Unraveling the genetic components of voluntary exercise behavior in adolescents and young adults

Schutte, N.M.

2017

document version
Publisher's PDF, also known as Version of record

CHAPTER 6

A TWIN-SIBLING STUDY AND META-ANALYSIS ON THE HERITABILITY OF MAXIMAL OXYGEN CONSUMPTION

PUBLISHED AS

Schutte NM, Nederend I, Hudziak JJ, Bartels M, de Geus EJC (2016)

A twin-sibling study and meta-analysis on the heritability of maximal oxygen consumption

Physiological Genomics 48(3), 210-219

ABSTRACT

Large individual differences exist in aerobic fitness in childhood and adolescence, but the relative contribution of genetic factors to this variation remains to be established. In a sample of adolescent twins and siblings (*N* = 479), heart rate (HR) and maximal oxygen uptake ($\dot{V}O_{2max}$) were recorded during the climax of a graded maximal exercise test. In addition, $\dot{V}O_{2max}$ was predicted in two graded submaximal exercise tests on the cycle ergometer and the treadmill, using extrapolation of the HR/$\dot{V}O_2$ curve to the predicted HR_{max}. Heritability estimates for measured $\dot{V}O_{2max}$ were 60% in mL/min and 55% for $\dot{V}O_{2max}$ in mL/min/kg. Phenotypic correlations between measured $\dot{V}O_{2max}$ and predicted $\dot{V}O_{2max}$ from either submaximal treadmill or cycle ergometer tests were modest ($.57 < r < .70$), in part because of the poor agreement between predicted and actual HR_{max}. The majority of this correlation was explained by genetic factors, therefore the submaximal exercise tests still led to very comparable estimates of heritability of $\dot{V}O_{2max}$. To arrive at a robust estimate for the heritability of $\dot{V}O_{2max}$ in children to young adults, a sample size weighted meta-analysis was performed on all extant twin and sibling studies in this age range. Eight studies, including the current study, were meta-analyzed and resulted in a weighted heritability estimate of 59% (mL/min) and 72% (mL/min/kg) for $\dot{V}O_{2max}$. Taken together, the twin-sibling study and meta-analyses showed that from childhood to early adulthood genetic factors determine more than half of the individual differences in $\dot{V}O_{2max}$.

INTRODUCTION

Maximal oxygen uptake ($\dot{V}O_{2max}$) is defined as the highest rate of oxygen consumption during maximal intensity exercise performed until exhaustion (Kenney et al., 2012) and is considered a good index of aerobic fitness and endurance capacity. Direct measurement of oxygen consumption and carbon dioxide production during the climax of a graded maximal exercise test is the golden standard to measure $\dot{V}O_{2max}$. Large individual differences exist in maximal exercise test derived $\dot{V}O_{2max}$, and, although these are significantly correlated to the regular exercise status of a subject, this correlation is not as strong as generally assumed. Various measures of total physical activity or regular leisure time sports and exercise behavior generally show only modest association with $\dot{V}O_{2max}$ (Aadahl et al., 2007; Bonen & Shaw, 1995; Siconolfi et al., 1985; Talbot et al., 2000). The variation in baseline $\dot{V}O_{2max}$ in sedentary subjects is often already much larger than the training-induced increase over this baseline, which is on average only about 25% (Church et al., 2007; Grant et al., 1995; Payne & Morrow, Jr., 1993; Wilmore et al., 2001). Training furthermore increases rather than decreases the individual differences seen at baseline, as the $\dot{V}O_{2max}$ response to training itself shows large variation (Bouchard & Rankinen, 2001; Skinner et al., 2001).

The above pattern suggests an important role for innate factors in the population variation in $\dot{V}O_{2max}$ and twin and family studies seem to confirm this (Bouchard et al., 1986; Bouchard et al., 1998; Fagard et al., 1991; Klissouras, 1971; Klissouras et al., 1973; Lesage et al., 1985; Lortie et al., 1982; Maes et al., 1996; Montoye & Gayle, 1978; Mustelin et al., 2011; Sundet et al., 1994). Table 6.1 provides an overview of correlations among relatives i.e. monozygotic (identical, MZ) and dizygotic (fraternal, DZ) twins, siblings and parents with their offspring. The monozygotic twin correlations in Table 6.1 range from .62 to .95. The dizygotic (DZ) twin correlations, sibling correlations and parent-offspring correlations also vary substantially across studies, but are systematically lower than the MZ twin correlations. In line with the variability in twin correlations, heritability estimates have varied widely. Possible sources of this variation are differences in the age of the subjects, differential approaches to adjustment for body mass and/or body composition, training status of the subjects, or differences in protocol or fitness equipment (i.e. cycle ergometer or treadmill) that was used to measure or predict $\dot{V}O_{2max}$ between the various studies. A major source, however, seems to be the rather modest sample sizes. As is clear from Table 6.1 there are only two studies with large samples

(Lortie et al., 1982; Sundet et al., 1994) but both these larger studies used a submaximal instead of a maximal exercise test. These tests do not measure $\dot{V}O_{2max}$ directly, but predict it from an exercise test that is halted at a predetermined point (certain percent of the predicted maximal heart rate) below the maximal exercise capability of the individual. Since it does not demand $\dot{V}O_{2max}$ measurement during exhaustive exercise, the submaximal exercise test is better suited in larger (genetic) epidemiological studies. However, it is currently unknown whether a submaximal exercise test correctly captures the genetic factors influencing $\dot{V}O_{2max}$.

One of the most used submaximal exercise test is the nomogram of Åstrand, which requires cycling on a constant individually chosen work rate. $\dot{V}O_{2max}$ is predicted using the steady-state heart rate (HR) achieved after 6 minutes (Åstrand & Rhyming, 1960). This method has clear limitations as results may be influenced by individual differences in submaximal HR at a given work rate due to training status, resting HR and body composition. Estimated $\dot{V}O_{2max}$ with this method showed correlations in the range of .47 and .82 with measured $\dot{V}O_{2max}$ in adult populations (Cink & Thomas, 1981; Ekblom-Bak et al., 2014; Jette, 1979; Kasch, 1984; Siconolfi et al., 1982). More promising is the $\dot{V}O_{2max}$ prediction using a graded submaximal exercise protocol in which the intensity increases at regular intervals up to but never exceeding a certain percent of the maximal heart rate (HR_{max}). $\dot{V}O_{2max}$ can be obtained by extrapolating the HR/$\dot{V}O_2$ curve to the predicted HR_{max}, allowing for individual differences in $\dot{V}O_2$/HR slope. This estimation method showed correlations in the range of .76 and .98 with measured $\dot{V}O_{2max}$ in adult populations (Ekblom-Bak et al., 2014; Grant et al., 1995; Legge & Banister, 1986), although it is sensitive to the protocol used. Submaximal tests on a cycle ergometer yield lower predicted $\dot{V}O_{2max}$ values than tests on a treadmill (Grant et al., 1995; Mays et al., 2010).

