

Reproducibility of cardiovascular, respiratory, and metabolic responses to submaximal exercise: The HERITAGE Family Study

JACK H. WILMORE, PHILIP R. STANFORTH, KENNETH R. TURLEY, JACQUES GAGNON, E. WARWICK DAW, ARTHUR S. LEON, D. C. RAO, JAMES S. SKINNER, and CLAUDE BOUCHARD

The University of Texas at Austin, Austin, TX; Laval University, Montreal, CANADA; Washington University, St. Louis, MO; University of Minnesota, Minneapolis, MN; Indiana University, Bloomington, IN

ABSTRACT

WILMORE, J. H., P. R. STANFORTH, K. R. TURLEY, J. GAGNON, E. W. DAW, A. S. LEON, D. C. RAO, J. S. SKINNER, and C. BOUCHARD. Reproducibility of cardiovascular, respiratory, and metabolic responses to submaximal exercise: The HERITAGE Family Study. *Med. Sci. Sports Sci.*, Vol. 30, No. 2, pp. 259–265, 1998. This study determined the reproducibility of cardiovascular, respiratory, and metabolic responses to submaximal cycle ergometer exercise at two power outputs (50 W and 60% $\dot{V}O_{2max}$) on each of two separate days in a sample of 390 subjects (198 men and 192 women) participating in the HERITAGE Family Study. The same protocol was conducted across 3 d in an intracenter quality control substudy which included an additional 55 subjects. Reproducibility estimates included technical error, coefficient of variation, and intraclass correlation for each of the selected variables for both subject populations. Further, since the data were collected across four clinical centers, intraclass correlations were also computed separately by clinical center. Most variables were highly reproducible, with coefficients of variation below 9% and intraclass correlations over 0.80. These results were consistent for both subject populations and across all four clinical centers. Reproducibility indicators were generally better at the higher power output. It is concluded that within-subject day-to-day variation and measurement unreliabilities are generally small compared with the between-subject variance in the response to submaximal exercise at each of the clinical centers of the HERITAGE Family Study. **Key Words:** REPRODUCIBILITY, CARDIAC OUTPUT, STROKE VOLUME, HEART RATE, BLOOD PRESSURE, O_2 UPTAKE, CO_2 PRODUCTION, RER, SUBMAXIMAL EXERCISE

The ability to track accurately change in physiological variables consequent to an exercise training intervention is in large part dependent on the reproducibility of the measurements, which reflects both trial-to-trial reliability as well as day-to-day biological variability. When measurements are less reliable (i.e., with large measurement error) or not highly reproducible, it is difficult to demonstrate change accurately. Unfortunately, the literature is replete with exercise training studies that have conducted only a single assessment pre-intervention and post-intervention without establishing reproducibility of their measurements.

The HERITAGE Family Study is a large multicenter clinical trial investigating possible genetic basis for the variability in the response of physiological measures and risk factors for cardiovascular disease and non-insulin-dependent diabetes mellitus to endurance exercise training. With a multicenter trial, it is critical that all clinical centers

use exactly the same protocols for measurements. Further, to detect relatively small changes in critical variables and to establish the possible genetic basis for these changes, it is imperative to have extremely high reproducibility. Thus, the purpose of this study was to establish the reproducibility of important cardiovascular and metabolic variables obtained during submaximal exercise at two different rates of work, 50 W and 60% of maximal oxygen uptake ($\dot{V}O_{2max}$), before and after a 20-wk endurance training program in HERITAGE Family Study subjects, and in a separate group of quality control subjects, across four participating clinical centers. Details of the HERITAGE Family Study have been presented previously (2).

METHODS

Subjects. The HERITAGE subject population is composed of families, including the natural father and mother and at least two (African-American families) or three (Caucasian families) offspring 17 yr of age or older. Subjects were recruited by each of the four clinical centers, located at Indiana University, Laval University, the University of Minnesota, and The University of Texas at Austin. Subjects for this study represent the first 390 subjects to participate in the

0195-9131/98/3002-0259\$3.00/0

MEDICINE & SCIENCE IN SPORTS & EXERCISE®
Copyright © 1998 by the American College of Sports Medicine

Submitted for publication May 1996.

Accepted for publication June 1997.

TABLE 1. Physical characteristics of the subjects in the HERITAGE Family Study and in the Intracenter Quality Control (ICQC) Substudy.

