



## MATERIALS AND METHODS

**Patients.** Sixteen patients with JDM participated in this study. Patients were recruited from the pediatric rheumatology outpatient clinic of the Wilhelmina Children's Hospital of the University Medical Center Utrecht and fulfilled the criteria for the diagnosis of JDM as described by Bohan and Peter<sup>18</sup>. The patients were considered to be in a stable clinical condition. Their characteristics are described in Table 1. Parents and patients gave their informed consent for participating in the study. The Institutional Review Board approved all procedures. Patients were tested and retested within  $12.8 \pm 5.7$  days during the summer of 2003. All tests were performed by the same clinical exercise physiologist (TT), with 6 years' experience in exercise testing of patients with a pediatric rheumatic disease.

**Anthropometry.** The participants' body mass and height were determined using an electronic scale and a measuring stick. Body composition was assessed using the sum of 7 skin folds ( $\Sigma 7SF$ ) method according to Pollack, *et al*<sup>19</sup>. The measurements were taken at 7 sites (on the right side of the body: triceps, biceps, subscapular, suprailiac, mid-abdominal, medial calf, and thigh) by the investigator (TT) in accord with the American College of Sports Medicine guidelines<sup>20</sup>.

**Anaerobic exercise capacity.** The WAnT was performed as described by Bar-Or<sup>13</sup> on a recently calibrated electromagnetic braked cycle ergometer (Lode Examiner, Lode BV, Groningen, The Netherlands). The ergometer was upgraded and calibrated by the manufacturer to a maximal resistance of 800 W instead of the standard 400 W. The external resistance was controlled and the power output was measured using the Lode Wingate software package<sup>21</sup>. The seat height was adjusted to the patient's leg length (comfortable cycling height).

The external load (torque; in  $N \cdot m$ ) was determined, dependent on body weight (at  $0.53 \times$  body weight and  $0.55 \times$  body weight for girls and boys under 14 years of age, and  $0.67 \times$  body weight and  $0.7 \times$  body weight for older girls and boys, respectively) according to the user manual.

The patients were instructed to exercise for 1 min at the cycle ergometer with an external load of 15 W at 50 RPM. Then the sprint protocol started; patients were instructed to cycle all-out for 30 s. The measured variables were mean power, peak power, lowest power, and fatigue index. Mean power represented the average power output over the 30 s sprint. Peak power was the highest recorded power output achieved during the 30 s sprint, lowest power was the lowest power recorded during the 30 s sprint. The fatigue index represented the decline in power output (fatigue) during the 30 s sprint, and was calculated as highest power minus lowest power divided by the time interval between highest power and lowest power. Power output during the WAnT was corrected for the inertia of the mass of the flywheel ( $23.11 \text{ kg} \cdot \text{m}^2$ ).

**Aerobic exercise test.** Subjects performed an aerobic exercise test using an electronically braked cycle ergometer (Lode Examiner). Three of the smallest patients did not fit this ergometer; they were tested on a smaller mechanically braked ergometer (Tunturi, Tampere, Finland).

The seat height of the ergometer was adjusted to the patient's leg length. One minute of unloaded cycling preceded the application of resistance to the ergometer. Then the workload was increased by constant increment of 20 W every minute. This protocol continued until the patient stopped

Table 1. Subject's characteristics. Six patients had active disease; 2 were in clinical remission (remission while taking medication); 8 were in remission taking no medication.

	Mean	SD	Range
Age, yrs	13.85	6.4	6.7–27.2
Age at onset, yrs	7.6	3.5	3.1–14.6
Height, m	1.49	0.18	1.15–1.76
Body mass, kg	45.1	16.7	24.0–79.0
$\Sigma$ Skin folds, mm	139.0	77.5	34.7–287.0

because of volitional exhaustion, despite strong verbal encouragement from the investigators. The highest achieved workload during the test ( $W_{\text{max}}$ ) was recorded.