Adolescent $\dot{V}O_{2max}$ has been measured in parent-offspring studies using submaximal exercise tests (Lesage et al., 1985; Lortie et al., 1982) but a striking omission in Table 6.1 is adolescent twin studies using a maximal exercise test to examine $\dot{V}O_{2max}$ in an adolescent population. The aim of the current study is to address this gap in the extant literature. In a large sample of adolescent twins and siblings, HR and $\dot{V}O_2$ were recorded during the climax of a graded maximal exercise test. $\dot{V}O_{2max}$ was further predicted from two graded submaximal exercise tests on the cycle ergometer and the treadmill, using extrapolation of the HR/ $\dot{V}O_2$ curve to the predicted HR_{max}. This allowed us to address our second aim: to test the extent to which

Table 6.1 Overview of genetic studies on V̇O2max conducted in a twin and/or family design.

Study	Subjects	$\dot{V}O_{2max}$ measurements	r_{MZ}	r_{DZ}	$r_{sibling}$	$r_{parent-offspring}$	Heritability
Klissouras et al., 1971	• 15 MZ • 10 DZ • Age: 10 ± 2	• Maximal exercise test • Treadmill • mL/min	.91	.44			93%
Klissouras et al., 1973	• 23 MZ • 16 DZ • Age: 27 ± 14	• Maximal exercise test • Treadmill • mL/min/kg	.95	.36			
Montoye & Gayle 1978	• 93 father-son pairs • 70 brother pairs • Age: 10 – 69	• <39y Maximal exercise test • >39y Submaximal exercise test (until HR=160) • Treadmill • L/min • Corrected for age, weight, skinfolds			.18	.34	
Lortie et al., 1982	• 594 parent-offspring pairs • 223 sibling pairs • Age: 42 ± 4 (parents) • Age: 15 ± 3 (children)	• Submaximal exercise test (until HR=150/170) • Treadmill • mL/min • mL/min/kg • VO_{2max} predicted • Corrected for sex, age, skinfolds, physical activity, cigarette smoking and SES			.36[a] .33[b]	.21[a] .17[b]	
Lesage et al., 1985	• 96 parent-offspring pairs • 39 sibling pairs • Age: 43 ± 5 (parents) • Age: 16 ± 4 (children)	• Maximal exercise test • Treadmill • mL/min • mL/min/kg • Corrected for sex and age			.14[a] .19[b]	.06[a] .03[b]	
Bouchard et al., 1986	• 106 MZ • 66 DZ • 27 sibling pairs • Age: 22 ± 3	• Maximal exercise test • Cycle ergometer • mL/min/kg • Corrected for sex and age	.70	.51	.41		38%

Study	Subjects	$\dot{V}O_{2max}$ measurements	r_{MZ}	r_{DZ}	$r_{sibling}$	$r_{parent\text{-}offspring}$	Heritability
Fagard et al., 1991	• 29 MZ • 19 DZ • Age: 22 ± 4	• Maximal exercise test • Cycle ergometer • mL/min • mL/min/kg • Only males, restricted age range	.77[a] .77[b]	.05[a] .04[b]			77% 68%
Sundet et al., 1994	• 436 MZ • 622 DZ • Age: late teens/ early twenties	• Submaximal exercise test (until HR=140) • Cycle ergometer • mL/min/kg • VO_{2max} predicted[c] • Only males, restricted age range	.62	.29			62%
Maes et al., 1996	• 43 MZ • 61 DZ • 84 fathers • 97 mothers • Age: 39 ± 4 (parents) • Age: 10 (children)	• Maximal exercise test • Treadmill • L/min • Restricted age range	.75	.32		.25 / .31[d]	69% / 87%[e]
Bouchard et al., 1998	• 125 sons • 134 daughters • 85 fathers • 85 mothers • Age: 52 ± 5 (parents) • Age: 25 ± 6 (children)	• Maximal exercise test • Cycle ergometer • mL/min • Corrected for sex and age			.36	.14 / .36[d]	59%
Mustelin et al., 2011	• 59 MZ • 92 DZ • Age: 27 ± 2	• Maximal exercise test • Cycle ergometer • mL/min • Corrected for sex, restricted age range	.64	.21			65%

Note. r_{MZ} = Monozygotic twin correlation; r_{DZ} = Dizygotic twin correlation; $r_{sibling}$ = Sibling correlation; $r_{parent\text{-}offspring}$ = Parent-offspring correlation; [a] mL/min; [b] mL/min/kg; [c] Predicted VO_{2max} was transformed to a categorical score from 1 to 9. The correlations are based upon those categorical scores; [d] father-child correlation/mother-child correlation; [e] heritability estimate for males/heritability estimate for females.

the genetic factors influencing measured $\dot{V}O_{2max}$ during a maximal exercise test overlap with those influencing predicted $\dot{V}O_{2max}$ from submaximal exercise tests. Information on the genetic overlap between measured and predicted $\dot{V}O_{2max}$ can reveal whether they can be used interchangeably in genetic association studies to examine the association of aiming to identify the genetic variants underlying $\dot{V}O_{2max}$. A high degree of overlap would mean that submaximal exercise tests, which are easier to implement in large scale genetic studies, might suffice for such studies. Twin correlations, heritability of the measured and predicted $\dot{V}O_{2max}$ as well as the genetic covariance among these parameters were estimated in a multivariate design. We hypothesize that a substantial part of the variation in $\dot{V}O_{2max}$ in our adolescent sample is explained by genetic factors. As previous studies in adults showed high correlations between $\dot{V}O_{2max}$ predicted using a graded submaximal exercise protocol and measured $\dot{V}O_{2max}$, we expect moderate to high phenotypic correlations, and a significant contribution of genes to this correlation. Finally, a sample size weighted meta-analysis was performed on the univariate analysis obtained from all twin studies in the age range of 10 to 30 years (including the current study) that measured $\dot{V}O_{2max}$, aiming to arrive at a more robust estimate for the heritability of this crucial trait in exercise physiology.

METHODS

Sample

Healthy adolescent twin pairs aged between 16 and 18 and their siblings (age range 12 – 25) from the Netherlands Twin Register (van Beijsterveldt et al., 2013) were invited to participate in a study on the determinants of adolescent exercise behavior. Selection for invitation was based on the availability of longitudinal survey data on zygosity and regular leisure time exercise behavior. The aim was to have sufficient individuals present from the entire spectrum of sedentary to vigorous leisure time exerciser and for each zygosity group. We started with a random selection, but if a zygosity group was underrepresented or if there were too few sedentary or vigorous exercisers, invitations were biased towards the underrepresented groups. This was mainly the case for sedentary subjects; twins who reported no engagement in exercise behavior on a previously filled out survey were selected for invitation. The co-twin was then selected as well, regardless of her or his exercise status.

In order to be eligible for the study, subjects had to have no history of cardiovascular or respiratory disease, and being physically capable of engaging in exercise activities.

Participants were invited by sending a letter advertising the opportunity to test their fitness in addition to earning a gift voucher. All invitees had to be able and willing to visit the VU University in Amsterdam for lab testing. For the current study, a complete dataset was available for 479 subjects: 221 complete twin pairs: 112 monozygotic pairs (MZ) and 109 dizygotic pairs (DZ) and 33 of their singleton siblings. In addition, two non-twin sibling pairs participated. This sample size should be sufficient to detect univariate genetic influences with a power of 80% (assuming substantial heritability estimates of 60%, based on previous studies) (Posthuma & Boomsma, 2000).

All subjects provided written informed consent and if the subjects were under 18 consent was given by both of their parents/guardians. All study procedures submitted to and approved by the Medical Ethics Review Committee of the VU University Medical Center Amsterdam (NL35634.029.10).