Subjects	Age (yr)	Height (cm)	Weight (kg)
HERITAGE Family Study			
Women (N = 198)	34.8 ± 13.9	163.7 ± 6.5	68.1 ± 13.8
Men (N = 192)	35.1 ± 14.7	177.6 ± 6.5	84.0 ± 16.0
All subjects (N = 390)	34.9 ± 14.3	170.8 ± 9.5	76.2 ± 16.9
ICQC Substudy			
Women (N = 30)	27.4 ± 8.4	165.6 ± 7.0	64.5 ± 11.5
Men (N = 25)	28.8 ± 9.9	176.5 ± 6.3	76.1 ± 10.9
All subjects (N = 55)	28.1 ± 9.2	171.4 ± 8.6	70.8 ± 12.5

* Data are expressed as mean ± SD.

HERITAGE Family Study protocol. Of these subjects, 192 were women and 198 were men, and their ages ranged from 17 to 65 yr. Subject characteristics are presented in Table 1. In addition to the above criteria, all subjects had to meet a set of inclusionary criteria (2) and to pass a physician-administered physical examination. This included a resting and exercise 12-lead electrocardiogram, the latter obtained during a maximal exercise test (2). The study protocol had been previously approved by each clinical center's institutional review board, and informed consent was obtained from each subject.

Each clinical center recruited additional subjects every 6 months over three consecutive 6-month periods to participate in an intracenter quality control substudy (ICQC substudy) during the second and third year of data collection. Subjects in the ICQC substudy were required to meet all criteria for admission to the HERITAGE Family Study with the exception of family membership. Data were available for 55 subjects across all four centers. Their characteristics are presented in Table 1.

Experimental design. Each subject in the HERITAGE Family Study completed an extensive battery of tests prior to starting the 20-wk endurance training program, including three cycle ergometer exercise tests conducted on separate days: a maximal exercise test (Max), a submaximal exercise test (Submax), and a submaximal/maximal exercise test (Submax/Max) which combined the previous Submax and Max tests (2). The training program was conducted on cycle ergometers (Universal Aerobicycle, Cedar Rapids, IO) interfaced with a Mednet computer system (Universal Gym Mednet, Cedar Rapids, IO) to control the power output of the ergometers to maintain constant training heart rates. Subjects started training at the heart rate associated with 55% of their initial maximal oxygen uptake ($\dot{V}O_{2max}$) for 30 $\text{min}\cdot\text{d}^{-1}$ and gradually progressed to the heart rate associated with 75% of their initial $\dot{V}O_{2max}$ for 50 $\text{min}\cdot\text{d}^{-1}$ at the end of 14 wk. They maintained this intensity and duration throughout the remaining 6 wk. Frequency was maintained at three sessions per week throughout the 20-wk training program. See Bouchard et al. (2) for further details concerning the training program. At the end of the 20-wk training program, the subjects completed an identical post-training test battery.

Subjects enrolled in the ICQC substudy completed the initial maximal exercise test to establish the power output necessary to elicit 60% of their $\dot{V}O_{2max}$. They then per-

formed three Submax/Max tests within a 3-wk period to establish reproducibility of the test results at each clinical center. These tests were separated by at least 4 d, but not by more than 10 d.

Exercise test methodology. Maximal and submaximal exercise tests were conducted on a cycle ergometer (SensorMedics Ergo-Metrics 800S, Yorba Linda, CA). Subjects completed an initial maximal exercise test using a graded exercise test protocol starting at 50 W for 3 min. The rate of work was then increased by 25 W every 2 min thereafter to the point of exhaustion. For older, smaller, or less fit subjects, the test was started at 40 W and increased by 10 to 20 W increments. Using the results of this initial maximal test, subjects then performed the Submax exercise test at 50 W and at 60% of their initial $\dot{V}O_{2max}$. The Submax/Max exercise test was then performed, starting with the Submax protocol, i.e., 50 W and 60% of the initial $\dot{V}O_{2max}$, and progressing to a maximal level of exertion. For this last test, a venous catheter was inserted into the left arm and blood samples were obtained at rest, during exercise at 50 W, 60% of $\dot{V}O_{2max}$, and 80% of $\dot{V}O_{2max}$, and immediately upon completion of the maximal test. The results of these tests were used to establish the endurance training program work rates and to quantify the magnitude of the training response (2).