During the aerobic exercise test, subjects breathed through a facemask (Hans Rudolph Inc., Kansas City, MO, USA) connected to a calibrated metabolic cart (Oxycon Champion, Jaeger, Viasys, Balthoven, The Netherlands). Expired gas was passed through a flow meter, an oxygen ( $O_2$ ) analyzer, and a carbon dioxide ( $CO_2$ ) analyzer. The flow meter and gas analyzers were connected to a computer, which calculated breath-by-breath minute ventilation ( $V_e$ ), oxygen consumption ( $VO_2$ ), carbon dioxide production ( $VCO_2$ ), and respiratory exchange ratio (RER) from conventional equations. Heart rate (HR) was measured continuously during the aerobic exercise test by a bipolar electrocardiogram. Peak oxygen consumption ( $VO_{2\text{peak}}$ ) was taken as the average value over the last 30 s during the aerobic exercise test. Heart rate recovery ( $HR_{\text{recov}}$ ) was calculated as the maximum HR ( $HR_{\text{max}}$ ) attained during the exercise test minus the HR after 1 min recovery.

**Statistics.** The data were analyzed using SPSS 11.0 and MS Excel 98 for Windows. Correlation coefficients [Pearson correlations (R) and intraclass correlations (ICC)] were computed for test-retest reliability. Acceptable reliability criteria for ICC values were values  $> 0.75$ <sup>22</sup>.

Moreover, limits of agreement (LOA) were calculated to conform to the procedure described by Bland and Altman<sup>23</sup>. Bland-Altman analysis described the level of agreement between 2 measurements<sup>23</sup>. In this analysis, the "bias" is an estimate of how closely on average the 2 measurements agree and the "precision" indicates how well the methods agree for an individual<sup>23</sup>. The LOA are calculated by multiplying the precision by 1.96<sup>23</sup>. Typical error and total error were calculated according to Hopkins<sup>24</sup>. Typical error was calculated as the standard deviation in each subject's measurements between tests, after any shifts in the mean have been taken into account. Here the typical error is expressed as a percentage of the subject's mean score to obtain an interpretable percentage score. This percentage is also known as the coefficient of variation<sup>24</sup>. Total error was calculated as the mean of each subject's standard deviation between the trials<sup>25</sup>. The level of statistical significance was set at  $p = 0.05$ .

## RESULTS

Three of the youngest patients did not fit on the electromagnetic braked ergometer, therefore only data of 13 patients were available for analysis of the reliability of the WAnT. However, the youngest patients were able to perform the aerobic exercise test when they were tested on a suitable size mechanically braked ergometer (see Methods).

The physiological variables measured during the exercise test on both occasions are described in Table 2. The reliability statistics of both the WAnT and the aerobic exercise test are given in Table 3. As shown by Bland-Altman plots (Figures 1 and 2), there was one obvious outlier. Even when this outlier was included in the calculations, the reliability statistics are still acceptable. We found R and ICC values for mean power, peak power, and fatigue index were  $\pm 0.85$  or above. The typical error (expressed as a percentage) showed that peak power and mean power were the 2 variables of the WAnT with the lowest error of measurement. The lowest power and the fatigue index had almost twice as much variability compared to mean and peak power.

When this outlier was omitted from the analysis, the reliability statistics for peak power were  $R = 0.938$ ,  $ICC = 0.937$ ,  $LOA = 137.86 \text{ W}$ , typical error = 49.77 W, total error = 58.40 W, and typical error = 14.8%. The reliability statis-

Table 2. Reproducibility of performance in the duplicate anaerobic and aerobic exercise tests.