Procedure

On arrival at the laboratory, height and weight were measured and a short lifestyle interview was completed, including detailed questions on current levels of regular exercise. Next, two exercise test were conducted (in fixed order) on an electromechanically braked Lode cycle ergometer (type Corival) and a Lode treadmill (type Valiant) at fixed loads that are below the intensity of the ventilatory threshold for most adolescents.

The first session on the cycle ergometer started with a 2-minute warming up period, followed by 4 incremental stages of 5 minutes each (males: 70W, 90W, 110W, 130W; females: 40W, 60W, 80W, 100W). Subjects were instructed to pedal at fixed rounds per minute (RPM): between 60 and 70 RPM. The test ended with a 1-minute cooling-down phase, followed by a 5-minute recovery period. The second session on the treadmill consisted of a 1-minute warm-up period, followed by 4 incremental stages of 5 minutes each (males: 6, 6.5, 7 and 8 km/h; females: 5.5, 6, 6.5 and 7 km/h). Again, the test ended with a 1-minute cooling-down phase, followed by a 5-minute recovery period. To ensure that the intensity of every stage was below the intensity of the ventilatory threshold for most adolescents, the ratio of the oxygen

consumption and carbon dioxide production ($\dot{V}CO_2/\dot{V}O_2$) was monitored. This respiratory exchange ratio (RER) can be used to estimate the ventilatory threshold (Solberg et al., 2005). This threshold is passes when exhalation of CO_2 exceeds inhalation of O_2, which is visualized by a RER > 1.00. For each test the load of each stage was adjusted when necessary to keep the intensity below an RER of 0.95.

Finally, an incremental maximal exercise test was conducted on a cycle ergometer to establish $\dot{V}O_{2max.}$ The work rate was increased every minute until exhaustion while subjects pedaled at 60-100 RPM. In the standard protocol male started at 75 Watt with increments of 25 Watt per minute. For females stage one started at 70 Watt and work load was increased by 20 Watt per minute. Adjustments to this protocol (higher increasing workloads every step) were done by experienced researchers based on the exercise behavior, age, height and weight of the subject. The test was terminated when the subject was not able to keep RPM above 50 despite serious attempts. After cessation of the test, every subject completed a mandatory cool-down phase on the cycle ergometer of 5 minutes on a low, individually chosen work rate.

Measurements

Regular exercise behavior Leisure time exercise behavior was measured by a short lifestyle interview, in which the subjects indicated what types of regular sports or exercise activities they were involved in. Subjects were asked to indicate for each activity for how many years the subject participated in the activity, for how many months a year, how many times a week, and how many minutes each time. Each activity was recoded into a metabolic equivalent (MET) score, based on the compendium of energy expenditure (Ainsworth et al., 1993). By multiplying the MET score, the frequency, and the duration of each exercise activity, weekly MET-hours spent on exercise activities were calculated for each subject. We only included activities that were conducted for at least 3 months a year and since at least half a year (thereby excluding ski holidays, sailing camps, and similar). In addition, subjects were asked to indicate how much time per week was spent on physical activity related to active transportation (walking, cycling) and compulsory physical education classes, but MET-hours spent on these activities were kept separate and not used in our index of voluntary exercise behaviour in leisure time.

Gas exchange Oxygen uptake ($\dot{V}O_2$) and carbon dioxide production ($\dot{V}CO_2$) were recorded breath-by-breath by means of a telemetric gas exchange system (Cosmed K4b^2, Cosmed Benelux, Nieuwegein, The Netherlands). During the course of the experiment, the main sample unit and the battery pack were attached to the back of the subject. Before each test, the O_2/CO_2 analysis system was calibrated using ambient air and a gas mixture that had an O_2 concentration of 16% and a CO_2 concentration of 5%. The calibration of the turbine flowmeter was performed by using a 3 liter syringe (all according to the manufacturer's instructions). Figure 6.1 illustrates the changes in $\dot{V}CO_2$ and $\dot{V}O_2$ across the entire experimental protocol for a pair of MZ and a pair of DZ twins.

Heart rate The electrocardiogram (ECG) was recorded continuously with the VU-AMS5fs device (VU University, Amsterdam, The Netherlands). This device was developed to study autonomic nervous system activity in naturalistic settings (de Geus et al., 1995). The version used here measured the ECG together with the impedance cardiogram (ICG) from five disposable, pre-gelled Ag/AgCl electrodes. Due to the portable nature of this device, the subjects were not discomforted by wearing this on the hip during the exercise tests. Heart rate was obtained from the ECG by an automated R-wave peak detector in the VU-AMS software suite (VU-DAMS version 3.1, VU University, Amsterdam, the Netherlands, www.vu-ams.nl) and shown online during testing. Data analysis was based on automated offline scoring of the R-waves, with suspicious inter beat intervals (too short or too long taken the local mean and variance) corrected by interpolation or excluded by marking these beats as artifacts during visual inspection of the ECG signal.

Data processing

Measuring $\dot{V}O_{2max}$ during maximal exercise To obtain VO_{2max}, only $\dot{V}O_2$ data with a corresponding RER of at least 1.10 was selected to ensure good effort above the intensity of the ventilatory threshold. Breath-by-breath VO_2 data was cut into 20-second blocks. For every 20 second block, the mean $\dot{V}O_2$ was calculated, after discarding deviant breaths. $\dot{V}O_{2max}$ was determined as the highest mean value of $\dot{V}O_2$ of all the 20-second blocks. The maximal HR in that specific block was taken as corresponding HR_{max}.