During the Submax and Submax/Max tests, two cardiac output, heart rate, and blood pressure measurements were obtained and averaged both at 50 W and at 60% of the initial $\dot{V}O_{2max}$. Subjects exercised for approximately 12–15 min at each work rate, with a 4-min period of seated rest between work rates. Cardiac output was determined using the Collier CO_2 rebreathing technique (3), as described by Wilmore et al. (18). Each clinical center used the same electronic mixing system to assure the proper volume and concentration of CO_2 for rebreathing dependent on the subjects' steady state $\dot{V}O_2$ and end tidal PCO_2 . For all tests, $\dot{V}O_2$, $\dot{V}CO_2$, expiratory minute ventilation ($\dot{V}E$), and the respiratory exchange ratio (RER) were determined every 20 s and reported as a rolling average of the three most recent 20-s values using a SensorMedics 2900 metabolic measurement cart (MMC). $\dot{V}O_{2max}$ was defined as the peak $\dot{V}O_2$ obtained during the test. Heart rate was determined by electrocardiography, and values were recorded during the last 15 s of each stage of the maximal test and once steady state had been achieved at each of the submaximal work rates during the Submax and Submax/Max tests. Blood pressures were obtained during the last minute of each stage of the maximal test protocol, and once when steady state was achieved at each of the submaximal work rates using a Colin STBP-780 (San Antonio, TX) automated blood pressure unit. Two electrocondenser microphones are embedded in the cuff, and the sound signal is synchronized to the ECG R-wave. Earphones allow the technician to confirm the blood pressure values selected by the instrument's detection algorithm.

Quality assurance, quality control, and statistical methodology. Several important quality assurance and quality control procedures were instituted across all four clinical centers, as detailed by Gagnon et al (4). Staff from

TABLE 2. Reproducibility of cardiovascular responses for the HERITAGE Family Study subjects ($N = 390$) across 2 d for two rates of submaximal exercise.

Variables	Submax Test Mean \pm SD	Submax/Max Test Mean \pm SD	TE	CV	ICC
50 W					
Heart rate (beats \cdot min $^{-1}$)	120.1 \pm 18.9	119.2 \ddagger \pm 18.6	6.1	5.1	0.89
Stroke volume (mL \cdot beat $^{-1}$)	97.8 \pm 20.8	98.1 \pm 20.2	7.4	7.6	0.87
Cardiac output (L \cdot min $^{-1}$)	11.5 \pm 1.7	11.4 \pm 1.7	0.8	7.2	0.76
Systolic BP (mmHg)	147.6 \pm 20.9	146.3 \pm 20.7	8.9	6.0	0.82
Diastolic BP (mmHg)	73.5 \pm 12.7	73.0 \pm 12.5	5.7	7.8	0.79
60% $\dot{V}O_{2max}$					
Heart rate (beats \cdot min $^{-1}$)	140.4 \pm 17.2	139.5 \pm 17.4	6.1	4.4	0.88
Stroke volume (mL \cdot beat $^{-1}$)	99.9 \pm 23.1	100.6 \pm 22.6	6.9	6.9	0.91
Cardiac output (L \cdot min $^{-1}$)	14.0 \pm 3.2	14.0 \pm 3.3	0.8	5.9	0.93
Systolic BP (mmHg)	164.0 \pm 21.6	163.3 \pm 22.0	9.3	5.7	0.82
Diastolic BP (mmHg)	74.2 \pm 12.9	74.0 \pm 12.8	6.1	8.3	0.77

TE, technical error; CV, coefficient of variation; ICC, intraclass correlation; \ddagger , significantly different from Submax test.

all clinical centers were trained centrally on several occasions, and all staff from each clinical center had to attain certification on each technique for which they were responsible. A detailed Manual of Procedures (MOP) was developed, and staff were required to review those sections of the MOP for which they were responsible every 6 months. Once each year for the first 2 yr of the study, a traveling crew of four subjects, two men and two women, went to each of the clinical centers over a 3- to 4-wk period, and were tested following the HERITAGE Family Study protocol at each clinical center, allowing comparisons to be made across the four clinical centers on these same four subjects (4).