	Measurement 1 Mean (SD)	Measurement 2 Mean (SD)	Change in Mean
Peak power, W	442.23 (183.9)	481.92 (212.7)	39.69
Mean power, W	244.9 (93.4)	270.0 (113.8)	25.15
Lowest power, W	132.7 (62.8)	117.62 (55.7)	-15.08
Fatigue index	12.45 (6.6)	14.8 (7.3)	2.42
VO <sub>2peak</sub> , L • min <sup>-1</sup>	1.3 (0.44)	1.3 (0.39)	0.0
W <sub>max</sub> , W	101.7 (38.8)	104.7 (41.9)	2.7
RER <sub>max</sub>	1.14 (0.11)	1.14 (0.06)	0.01
HR <sub>max</sub> , beats • min <sup>-1</sup>	178.4 (16.5)	175.3 (14.7)	-3.06
HR <sub>recov</sub> , beats • min <sup>-1</sup>	51.13 (22.67)	48.25 (18.72)	-2.88

Change in mean denotes the change between measurement 1 and measurement 2, VO<sub>2peak</sub>: peak oxygen uptake, W<sub>max</sub>: maximal workload, RER<sub>max</sub>: maximal respiratory exchange ratio, HR<sub>max</sub>: maximal heart rate, HR<sub>recov</sub>: recovery heart rate.

tics for mean power after omission of this patient were R = 0.974, ICC = 0.972, LOA = 47.25 W, typical error = 17.07 W, total error = 21.9 W, and typical error was 8.5%. The reliability statistics of the aerobic exercise test showed a moderately low variability (Figure 3, Figure 4, Table 3). VO<sub>2peak</sub> and W<sub>max</sub> showed an R and ICC > 0.95, which indicated a very good reliability. HR<sub>max</sub> was also a very reliable measure, with R and ICC of ± 0.90. RER<sub>max</sub> and HR<sub>recov</sub> had R and ICC values between 0.59 and 0.78, indicating moderate reliability compared to the other measured variables. In the VO<sub>2peak</sub> test, there was either no change (VO<sub>2peak</sub>) or a very small change in the mean (W<sub>max</sub>, RER<sub>max</sub>, HR<sub>max</sub>, and HR<sub>recov</sub>), which indicated that there was no systematic learning effect of this test (Table 2).

## DISCUSSION

This is the first study investigating the reliability of exercise tests in patients with JDM. We found that both anaerobic and aerobic exercise tests were reliable exercise tests in patients with JDM. Tirosh, *et al*<sup>14</sup> found very good reliabil-

ity (R > 0.95) of the WAnT in a variety of patients with disorders such as Duchenne muscular dystrophy, Becker's muscular dystrophy, cerebral palsy, and spinal muscular atrophy. When the one outlier was omitted from the analysis we found comparable results in patients with JDM. This patient was known to have a behavioral dysfunction, and he sometimes lacked the motivation to perform in clinical tests. This case highlights that exercise tests are influenced by the motivation of the patient to give a maximal effort and that lack of motivation can influence the final test result.

The change in mean in both mean power and peak power during the WAnT showed that both increased during the second assessment, while the lowest power decreased and consequently the fatigue index increased. This result might be because one of the patients scored significantly higher in the second assessment compared to the first. When this patient was omitted from the analysis, the change in mean was reduced to 18 W for peak power and 10 W for mean power. When using the WAnT as an outcome for an intervention study, one should consider a learning session to rule out this small systematic learning effect, which could affect the power of a study.

The reliability statistics of the aerobic exercise test showed that this test was very accurate in determining VO<sub>2peak</sub> and W<sub>max</sub> in patients with JDM. In previous studies in patients with juvenile arthritis, error of measurement in VO<sub>2peak</sub> of 10% to 12% were found<sup>26,27</sup>. In our sample of patients with JDM the error of measurement was 5.5%. This difference might reflect the day-to-day variation in disease expression in patients with juvenile arthritis, while patients with JDM might have less short-term fluctuation in their disease state. The total error of VO<sub>2peak</sub> was 0.8 l • min<sup>-1</sup>, and was well within the quality range of certified exercise physiology laboratories<sup>28</sup>, and indicates the reliability of the aerobic exercise test. Our error level was comparable with observations in healthy children<sup>29</sup>.

Since JDM is a very rare disease, our sample size was small. To improve sample size, multicenter studies could be

Table 3. Reliability statistics of the anaerobic and the aerobic exercise tests.