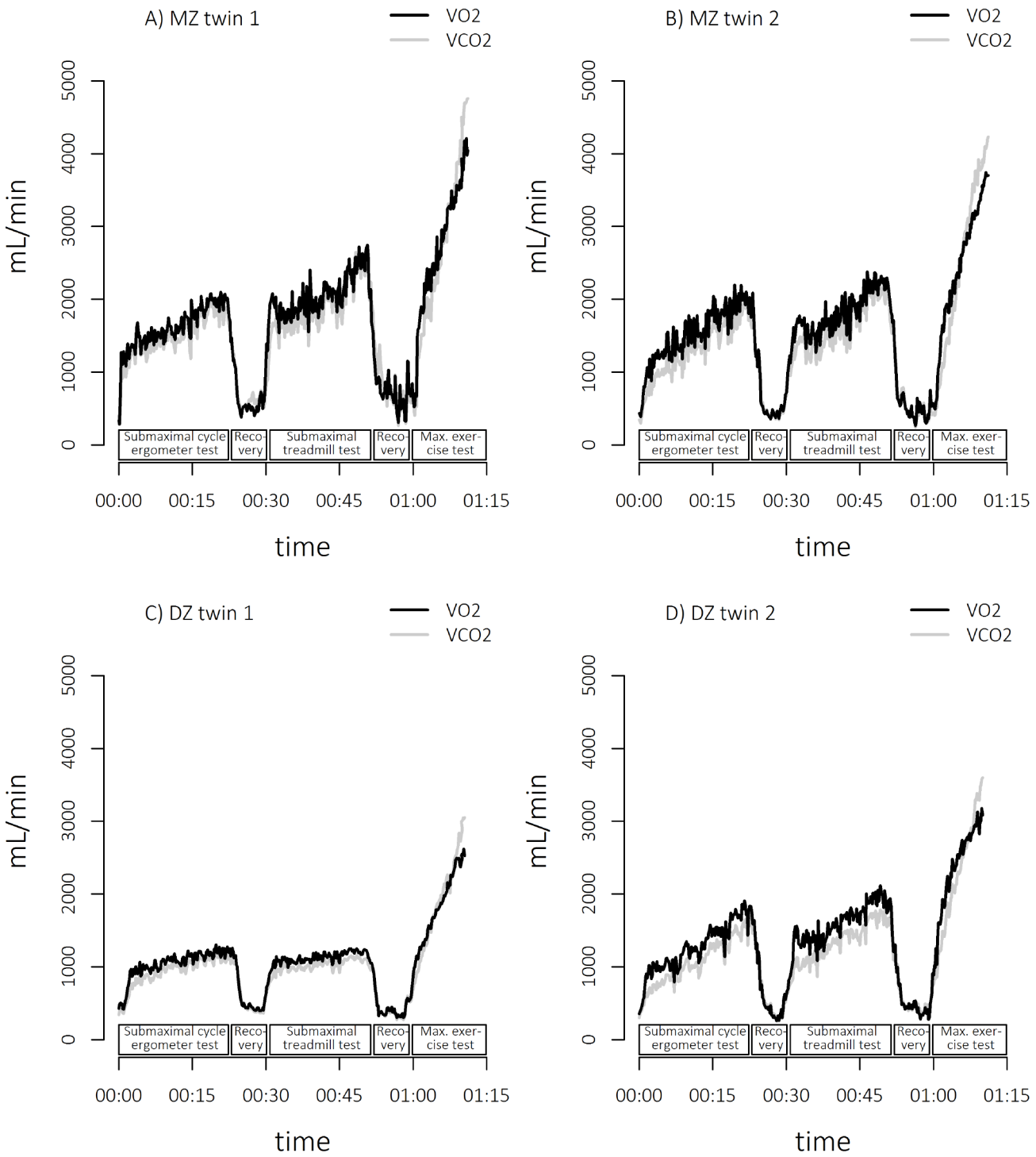

Figure 6.1 Changes in $\dot{V}O_2$ and $\dot{V}CO_2$ across the entire experimental protocol for a pair of MZ (A and B) and a pair of DZ twins (C and D). The two submaximal exercise tests on the cycle ergometer and treadmill and the final maximal exercise test are clearly visible as $\dot{V}O_2$ and $\dot{V}CO_2$ increase when subjects start exercising. The MZ twins resemble each other more than DZ twins in absolute $\dot{V}O_2$ and $\dot{V}CO_2$.

Predicting $\dot{V}O_{2max}$ from submaximal exercise To predict $\dot{V}O_{2max}$, breath-by-breath $\dot{V}O_2$ data and beat-to-beat HR data were synchronized and the mean of every 5-second block was calculated for submaximal cycle and treadmill exercise tests separately. Using the univariate regression function in SPSS (IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp), the relationship between $\dot{V}O_2$ (dependent variable) and HR (independent variable) was examined and a slope and intercept were calculated for every subject for the submaximal cycle ergometer test as well as for the submaximal treadmill test. Using these parameter estimates together with HR_{max} the predicted $\dot{V}O_{2max}$ was calculated for every subject. Because we wanted to test the feasibility of using submaximal tests only, HR_{max} was obtained by using the formula 208 – 0.7*age (Tanaka et al., 2001) rather than using the actual measured HR_{max}, although analyses were repeated using the actual measured HR_{max}.

Genetic analyses

Genetic structural equation modeling was done in OpenMx (Boker et al., 2011) under R (R Development Core Team, 2011) with the raw-data ML procedure for estimation of parameters. For all analyses, a threshold of $p < 0.05$ was considered for statistical significance. All $\dot{V}O_{2max}$ values were Z-transformed. Since (non-twin) siblings share, like DZ twins, on average 50% of their genes, parameter estimates were constrained to be equal for DZ twins and siblings. First, a trivariate model that estimated all parameters freely (a saturated model) was fitted, including the measured $\dot{V}O_{2max}$ and the $\dot{V}O_{2max}$ predicted from the submaximal cycle and treadmill test. Main effects of sex and age on mean levels of these phenotypes were considered in the model. The significance of these covariates was tested by comparing the model including the specific component to a model in which the component is constraint to be equal to zero. These nested submodels were compared by hierarchic χ^2 tests. The χ^2statistic is computed by subtracting log-likelihood (–2LL) for a reduced model from the -2LL for the full model (χ^2= $-2LL_{full\ model} - -2LL_{reduced\ model}$). This χ^2statistic is distributed with degrees of freedom (*df*) equal to the difference in the number of parameters estimated in the two models ($\Delta df = df_{full\ model} - df_{reduced\ model}$). If the difference test is significant the constraints on the reduced model cause a significant deterioration of the fit of model. Twin and cross-twin/cross-trait correlations and their 95% confidence intervals were estimated for the MZ and DZ twins/siblings.

Subsequently, a trivariate Cholesky decomposition was fitted to the data. This decomposition model decomposes the total phenotypic variance into sources of additive genetic variance/covariance (A), dominant genetic variance/covariance (D) or shared (familial) environmental variance/covariance (C) and person-specific environmental variance/covariance (E). The C and D effects cannot be estimated simultaneously in a twin/sibling model. Therefore, the ratio of the MZ correlations to the DZ/sibling correlations was used to determine which model (ACE or ADE) is most appropriate. The significance of the variance-covariance components was tested by comparing the model including the specific component to a model in which the component is constraint to be equal to zero.

Meta-analysis

A search of the electronic databases ISI Web of Knowledge and PubMed was conducted using the key words: *maximal oxygen uptake, $\dot{V}O_{2max}$, aerobic capacity, aerobic performance, cardiorespiratory (fitness) and genes, heritability, twin(s), family* (date last searched: January 2015). Furthermore, the reference lists of these articles were inspected. Articles published in English, and reporting twin, sibling and/or parent-offspring correlations and corresponding sample sizes, and with subjects with an age < 30y were selected. Only articles in which $\dot{V}O_2$ was measured in a maximal exercise protocol or predicted using a submaximal exercise protocol were included. All twin and sibling correlations of these articles (including the current study) were included in a sample size weighted meta-analysis for $\dot{V}O_{2max}$ expressed in mL/min and $\dot{V}O_{2max}$ expressed in mL/min/kg. Twin correlations from the current study were calculated in univariate models without the siblings to be comparable to the twin correlations included in the meta-analysis.

In OpenMx, a variance decomposition model was fitted to the twin correlations (weighted for sample size) to estimate the influence of additive genetic (A) and shared environmental influences (C) on $\dot{V}O_{2max}$ in mL/min and $\dot{V}O_{2max}$ in mL/min/kg according to the approach of Bartels et al. (2003). First, the twin and sibling correlations were used to estimate the genetic and environmental influences for each study separately. Subsequently, all studies were taken together to estimate one weighted heritability estimate for $\dot{V}O_{2max}$. These two models were compared using the hierarchic χ^2 test. A significant deterioration of the fit of model indicated significant heterogeneity across the studies (Bartels et al., 2003). We repeated the meta-

analysis by excluding the study by Sundet et al. (1994) that used predicted $\dot{V}O_{2max}$ from submaximal exercise testing to also provide a weighted heritability estimate of actual measured $\dot{V}O_{2max}$.