All data were analyzed using the SAS statistical package. Except where noted, data are expressed as mean \pm SD. Technical errors, coefficients of variation, and intraclass correlations were computed to evaluate the reproducibility in both the HERITAGE sample and the ICQC sample using the model of Shrout and Fleiss (13). With this model, the i^{th} measurement on the j^{th} subject, x_{ij} , is given by:

$$x_{ij} = \mu + b_j + w_{ij}$$

where μ is the population mean, b_j is the difference from μ of the mean of the measurements on the j^{th} subject, and w_{ij} is the difference from $\mu + b_j$ of the i^{th} measurement on the j^{th} subject. Both b_j and w_{ij} are assumed to be normally distributed and independent with SDs of σ_τ and σ_ω , respectively. σ_ω is the within-subjects SD, also called technical error. The coefficient of variation within subjects was computed as:

$$CVW = (100 \cdot \sigma_\omega) / \mu$$

To compute the intraclass correlation coefficient, PROC GLM in SAS was used to run an ANOVA, providing a between-subjects mean square (BMS) and a within-subjects mean square (WMS). These were used to estimate the intraclass correlation (ICC) according to Shrout and Fleiss (13):

$$ICC = (BMS - WMS) / [BMS + (k - 1) \cdot WMS],$$

where k is the number of replicate measurements on a subject.

A multiple testing ANOVA was implemented using the General Linear Models Procedure to assess whether there were differences across the two pre-training and two post-training tests for the total sample and across the three

Submax/Max tests for the ICQC sample. Tukey's Studentized Range (HSD) Test was used to determine between which trials there were significant differences. The multiple testing ANOVA used controls for all potential sources of variation. Statistical significance was set at the 0.05 level.

RESULTS

The pre-training cardiovascular, respiratory, and metabolic responses for the total sample at 50 W and 60% of $\dot{V}O_{2max}$ during the Submax and Submax/Max tests are presented in Tables 2 and 3. The data in these tables represent the means of the two trials at each work rate for each test. There was excellent agreement between the Submax and Submax/Max test for all variables. While there were significant mean differences for pre-training heart rate at 50 W and pre-training $\dot{V}CO_2$ at 50 W, these differences were small and physiologically insignificant. The technical errors, coefficients of variation, and intraclass correlations were well within the expected range at both power outputs, but the data suggest a better reproducibility of test results at 60% of $\dot{V}O_{2max}$ than at 50 W.

Table 4 presents the intraclass correlations for each of the four clinical centers for the cardiovascular, respiratory, and metabolic variables, both pre- and post-training. With the exception of the RER, the intraclass correlations were reasonably high and consistent across the four centers.

Table 5 presents the results for the ICQC substudy. There was excellent agreement across the three trials. The only significant differences were observed for $\dot{V}O_2$ and RER at 60% of $\dot{V}O_{2max}$. Again, although these differences were statistically significant, they were small and of no physiological significance. The technical errors, coefficients of variation, and intraclass correlations were similar to those for the HERITAGE Family Study sample presented in Tables 2 and 3.

DISCUSSION

At its conclusion, approximately 750 subjects will have completed the entire HERITAGE Family Study protocol, including the assessment of cardiovascular, respiratory, and metabolic responses to three different submaximal and maximal exercise test protocols administered on separate days,

TABLE 3. Reproducibility of respiratory and metabolic responses for the HERITAGE Family Study subjects (*N* = 390) across 2 d for two rates of submaximal exercise.

Variables	Submax Test Mean ± SD	Submax/Max Test Mean ± SD	TE	CV	ICC
50 W					
VE (L·min ⁻¹)	31.6 ± 5.6	31.6 ± 5.9	2.1	6.7	0.87
VO ₂ (mL·min ⁻¹)	1055 ± 129	1049 ± 128	49.5	4.7	0.85
VCO ₂ (mL·min ⁻¹)	977 ± 127	969 ± 126	53.9	5.5	0.82
RER	0.93 ± 0.06	0.92 ± 0.05	0.03	3.3	0.69
60% VO _{2max}					
VE (L·min)	43.8 ± 11.3	43.8 ± 11.4	2.3	5.3	0.96
VO ₂ (mL·min ⁻¹)	1474 ± 451	1460 ± 455	53.4	3.6	0.99
VCO ₂ (mL·min ⁻¹)	1421 ± 451	1406 ± 453	59.3	4.2	0.98
RER	0.96 ± 0.05	0.96 ± 0.04	0.03	2.9	0.64