	R	ICC	Typical Error	Total Error	LOA	Typical Error, % (CV)
Peak power, W	0.879	0.869	71.88	91.38	199.11	18.7
Mean power, W	0.860	0.851	41.17	54.76	114	16.8
Lowest power, W	0.735	0.73	30.84	31.06	85.42	41.0
Fatigue index	0.849	0.845	2.74	3.79	7.6	28.3
VO <sub>2peak</sub> , L • min <sup>-1</sup>	0.962	0.955	0.09	0.08	0.24	5.8
W <sub>max</sub> , W	0.969	0.969	7.28	4.0	20.15	3.6
RER <sub>max</sub>	0.677	0.589	0.06	0.05	0.15	4.8
HR <sub>max</sub> , beats • min <sup>-1</sup>	0.904	0.898	4.99	4.27	13.82	2.4
HR <sub>recov</sub> , beats • min <sup>-1</sup>	0.78	0.767	10.04	10.61	27.8	28.8

R: Pearson product moment correlation, ICC: intraclass correlation coefficient, LOA: limits of agreement, CV: coefficient of variation, VO<sub>2peak</sub>: peak oxygen uptake, W<sub>max</sub>: maximal workload, RER<sub>max</sub>: maximal respiratory exchange ratio, HR<sub>max</sub>: maximal heart rate, HR<sub>recov</sub>: recovery heart rate.

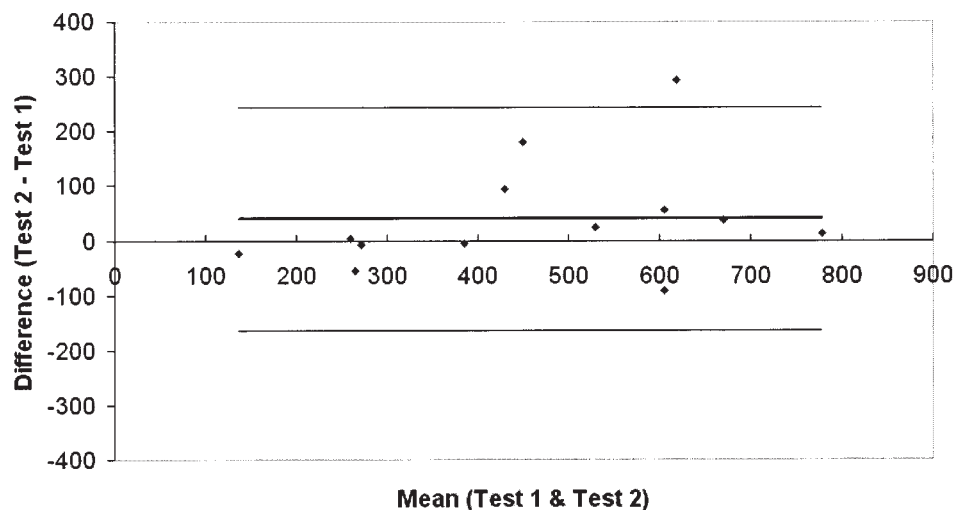


Figure 1. Bland–Altman plot of peak power during test and retest on the WAnT. Bold line shows the mean difference between the 2 measurement methods, the 2 thin lines indicate  $\pm 2$  SD. X axis: average peak power value from both tests. Y axis: difference between peak power during the test and peak power during the retest.