RESULTS

General descriptives

Means and standard deviations for measured $\dot{V}O_{2max}$ and $\dot{V}O_{2max}$ predicted from the submaximal cycle and treadmill test, and measured and predicted HR_{max} of males and females are shown in Table 6.2. Fifteen subjects did not meet the RER > 1.10 criterion. For 9 of these subjects there was no sufficient evidence that they did exercise until exhaustion according to the experimental researcher report and/or the HR_{max} was less than 85% of HR_{max}. Therefore, these 9 subjects and their coinciding twin/sibling were excluded from further analyses involving measured $\dot{V}O_{2max}$. The final sample size consisted of 463 subjects. Means and standard deviations of minutes spent on walking, cycling, physical education class and MET scores for leisure time exercise behavior are presented in Table 6.2. Although the age-range was small, significant age effects ($\dot{V}O_{2max}$ increases with age) were found on $\dot{V}O_{2max}$ in mL/min ($p < .001$), but not for $\dot{V}O_{2max}$ expressed in mL/min/kg. Both predicted and measured $\dot{V}O_{2max}$ were higher in males than in females (all $p < .001$). Furthermore, males were more engaged in weekly exercise behavior ($p = .002$). As expected, weekly leisure time exercise behavior correlated significantly with $\dot{V}O_{2max}$, but the correlation was modest: $r = .28$ with $\dot{V}O_{2max}$ in mL/min and $r = .34$ (both $p < .001$) with $\dot{V}O_{2max}$ mL/min/kg. The correlation between measured $\dot{V}O_{2max}$ and weekly minutes of cycling was significant ($r = .25$ for $\dot{V}O_{2max}$ expressed in mL/min and $r = .22$ for $\dot{V}O_{2max}$ expressed in mL/min/kg, both $p < .001$), whereas the correlations between measured $\dot{V}O_{2max}$ and weekly minutes of walking or weekly hours of physical education class were only small ($-.12 < r < .03$).

Table 6.2 Means and standard deviations (SD) of measured and predicted $\dot{V}O_{2maxl}$ in mL/min and mL/min/kg, measured and predicted HR, and minutes per week spent on walking and cycling (transportation), physical education class and leisure time exercise behavior in METs of males and females.

		Males (N = 233)		Females (N = 230)	
		Mean	*SD*	*Mean*	*SD*
Body composition	Height (cm)	180.4	7.8	168.3	6.6
	Weight (cm)	67.3	10.3	61.8	9.7
	BMI ($kg \cdot m^{-1}$)	20.6	2.5	21.8	3.3
$\dot{V}O_{2max}$ in mL/min	Measured	3132	540	2240	316
	Predicted from cycle ergometer test	2933	648	2021	389
	Predicted from treadmill test	2968	606	2029	400
$\dot{V}O_{2max}$ in mL/min/kg	Measured	46.9	6.9	36.7	5.6
	Predicted from cycle ergometer test	43.8	8.1	33.0	6.0
	Predicted from treadmill test	44.5	8.1	34.1	6.3
Heart rate (bpm)	Resting Heart Rate	72.6	11.4	75.6	11.2
	Maximal Heart Rate Measured	195.4	10.1	195.2	8.9
	Maximal Heart Rate Predicted (Tanaka)	196.1	0.8	195.9	0.9
Regular exercise	Walking (minutes/week)	38.9	71.8	43.5	79.8
	Cycling (minutes/week)	233.1	156.3	209.4	163.3
	Physical education (minutes/week)	151.6	139.9	132.4	112.4
	Leisure-time exercise (METs/week)	25.7	22.5	19.2	22.1

Correlation between measured and predicted $\dot{V}O_{2max}$

Measured $\dot{V}O_{2max}$ in mL/min showed a correlation of .70 (95% CI: .65 – .75) with $\dot{V}O_{2max}$ predicted from the submaximal cycle test and .64 (95% CI: .58 – .70) with $\dot{V}O_{2max}$ predicted from the treadmill test. Likewise, measured $\dot{V}O_{2max}$ in ml/min/kg was significantly correlated with $\dot{V}O_{2max}$ predicted from the submaximal cycle test (r = .61, 95% CI: .55 – .68) and with $\dot{V}O_{2max}$ predicted from the submaximal treadmill test (r = .57, 95% CI: .50 – .64). In spite of the significant relationship between predicted and measured $\dot{V}O_{2max}$, Bland Altman plots in Figure 6.2 show considerably discrepancy between these measures, expressed in mL/min. Regression of the mean of the two measurements (measured and predicted $\dot{V}O_{2max}$) on the difference between the two values (y-axis), showed that the discrepancy increases as

absolute $\dot{V}O_{2max}$ increases. In males the absolute differences in measured and predicted $\dot{V}O_{2max}$ were larger than in females. A potential source of error was the use of an age-predicted HR_{max}. Absolute mean differences between measured (195 ± 10) and predicted HR_{max} (202 ± 1) were greater than zero ($p < .001$). Repeating the analyses with measured HR_{max} significantly improved the correlation of measured to predicted $\dot{V}O_{2max}$ from the submaximal cycle test (in mL/min $r = .76$, 95% CI: .72 – .90; in mL/min/kg $r = .69$, 95% CI: .64 – .74) and to predicted $\dot{V}O_{2max}$ from the treadmill test (in mL/min $r = .71$, 95% CI: .66 – .76; in mL/min/kg $r = .65$, 95% CI: .59 – .70).

Genetic analyses

The twin and cross-twin/cross-trait correlations of measured $\dot{V}O_{2max}$ and predicted $\dot{V}O_{2max}$ are presented in Table 6.3. For $\dot{V}O_{2max}$ in mL/min, MZ correlations ($r = .61$ for measured $\dot{V}O_{2max}$ and $r = .67$ and $r = .65$ for the $\dot{V}O_{2max}$ predicted from the submaximal cycle and treadmill tests) almost twice as high as the DZ/sibling correlation ($r = .26$, $r = .45$ and $r = .37$). When the MZ resemblance is higher than the DZ resemblance this constitutes evidence for genetic influences on $\dot{V}O_{2max}$. For $\dot{V}O_{2max}$ in mL/min/kg, twin correlations were also higher for MZ twins ($r = .53$, $r = .58$ and $r = .59$) than for DZ twins/siblings ($r = .43$, $r = .52$ and $r = .38$) but much less than half, providing evidence for genetic as well as shared environmental factors underlying familial aggregation. The cross-twin/cross-trait correlations (off-diagonal correlations in Table 6.3) were higher for MZ twins than for DZ twins/siblings for all phenotypes suggesting genetic influences on the covariance between measured $\dot{V}O_{2max}$ and $\dot{V}O_{2max}$ predicted from the submaximal cycle and treadmill tests.