TE, technical error; CV, coefficient of variation; ICC, intraclass correlation; ‡, significantly different from Submax test.

both pre- and post-training. To determine pre- to post-training changes in these responses to both submaximal and maximal exercise, it is important to have accurate and highly reproducible measurements. Whereas the reproducibility of maximal exercise test data, such as maximal heart rate and $\dot{V}O_{2max}$, have been well established (7,10,15,17), there is little information on the reproducibility of submaximal exercise responses. Since a primary purpose of the HERITAGE Family Study is to determine the genetic basis of change in submaximal as well as maximal physiological responses to exercise, it is imperative to have highly reproducible measures.

The reproducibility estimates presented in this paper reflect intra-individual variability across days and the variability associated with measurement error. Reproducibility estimates were also calculated across trials on the same day for the Submax/Max test only. Since we were interested in the day-to-day variability these data were not reported. However, the intraclass correlations across trials on the same day, when compared with the mean of the two trials across days, were generally higher by 0.01 to 0.10, with the exception of diastolic blood pressure and RER at 60% of $\dot{V}O_{2max}$, which were higher by 0.11 and 0.17, respectively.

The reproducibility of the cardiovascular response data for the HERITAGE subjects (Table 2) and for the ICQC subjects (Table 5) compare favorably with the limited data available from other studies using cycle ergometry and the Collier technique (3) for determining cardiac output. Zeidifard et al. (20) determined the reproducibility of cardiac output, heart rate, and stroke volume in seven adults and three children during exercise at a $\dot{V}O_2$ of 1,200 mL·min⁻¹, a $\dot{V}O_2$ approximately midway between our mean values for 50 W and 60% of $\dot{V}O_{2max}$ work rates. They reported coefficients of variation, across at least 4 d, of 5.7% for cardiac output, 6.8% for heart rate, and 5.6% for stroke volume. Wolfe et al. (19) determined reproducibility of several cardiovascular variables over two separate days at three power outputs in 20 men. Using Pearson product-moment correlations, they reported correlations ranging from 0.83 to 0.94 for cardiac output, 0.90 to 0.92 for heart rate, 0.81 to 0.94 for stroke volume, and 0.60 to 0.86 for systolic blood pressure.

Van Herwaarden et al. (16) reported a coefficient of variation of 4.1% for cardiac output and 2.6% for heart rate at a mean submaximal $\dot{V}O_2$ of 1.45 L·min⁻¹. Their values appear to be substantially lower than those reported in the

TABLE 4. Intraclass correlations for selected cardiovascular, respiratory and metabolic variables pre- and post-training across clinical centers for the HERITAGE Family Study Subjects.

Variables	Clinical Center A <i>N</i> = 85		Clinical Center B <i>N</i> = 99		Clinical Center C <i>N</i> = 98		Clinical Center D <i>N</i> = 93	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
50 W								
Heart rate	0.88	0.87	0.87	0.85	0.88	0.85	0.92	0.91
Stroke volume	0.86	0.88	0.81	0.79	0.89	0.87	0.91	0.91
Cardiac output	0.81	0.79	0.67	0.72	0.85	0.77	0.76	0.77
Systolic BP	0.83	0.80	0.75	0.72	0.78	0.77	0.83	0.72
Diastolic BP	0.85	0.75	0.72	0.76	0.74	0.69	0.77	0.67
VE	0.82	0.84	0.83	0.84	0.88	0.79	0.86	0.84
VO ₂	0.85	0.94	0.84	0.85	0.85	0.84	0.82	0.84
VCO ₂	0.81	0.85	0.75	0.78	0.86	0.79	0.76	0.71
RER	0.75	0.43	0.47	0.51	0.66	0.46	0.72	0.43
60% VO ₂								
Heart rate	0.87	0.88	0.85	0.83	0.90	0.91	0.85	0.93
Stroke volume	0.86	0.93	0.91	0.93	0.91	0.93	0.92	0.94
Cardiac output	0.95	0.94	0.92	0.95	0.95	0.95	0.92	0.95
Systolic BP	0.84	0.91	0.77	0.77	0.76	0.83	0.86	0.73
Diastolic BP	0.81	0.81	0.79	0.68	0.64	0.77	0.73	0.72
VE	0.95	0.95	0.95	0.93	0.96	0.96	0.95	0.96
VO ₂	0.99	0.98	0.98	0.98	0.99	0.99	0.99	0.99
VCO ₂	0.98	0.97	0.98	0.98	0.99	0.98	0.98	0.99
RER	0.61	0.51	0.62	0.50	0.53	0.72	0.65	0.57

TABLE 5. Reproducibility of selected cardiovascular, respiratory, and metabolic variables across three trials for the ICQC substudy subjects ($N = 55$).