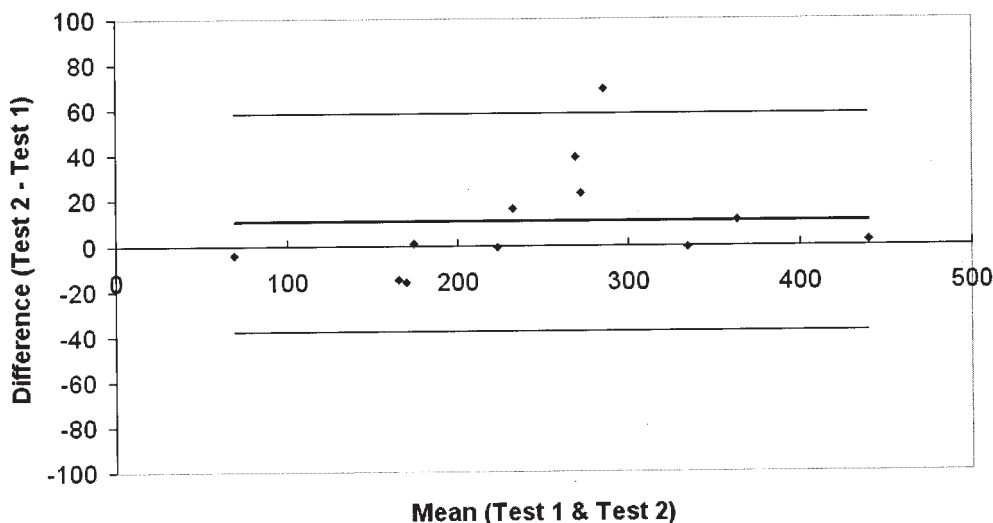


Figure 2. Bland–Altman plot of mean power during test and retest on the WAnT. Bold line shows the mean difference between the 2 measurement methods, the 2 thin lines indicate  $\pm 2$  SD. X-axis: average mean power value from both tests. Y axis: difference between mean power during the test and mean power during the retest.

initiated. However, this could introduce bias due to the use of different equipment.

The introduction of clinical exercise physiology as a diagnostic and therapeutic tool in the management of JDM seems promising. Not only was aerobic and anaerobic testing found to be reliable, in our clinical practice it also is feasible to conduct both exercise testing modes in patients with JDM over a large age span. This will facilitate followup of these patients well into adulthood, since most commonly used outcome measures (e.g., the Childhood Health Assessment Questionnaire and Childhood Myositis Assessment Score) are age-limited, as they are designed and

validated for use in children and adolescents only. However, exercise tests have been successfully performed over a very wide age span, from as young as 5 years of age<sup>30</sup> to 75 years<sup>31</sup>, and thus will facilitate the longitudinal followup of our patients. Future studies should determine the sensitivity to change of both anaerobic and aerobic exercise tests during exercise therapy and medical treatment in patients with JDM.

We found acceptable reliability for the WAnT and very good reliability for the aerobic exercise test ( $VO_{2peak}$  and  $W_{max}$ ). This indicates that exercise tests can be reliably performed in patients with JDM.

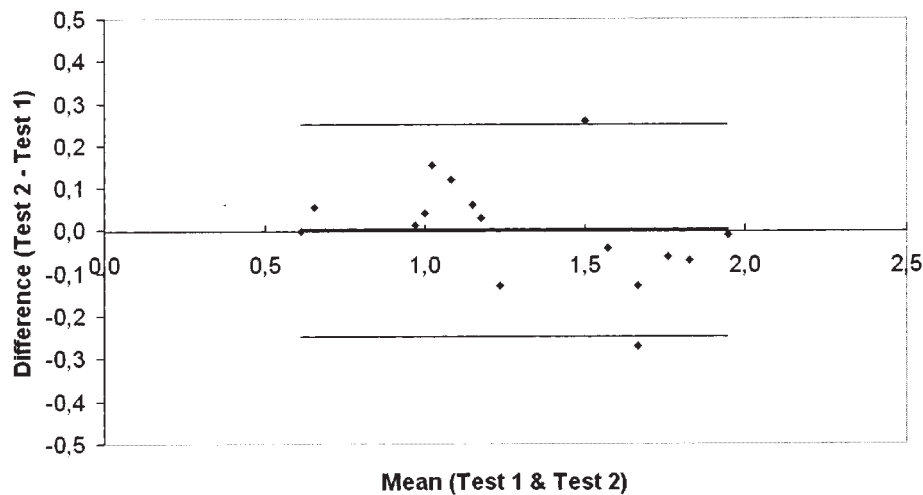


Figure 3. Bland-Altman plot of  $VO_{2peak}$  during test and retest. Bold line shows mean difference between the 2 measurement methods, the 2 thin lines indicate  $\pm 2$  SD. X axis: average  $VO_{2peak}$  value from both tests. Y axis: difference of  $VO_{2peak}$  during the test and the retest.

