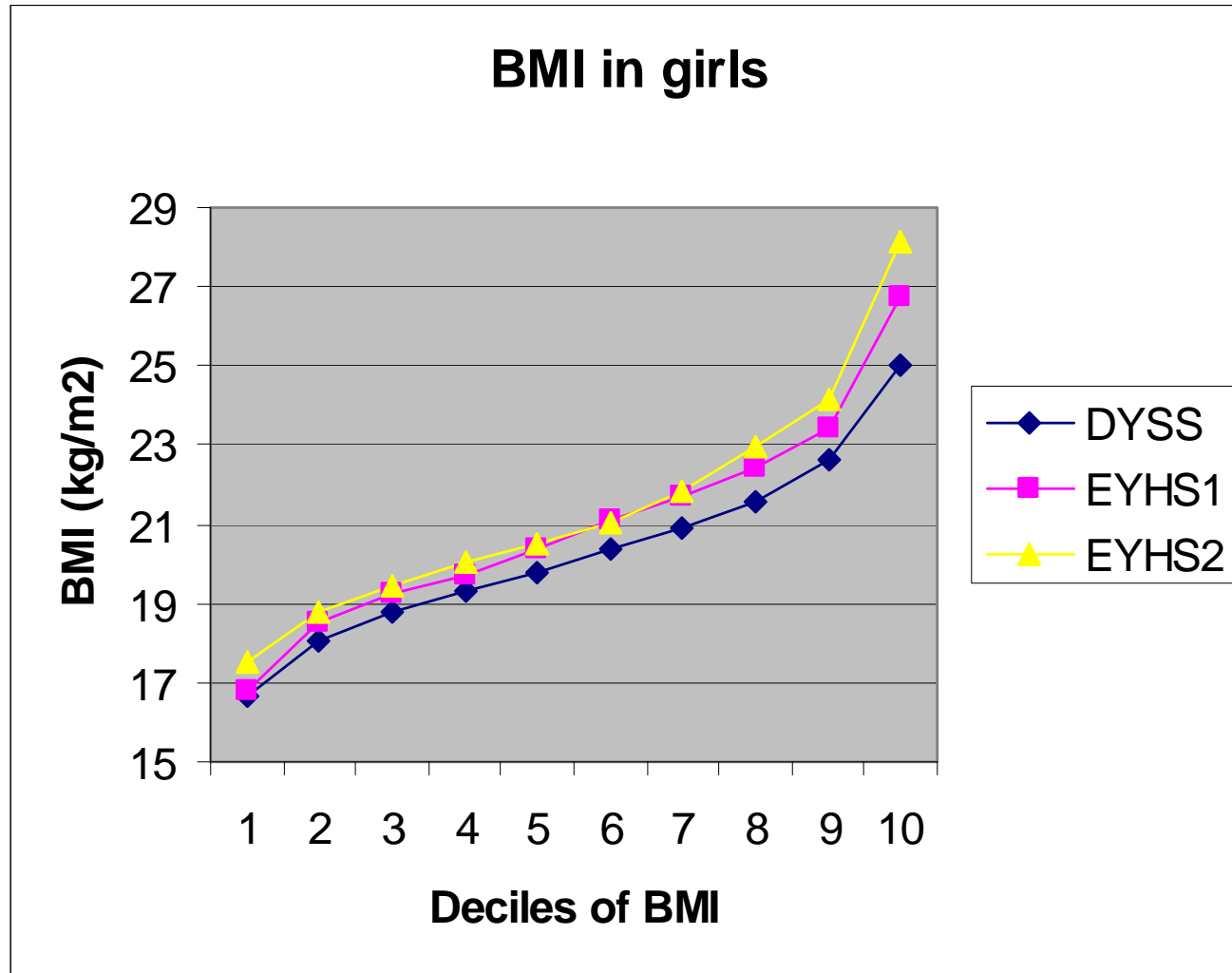
Legend: DYSS, EYHS1, EYHS2

Table 1: Descriptive of the 3 cohorts, mean (SD): DYSS (Danish Youth and Sport Study) and the randomly selected subgroup participating in Lab tests; EYHS1 (European Youth Heart Study 1997); EYHS2 (European Youth Heart Study 2003). Only subjects with valid fitness test are included.