Variable	Trial 1 (mean \pm SD)	Trial 2 (mean \pm SD)	Trial 3 (mean \pm SD)	TE	CV	ICC
50 W						
Heart rate (beats·min ⁻¹)	107.5 \pm 16.4	106.7 \pm 13.5	106.3 \pm 13.2	6.5	6.1	0.79
Stroke volume (mL·beat ⁻¹)	104.9 \pm 20.8	106.2 \pm 21.6	103.3 \pm 20.9	8.9	8.5	0.82
Cardiac output (L·min ⁻¹)	11.1 \pm 1.8	11.2 \pm 2.0	10.8 \pm 1.7	0.9	8.0	0.77
Systolic BP (mmHg)	139.8 \pm 14.5	138.0 \pm 15.3	138.7 \pm 13.1	7.8	5.6	0.70
DBP (mmHg)	68.6 \pm 9.8	67.0 \pm 10.1	66.7 \pm 9.5	5.1	7.6	0.73
VE (L·min ⁻¹)	28.6 \pm 4.7	28.6 \pm 4.4	28.0 \pm 4.2	2.0	6.9	0.81
VO ₂ (mL·min ⁻¹)	1010 \pm 130	1007 \pm 116	982 \pm 111	46.9	4.7	0.85
VCO ₂ (mL·min ⁻¹)	912 \pm 124	909 \pm 115	882 \pm 95	52.0	5.8	0.79
RER	0.90 \pm 0.06	0.90 \pm 0.04	0.90 \pm 0.05	0.03	3.6	0.58
60% of $\dot{V}O_{2max}$						
Heart rate (beats·min ⁻¹)	144.4 \pm 14.4	142.9 \pm 14.0	142.9 \pm 12.3	5.7	4.0	0.82
Stroke volume (mL·beat ⁻¹)	111.3 \pm 21.3	113.1 \pm 24.9	112.4 \pm 21.4	7.2	6.5	0.90
Cardiac output (L·min ⁻¹)	16.0 \pm 3.0	16.0 \pm 3.4	16.0 \pm 3.0	0.7	4.5	0.95
SBP (mmHg)	166.3 \pm 21.0	165.9 \pm 19.5	166.4 \pm 19.2	8.4	5.1	0.82
DBP (mmHg)	72.1 \pm 11.1	70.5 \pm 11.1	69.7 \pm 10.4	5.9	8.4	0.70
VE (L·min ⁻¹)	49.6 \pm 9.1	50.5 \pm 9.9	50.1 \pm 10.1	2.5	5.1	0.93
VO ₂ (mL·min ⁻¹)	1787 \pm 445	1773 \pm 436	1744 \pm 421 \ddagger	62.2	3.5	0.98
VCO ₂ (mL·min ⁻¹)	1717 \pm 413	1722 \pm 425	1699 \pm 410	64.3	3.8	0.98
RER	0.96 \pm 0.05 \ddagger	0.97 \pm 0.05	0.97 \pm 0.05	0.03	2.6	0.72

\ddagger Significantly different from the other two trials ($P < 0.05$).

present study, but the details of their study have not been provided and it is unclear whether their coefficients of variation represent comparisons across trials or across days. Paterson et al. (12) conducted five repeat tests within 2 to 3 wk in 12 boys to determine the reproducibility of cardiac output, stroke volume, and heart rate at three submaximal power outputs. The coefficients of variation for cardiac output varied from 6.6% to 8.5%, for stroke volume from 7.2% to 10.8%, and for heart rate from 4.3% to 6.0%. Kirby and Shea (8) tested 15 subjects across three separate days and reported intraclass correlation coefficients for cardiac output of 0.69 at rest and at a low workload, and 0.87 at higher workloads. Our results compare favorably with these.