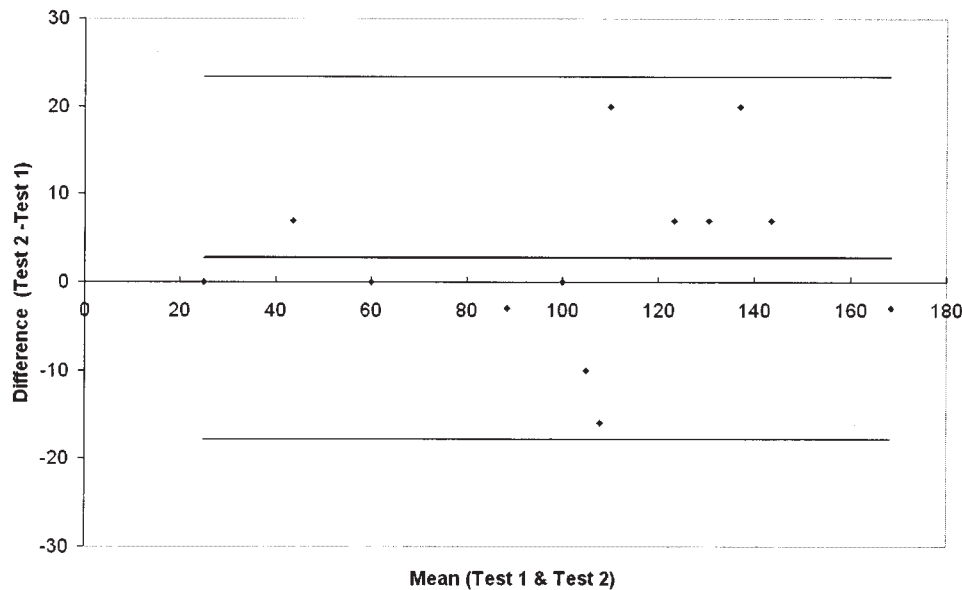


Figure 4. Bland-Altman plot of  $W_{max}$  during test and retest. Bold line shows mean difference between the 2 measurement methods, the 2 thin lines indicate  $\pm 2$  SD. X axis: average  $W_{max}$  value from both tests. Y axis: difference of  $W_{max}$  during the test and the retest.

## ACKNOWLEDGMENT

We thank Susan Maillard, MSc, for her review of the manuscript.

## REFERENCES

- Pachman LM. Juvenile dermatomyositis. Pathophysiology and disease expression. *Pediatr Clin North Am* 1995;42:1071-98.
- Resnick JS, Mammel M, Mundale MO, Kottke FJ. Muscular strength as an index of response to therapy in childhood dermatomyositis. *Arch Phys Med Rehabil* 1981;62:12-9.
- Takken T, Spermon N, Helder PJ, Prakken AB, van der Net J. Aerobic exercise capacity in patients with juvenile dermatomyositis. *J Rheumatol* 2003;30:1075-80.
- Hicks JE, Drinkard B, Summers RM, Rider LG. Decreased aerobic capacity in children with juvenile dermatomyositis. *Arthritis Rheum* 2002;47:118-23.
- Wiesinger GF, Quittan M, Nuhr M, et al. Aerobic capacity in adult dermatomyositis/polymyositis patients and healthy controls. *Arch Phys Med Rehabil* 2000;81:1-5.
- Hebert CA, Byrnes TJ, Baethge BA, Wolf RE, Kinasewitz GT. Exercise limitation in patients with polymyositis. *Chest* 1990;98:352-7.
- Newman ED, Kurland RJ. P-31 magnetic resonance spectroscopy in polymyositis and dermatomyositis. *Arthritis Rheum* 1992; 35:199-203.