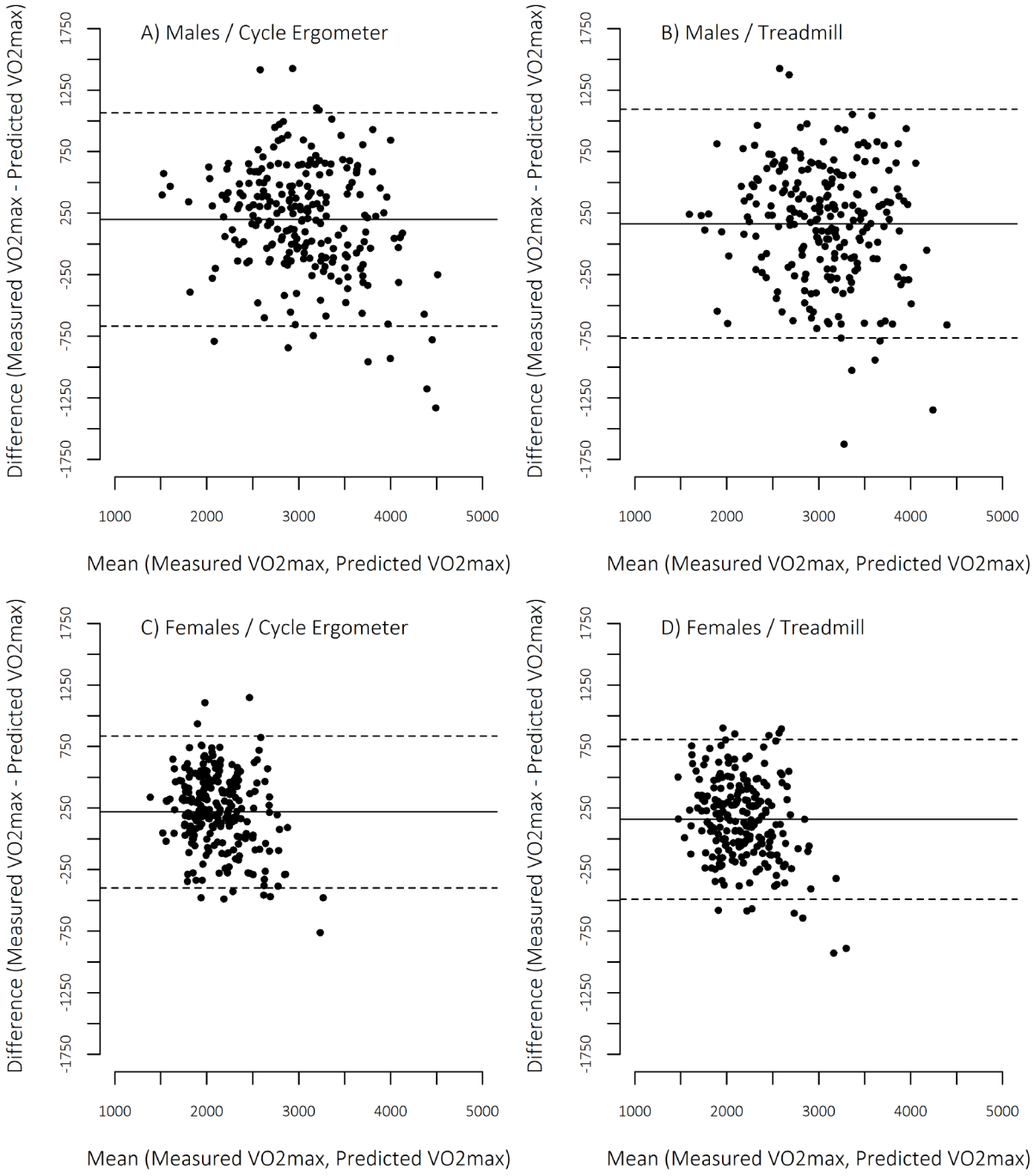

Figure 6.2 Bland-Altman plots for VO_{2max} in mL/min. The x-axis shows the mean of the two measurements (measured and predicted $\dot{V}O_{2max}$) and the y-axis the difference between the two values. The solid line represents the mean difference. The dotted lines represent the average difference ± 1.96 standard deviation of the difference. A) Male $\dot{V}O_{2max}$ submaximal cycle test; B) Female $\dot{V}O_{2max}$ submaximal cycle test; C) Male $\dot{V}O_{2max}$ submaximal treadmill test; D) Female $\dot{V}O_{2max}$ submaximal treadmill test.

Table 6.3 Twin (diagonal) and cross-twin/cross-trait (off diagonal) correlations (95% CI) estimated from the saturated model for measured $\dot{V}O_{2max}$ and $\dot{V}O_{2max}$ predicted from the submaximal cycle and treadmill tests.

	V̇O2max (mL/min)		
	Measured	Predicted cycle test	Predicted treadmill test
	MZ correlations		
Measured	.61 (.50, .70)		
Predicted cycle test	.59 (.51, .66)	.67 (.57, .75)	
Predicted treadmill test	.53 (.44, .61)	.69 (.62, .74)	.65 (.54, .73)
	DZ/sibling correlations		
Measured	.26 (.11, .40)		
Predicted cycle test	.45 (.37, .53)	.45 (.34, .55)	
Predicted treadmill test	.43 (.35, .51)	.53 (.45, .60)	.37 (.23, .49)
	V̇O2max (mL/min/kg)		
	Measured	Predicted cycle test	Predicted treadmill test
	MZ correlations		
Measured	.53 (.40, .63)		
Predicted cycle test	.52 (.43, .59)	.58 (.47, .68)	
Predicted treadmill test	.44 (.35, .53)	.63 (.56, .69)	.59 (.46, .68)
	DZ/sibling correlations		
Measured	.43 (.29, .55)		
Predicted cycle test	.45 (.36, .54)	.52 (.41, .61)	
Predicted treadmill test	.42 (.32, .50)	.53 (.45, .61)	.38 (.24, .50)

Genetic modeling started with an ACE model, as in all cases the DZ/sibling correlation was higher than half the MZ correlation, except for measured $\dot{V}O_{2max}$ in mL/min. Shared environmental influences were not significant for measured and predicted $\dot{V}O_{2max}$ in mL/min ($X^2(6) = 10.8$, $p = .096$). For $\dot{V}O_{2max}$ in mL/min/kg, shared environmental factors were not significant for the measured $\dot{V}O_{2max}$, but for predicted $\dot{V}O_{2max}$ a small but significant effect of shared environmental factors was detected. Standardized components from the best fitting model for additive genetic and shared and person-specific environmental influences on measured and predicted $\dot{V}O_{2max}$ and their covariances are presented in Table 6.4. Heritability estimates for measured $\dot{V}O_{2max}$ were 60% (95% CI: 47% – 69%) and 55% (95% CI: 43% – 64%) for $\dot{V}O_{2max}$ in mL/min (Table 6.4a) and mL/min/kg respectively (Table 6.4b). Heritability estimates for predicted $\dot{V}O_{2max}$ ranged from 47% for $\dot{V}O_{2max}$ in mL/min/kg to 67% for $\dot{V}O_{2max}$ in mL/min (both predicted from the cycle test). Shared environmental influences were small and

not significant for $\dot{V}O_{2max}$ in mL/min. For $\dot{V}O_{2max}$ in mL/min/kg, however, 12% (95% CI: 4% – 19%) of the variance in $\dot{V}O_{2max}$ predicted from the cycle protocol and 4% (95% CI: 4% – 19%) of the variance in $\dot{V}O_{2max}$ predicted from the treadmill protocol could be explained by shared environmental influences.