Cohort		N	Age (years)	Height (cm)	Weight (kg)	BMI (kg m^{-2})	HR$_{max}$ (bpm)	Watt$_{max}$	Watts kg^{-1}
DYSS	♀	584	17.1 (1.0)	168.1 (6.6)	58.5 (8.1)	20.7 (2.4)	195.4 (17.3)	193 (38)	3.32 (0.62)
	♂	466	17.1 (1.0)	179.4 (6.6)	66.8 (7.9)	20.7 (2.0)	197.0 (15.0)	269 (47)	4.05 (0.68)
DYSS∞	♀	151	17.1 (1.1)	168.1 (6.9)	58.5 (8.4)	20.7 (2.2)	198.2 (15.5)	199 (41)	3.41 (0.56)
	♂	120	17.2 (1.0)	178.8 (6.6)	66.4 (7.0)	20.8 (1.9)	196.0 (13.5)	271 (49)	4.10 (0.72)
EYHS1	♀	212	15.5 (0.5)	165.1 (6.4)	57.2 (8.8)	20.9 (2.8)	197.0 (7.3)	168 (29)	2.98 (0.51)
	♂	198	15.5 (0.4)	174.2 (6.9)	63.5 (10.1)	20.9 (2.6)	198.5 (6.7)	237 (40)	3.77 (0.55)
EYHS2	♀	234	15.7 (0.4)	164.7 (6.9)	57.6 (9.1)	21.2 (3.0)	197.4 (6.6)	170 (25)	2.98 (0.44)
	♂	184	15.8 (0.3)	176.0 (7.5)	65.3 (10.9)	21.0 (2.8)	197.5 (7.2)	243 (36)	3.76 (0.53)

∞Group with direct measurement of VO$_{2max}$. Data shown are from the measurements done by the teachers, who tested the main group

Table 2: Descriptive (mean (SD)) of Lab test from DYSS randomly selected subgroup.

Cohort		HR_{max} (bpm)	MPO (Watts)	MPO (Watts kg^{-1})	VO_{2max} (l min^{-1})	Fitness (ml kg^{-1} min^{-1})
DYSS∞	Girls (n=151)	199.2 (7.9)	204 (44)	3.44 (0.56)	2.39 (0.44)	40.4 (5.3)
	Boys (n=120)	200.0 (7.9)	284 (42)	4.26 (0.53)	3.46 (0.52)	51.6 (5.6)

∞Group with direct measurement of VO_{2max}.

Table 3. Comparison of the three cohorts where all data are adjusted to 16 years of age, and VO2max was estimated from the validation studies. P-values calculated from oneway anova using Tukey's posthoc analysis.

	Boys				Girls			
	DYSS 1983	EYHS1 1997	EYHS2 2003		DYSS 1983	EYHS1 1997	EYHS2 2003	
	Mean (SD)	Mean (SD)	Mean (SD)	P	Mean (SD)	Mean (SD)	Mean (SD)	P
BMI (kg m^{-2})	20,46 (2.01)	20,99 (2.58)	21,05 (2.81)	#0.01	20,31 (2.36)	21,02 (2.75)	21,45 (3.06)	#0.001
MPO (watts)	262,2 (46.5)	239,9 (39.7)	244,0 (36.5)	#0.001	190,5 (38.2)	169,1 (29.4)	171,0 (25.3)	#0.001
MPO (watts kg^{-1})	4,04 (0.68)	3,77 (0.55)	3,76 (0.53)	#0.001	3,36 (0.62)	2,96 (0.50)	2,97 (0.44)	#0.001
Estimated VO$_{2max}$	51,64 (6.10)	52,15 (6.45)	51,61 (6.68)	NS	41,36 (5.27)	40,99 (5.87)	40,99 (5.36)	NS

NS denotes that no difference was found between any groups; # is DYSS > EYHS1 and EYHS2 (no difference between EYHS1 and EYHS2 in any parameter).