8. Niermann KJ, Olsen NJ, Park JH. Magnesium abnormalities of skeletal muscle in dermatomyositis and juvenile dermatomyositis. *Arthritis Rheum* 2002;46:475-88.
9. Park JH, Niermann KJ, Ryder NM, et al. Muscle abnormalities in juvenile dermatomyositis patients: P-31 magnetic resonance spectroscopy studies. *Arthritis Rheum* 2000;43:2359-67.
10. Pfliegerer B, Lange J, Loske K, Sunderkotter C. Metabolic disturbances during short exercises in dermatomyositis revealed by real-time functional 31P magnetic resonance spectrometry. *Rheumatology Oxford* 2004;43:696-703.
11. Bailey RC, Olson J, Pepper SL, Porszasz J, Barstow TJ, Cooper DM. The level and tempo of children's physical activities: an observational study. *Med Sci Sports Exerc* 1995;27:1033-41.
12. Takken T, van der Net J, Helders PJ. Relationship between functional ability and physical fitness in juvenile idiopathic arthritis patients. *Scand J Rheumatol* 2003;32:174-8.
13. Bar-Or O. The Wingate anaerobic test. An update on methodology, reliability and validity. *Sports Med* 1987;4:381-94.
14. Tirosh E, Bar-Or O, Rosenbaum P. New muscle power test in neuromuscular disease. Feasibility and reliability. *Am J Dis Child* 1990;144:1083-7.
15. Malleon PN, Bennett SM, MacKinnon M, et al. Physical fitness and its relationship to other indices of health status in children with chronic arthritis. *J Rheumatol* 1996;23:1059-65.
16. Inbar O, Bar-Or O, Skinner JS. The Wingate anaerobic test. Champaign, IL: Human Kinetics; 1996.
17. Takken T, van der Net J, Helders PJM. Anaerobic exercise capacity in juvenile-onset idiopathic inflammatory myositis patients. *Arthritis Rheum* 2005; (in press).
18. Bohan A, Peter JB. Polymyositis and dermatomyositis (second of 2 parts). *N Engl J Med* 1975;292:403-7.
19. Pollack ML, Schmidt DH, Jackson AS. Measurement of cardio-respiratory fitness and body composition in the clinical setting. *Compr Ther* 1980;6:12-27.
20. Latin RW. Surface anatomy. In: Roitman JL, editor. ACSM's resource manual for guidelines for exercise testing and prescription. 3rd ed. Baltimore: Williams and Wilkins; 1998:89-100.
21. Lode. Wingate software. Version 1.09 for DOS. Groningen: Lode BV; 1999.
22. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 2nd ed. Upper Saddle River, NJ: Prentice Hall; 2000.
23. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-10.
24. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 2000;30:1-15.
25. Knapp TR. Technical error of measurement: a methodological critique. *Am J Phys Anthropol* 1992;87:235-6.
26. Takken T, van der Net J, Helders PJM. Aerobic exercise testing in juvenile rheumatoid arthritis patients. *Clin Exerc Physiol* 2002;4:38-43.
27. Giannini MJ, Protas EJ. Exercise response in children with and without juvenile rheumatoid arthritis: a case-comparison study. *Phys Ther* 1992;72:365-72.
28. Gore CJ. Quality assurance in exercise physiology laboratories. In: Gore CJ, editor. Physiological tests for elite athletes. Champaign, IL: Human Kinetics; 2000:3-11.
29. Boileau RA, Bonen A, Heyward VH, Massey BH. Maximal aerobic capacity on the treadmill and bicycle ergometer of boys 11-14 years of age. *J Sports Med Physical Fitness* 1977;17:153-62.
30. LeMura LM, von Duvillard SP, Cohen SL, et al. Treadmill and cycle ergometry testing in 5- to 6-year-old children. *Eur J Appl Physiol* 2001;85:472-8.
31. Shvartz E, Reibold RC. Aerobic fitness norms for males and females aged 6 to 75 years: a review. *Aviat Space Environ Med* 1990;61:3-11.