Table 6.4 Standardized estimates (95% CI) for additive genetic (A), shared environmental (C) and person-specific environmental (E) influences on measured $\dot{V}O_{2max}$ and $\dot{V}O_{2max}$ predicted from the submaximal cycle and treadmill tests and their covariances in (a) mL/min and (b) mL/min/kg.

a)	V̇O2max (mL/min)		
	Measured	Predicted cycle test	Predicted treadmill test
	Additive genetics (A)		
Measured	**.60 (.47, .69)**		
Predicted cycle test	.76 (.63, .85)	**.67 (.60, .75)**	
Predicted treadmill test	.70 (.56, .81)	.76 (.65, .84)	**.64 (.53, .72)**
	Unique environment (E)		
Measured	.40 (.31, .55)		
Predicted cycle test	.24 (.15, .37)	.33 (.25, .43)	
Predicted treadmill test	.30 (.19, .44)	.24 (.16, .35)	.36 (.28, .47)

b)	V̇O2max (mL/min/kg)		
	Measured	Predicted cycle test	Predicted treadmill test
	Additive genetics (A)		
Measured	**.55 (.43, .64)**		
Predicted cycle test	.70 (.55, .82)	**.47 (.32, .60)**	
Predicted treadmill test	.62 (.46, .75)	.61 (.45, .75)	**.55 (.42, .66)**
	Shared environment (C)		
Measured	-		
Predicted cycle test	-	.12 (.04, .19)	
Predicted treadmill test	-	.09 (.01, .16)	.04 (.00, .10)
	Unique environment (E)		
Measured	.44 (.35, .56)		
Predicted cycle test	.31 (.19, .46)	.41 (.32, .52)	
Predicted treadmill test	.37 (.24, .54)	.30 (.21, .43)	.40 (.31, .52)

Note. Heritability estimates in **bold**.

Significant genetic correlations were found for measured $\dot{V}O_{2max}$ and $\dot{V}O_{2max}$ predicted from the submaximal cycle test (r = .84 (95% CI: .76 – .91) for $\dot{V}O_{2max}$ in mL/min and .81 (95% CI: .68 – .95) for $\dot{V}O_{2max}$ in mL/min/kg). Measured $\dot{V}O_{2max}$ and $\dot{V}O_{2max}$ predicted from the submaximal treadmill test showed a genetic correlation of .73 (95% CI: .62 – .82) and .63 (95% CI: .48 – .76) for $\dot{V}O_{2max}$ in mL/min and in mL/min/kg respectively. A genetic correlation > 0 indicates that traits are influences by common genes. Therefore, these correlations suggest that the three $\dot{V}O_{2max}$ measures largely reflect the same set of underlying genetic variants. Furthermore, 61% – 76% of the phenotypic correlations between measured $\dot{V}O_{2max}$ and $\dot{V}O_{2max}$ predicted from the submaximal cycle and treadmill tests could be explained by genetic factors.

Meta-analysis

The literature search and screening resulted in 11 articles (see Table 6.1). Four studies were excluded from the meta-analysis. The studies by Montoye & Gayle (1978), Lortie et al. (1982), Lesage et al. (1985), and Bouchard et al. (1998) were parent-offspring studies and were excluded from the analysis because cohort effects and shared environment could be affecting the correlations. Moreover, whereas other studies either corrected for sex and age or used single-sex or age-restricted samples, Montoye & Gayle (1978) and Lortie et al. (1982) additionally corrected for skinfold thickness, physical activity, cigarette smoking and social-economic status (SES).

The seven studies included in the meta-analysis showed MZ twin correlations for $\dot{V}O_{2max}$ ranging from .62 to .95 whereas the DZ and sibling correlations were much lower (.04 to .51). In the study by Sundet et al. (1994), $\dot{V}O_{2max}$ was predicted using extrapolation of the $\dot{V}O_2$/HR slope, whereas the rest of the studies reported measured $\dot{V}O_{2max}$ values. All of these remaining studies corrected the $\dot{V}O_{2max}$ values for sex when the sample comprised both males and females, except for two studies by Klissouras (Klissouras, 1971; Klissouras et al., 1973). The age range in most studies was very restricted and two studies with a broader range corrected for age (Bouchard et al. (1986) and the current study). Univariate twin correlations (without siblings, estimated from a saturated model) from the current study were used in the meta-analysis (r_{MZ} = .58, r_{DZ} = .29 and r_{MZ} = .54, r_{DZ} = .38 for $\dot{V}O_{2max}$ in mL/min and mL/min/kg respectively).

Heterogeneity testing showed that all studies on the heritability of $\dot{V}O_{2max}$ expressed in mL/min (combined sample of 1088 individuals) could be taken together ($\chi^2(4) = 5.8$, $p = .218$) and a sample size weighted heritability estimate of 59% (95% CI: 52% – 66%) was found. For $\dot{V}O_{2max}$ expressed in mL/min/kg a weighted heritability estimate of 64% (95% CI: 60% – 69%) was found in a combined sample size of 3120 individuals, but heterogeneity testing showed that these studies could not be simply taken together ($\chi^2(4) = 12.2$, $p = .016$). Repeating the analysis without Sundet et al. (in which $\dot{V}O_{2max}$ was predicted) removed heterogeneity in the estimates of the four remaining studies ($p = .098$) and increased the weighted heritability estimate to 72% (N = 1004). For both $\dot{V}O_{2max}$ expressed in mL/min as $\dot{V}O_{2max}$ expressed in mL/min/kg, shared environmental influences were not significant ($p > .05$).

DISCUSSION

The main purpose of this paper was to estimate the heritability of aerobic fitness in an adolescent population, as assessed by $\dot{V}O_{2max}$ measured during a maximal exercise test. In concordance with previous literature, $\dot{V}O_{2max}$ was only moderately correlated with regular exercise behavior in leisure time. Genetic analysis revealed that 60% of the total variance in measured $\dot{V}O_{2max}$ in mL/min and 55% of the total variance in measured $\dot{V}O_{2max}$ in mL/min/kg can be explained by genetic factors.

In addition to measuring $\dot{V}O_{2max}$ during the climax of a graded maximal cycle ergometer test, $\dot{V}O_{2max}$ was predicted from submaximal tests on a cycle ergometer and a treadmill using extrapolation of the heart rate/oxygen uptake ($HR/\dot{V}O_2$) curve to the predicted HR_{max}. Only a moderate phenotypic relationship was found between predicted $\dot{V}O_{2max}$ and measured $\dot{V}O_{2max}$ in the current study ($.57 < r < .70$). This was lower than had been reported in previous studies using adult subjects (Ekblom-Bak et al., 2014; Grant et al., 1995; Legge & Banister, 1986). This difference can in part be attributed to the poor agreement between predicted and actual HR_{max}. Although there is substantial evidence that maximal heart rate is age-related in adults, it has been suggested that HR_{max} might be age-independent in children and adolescents (Rowland, 1996). When repeating the analysis using the *measured* HR_{max}, the phenotypic correlations between the measured $\dot{V}O_{2max}$ and the predicted $\dot{V}O_{2max}$ indeed increased. Nonetheless, correlations remained below those found for adults, suggesting that, apart from

the higher individual variation in HR_{max}, the variability in the HR/$\dot{V}O_2$ relationship may also be higher in adolescents.

In spite of the moderate phenotypic correlation to measured $\dot{V}O_{2max}$, heritability estimates from multivariate genetic analyses showed that heritability estimates for predicted $\dot{V}O_{2max}$ (46% to 67%) were very similar to those obtained for measured $\dot{V}O_{2max}$. For $\dot{V}O_{2max}$ in mL/min, the heritability estimates were higher than measured $\dot{V}O_{2max}$, but for $\dot{V}O_{2max}$ in mL/min/kg the heritability estimates were as high (treadmill test) or lower (cycle ergometer test) than measured $\dot{V}O_{2max}$. However, as all heritability estimates are within the confidence interval of measured $\dot{V}O_{2max}$, the differences were not significant. Moreover there was a substantial overlap in the genetic factors influencing predicted and measured $\dot{V}O_{2max}$. That genetic effects on $\dot{V}O_{2max}$ can be reliably estimated from submaximal tests is important as submaximal tests may be more amenable in large-scale studies. The graded maximal exercise test requires strenuous physical activity from the subject, which produces discomfort, and cannot be attained by or poses a health risk for some subgroups of the population (e.g. sedentary individuals, young children, elderly or patients suffering from cardiovascular or respiratory disease). It may also lead to a larger selection bias when recruiting volunteers from population-based samples (like twin registries) as not all subjects may be willing to exercise to exhaustion. This favors the participation of regular exercisers over sedentary subjects to exercise testing studies which will lead to biased estimates of both mean and variance in $\dot{V}O_{2max}$. The use of submaximal tests may lead to samples that are more representative of the general population.

Our sample size weighted meta-analysis on all heritability studies in children, adolescents and young adults to date showed that 59% (when expressed in mL/min) (N = 1088) and 72% (when expressed in mL/min/kg) (N = 1004) of the variance in measured $\dot{V}O_{2max}$ can be explained by genetic influences. All studies converge on the absence of detectable shared environmental factors (C). Shared environmental influences, including the family environment, were also low and not significant in the current study (except for predicted $\dot{V}O_{2max}$ expressed in mL/min/kg) but the power to detect C was low, even after adding siblings of the twins to the design. Power analysis suggests that our sample size had to be at least twice as big for C to be detected with 80% power (Posthuma & Boomsma, 2000). This leads us

to suggest that shared environmental influences on adolescent $\dot{V}O_{2max}$ cannot be excluded but at best play a very modest role.

The overarching conclusion from our (meta-)analyses is that $\dot{V}O_{2max}$ is a highly heritable phenotype from childhood to young adulthood. Heritability is likely to continue into adulthood, but there were no middle-aged or older adult twin samples that could be included in our meta-analysis. We did find four studies that measured $\dot{V}O_{2max}$ in parents and offspring. In these parent-offspring designs, however, heritability estimation can be affected by cohort effects, since different genetic variants affecting aerobic fitness can be expressed at different ages. To get a complete picture of the heritability of $\dot{V}O_{2max}$ across the whole life-span, twin studies focusing on middle-aged and older samples are direly needed.

A limitation of our study is that we cannot currently determine the exact contribution of the two different components that make up the heritability of $\dot{V}O_{2max}$: genetic factors that contribute to baseline (untrained) performance levels and those related to 'gain' in $\dot{V}O_{2max}$ (i.e. genetic factors contributing to aerobic trainability). The HERITAGE study showed that the variation in baseline performance, as well as the variance in trainability is larger between families than within families, confirming the role of genetic factors in baseline levels as well as in gain in $\dot{V}O_{2max}$ (Bouchard et al., 1998; Bouchard et al., 1999). Our study used a mixture of sedentary subjects and moderately and vigorous exercisers. $\dot{V}O_{2max}$ in the two latter groups will reflect a mixture of the baseline and trainability components. A possible way to discriminate between the two components is by estimating the heritability of $\dot{V}O_{2max}$ in untrained (persistent sedentary) individuals only. A further limitation is that even though the current study is the largest twin study on measured $\dot{V}O_{2max}$, our sample is still too small to have enough power to analyze sex differences in $\dot{V}O_{2max}$. It might be that the effects of genetic or environmental factors on $\dot{V}O_{2max}$ differ between males and females. A limitation of our meta-analysis is that there was significant heterogeneity across the studies, so that a single estimate therefore does not capture all individual studies adequately. However, recomputation of heritability in a restricted, more homogenous subset led to similar estimates. Finally, it should be noted that maximal exercise tests performed on a cycle ergometer generally yield lower $\dot{V}O_{2max}$ values than maximal exercise tests performed on a treadmill due to a larger excising muscle mass. Comparing the heritability studies of $\dot{V}O_{2max}$ performed on a treadmill and cycle ergometer showed that the two heritability estimates of

treadmill-derived $\dot{V}O_{2max}$ are slightly higher (Klissouras, 1971; Maes et al., 1996), but these estimates are based on small sample sizes consisting of 10-year-olds. Replication of these studies in other age groups is needed to examine the effect of exercise equipment on the heritability of $\dot{V}O_{2max}$.

Twin studies offer a unique opportunity to estimate the importance of genetic and environmental influences on a trait. Estimates of heritability inform us on how much of the variation in a phenotype in a population sample is due to genetic variation and generally define the upper limit of the percentage of variance that is explained by genetics, but does not reveal which and how many genes are involved. Therefore, an important next step is to identify the genetic variants underlying the heritability of $\dot{V}O_{2max}$. Thus far, studies reported case-control candidate gene and linkage studies, mostly characterized by small sample sizes and mixed results (Bouchard et al., 2011a). Two of the most studied polymorphisms is the R577X variation in the *ACTN3* gene and the I/D polymorphism in the *ACE* gene (MacArthur & North, 2011; Skipworth et al., 2011). The preferred approach to identify genetic variants for complex traits (which are known to be influences by multiple genetic factors) is a meta-analysis of genome-wide association (GWA) studies with a large cumulative sample size (Flint, 2013; Visscher et al., 2012). Only one GWA study on $\dot{V}O_{2max}$ has been conducted to date by Bouchard et al. (2011). Strikingly, in spite of the small sample size, this study revealed that 16 Single Nucleotide Polymorphisms (SNPs) accounted for 45% of the variance in gains in $\dot{V}O_{2max}$ after exposure to a standardized 20-weeks exercise program in a sample of 473 sedentary adults (Bouchard et al., 2011b). No GWA studies have yet been performed on $\dot{V}O_{2max}$ in the untrained or baseline state (before training). Such studies will need large samples with both $\dot{V}O_{2max}$ data and genome-wide genotyping. The feasibility of this increases greatly if submaximal exercise tests generate sufficiently valid estimates. Notwithstanding the imperfect correlation between predicted and measured $\dot{V}O_{2max}$, our results can be considered encouraging: The high genetic correlation between measured and predicted $\dot{V}O_{2max}$ in the current study suggests that they largely capture the same latent genetic factors and these genetic factors explained the largest part of the observed correlation between measured and predicted $\dot{V}O_{2max}$. GWA meta-analyses across studies using (graded) submaximal and maximal tests should be able to pick up these shared genetic variants.

To conclude, the results of the current study, together with the results of the meta-analyses, confirm that innate factors determine more than half of the individual differences in the $\dot{V}O_{2max}$ from childhood to young adulthood.