Becque et al. (1) investigated the reproducibility of heart rate, and systolic and diastolic blood pressures during submaximal cycle ergometry at power outputs of 50 W, 125 W, and 55% of each subject's maximum work rate. Four subjects exercised for 10 min at each power output and repeated each power output at least ten times on separate days. The coefficient of variation across the three power outputs varied from 2.9% to 3.4% for heart rate, 5.5% to 8.7% for systolic blood pressure, and from 7.6% to 13.1% for diastolic blood pressure, while the reliability coefficients ranged from 0.86 to 0.89 for heart rate, from 0.27 to 0.77 for systolic blood pressure, and from 0.64 to 0.84 for diastolic blood pressure. Finally, Hartung et al. (5) conducted repeat tests within approximately 1 wk of each other on 38 women who cycled at a submaximal work rate to attain steady-state heart rates. The correlation coefficient for heart rate between the two trials was 0.89, with a mean heart rate of 136.9 and 137.9 beats·min⁻¹ for the first and second trials, respectively.

The reproducibility of the respiratory and metabolic data for the HERITAGE subjects (Table 3) and for the ICQC subjects (Table 5) compare favorably with data from other studies. Taylor (14) was one of the first to look at reproducibility of submaximal exercise responses. He compared

the responses of 31 subjects to treadmill walking for 4 min on two separate days at a rate that resulted in a mean $\dot{V}O_2$ of 1.65 L·min⁻¹. Correlation coefficients for $\dot{V}O_2$, $\dot{V}E$, and heart rate were 0.59, 0.49, and 0.78, respectively, with coefficients of variation of 6.6%, 8.0%, and 4.1%, respectively. Henry (6) reported the variability of $\dot{V}O_2$ at 16 W ($N = 35$) and 100 W ($N = 25$) on a cycle ergometer, with reliability coefficients of 0.69 and 0.55, respectively. Both Taylor and Henry used closed circuit spirometry to measure $\dot{V}O_2$, a method known to have low reliability.

Wolfe et al. (19), as described previously, reported the reliability coefficients for $\dot{V}O_2$ at three power outputs on a cycle ergometer to range from 0.88 for light work up to 0.96 for heavy work. Becque et al. (1), as detailed earlier, found that the coefficient of variation across the three power outputs varied from 5.3% to 7.7% for $\dot{V}E$, and from 4.0% to 4.6% for $\dot{V}O_2$, while the reliability coefficients ranged from 0.69 to 0.97 for $\dot{V}E$ and from 0.79 to 0.95 for $\dot{V}O_2$. Melanson et al. (9) evaluated the reliability of $\dot{V}E$ and $\dot{V}O_2$ in 13 men at three rates of work (i.e., slow walk, slow jog, and run) on a treadmill across 2 d. Intraclass correlations for $\dot{V}E$ ranged from 0.87 to 0.93 and for $\dot{V}O_2$ from 0.77 to 0.94. Finally, Morgan et al. (11) reported an intraclass correlation of 0.95 for $\dot{V}O_2$ during level treadmill running (3.33 m·s⁻¹) across two test days, 1 to 3 d apart in 31 male runners.

We were unable to find studies that had reported any measure of reproducibility for submaximal RER. In the present data set for both the HERITAGE Family Study subjects (Table 3) and the ICQC subjects (Table 5), the intraclass correlations are generally low. However, the coefficients of variation and the technical errors are also low. At first, this would seem to be a contradiction, as the intraclass correlations would suggest poor reproducibility, whereas the coefficients of variation and technical errors would suggest the opposite. However, the real reason for the low intraclass correlation coefficients is the small range of variation across subjects for a given rate of work (i.e., the

SD within each trial is quite small; see Table 5). However, the diet 12 h prior to the exercise bout can also affect submaximal RER values, and there were no specific dietary controls enforced prior to exercise in these subjects.

There was excellent agreement across the four clinical centers with respect to the intraclass correlation coefficients (Table 4). For each variable, the intraclass correlation coefficients for both the 50 W and 60% $\dot{V}O_{2max}$ power outputs were, for the most part, very consistent across the clinical centers, and they were nearly identical to the intraclass correlation coefficients for the pooled data for each variable. Further, there was excellent agreement between the intraclass correlation coefficients pre- and post-training, indicating that reproducibility wasn't compromised longitudinally, which is particularly important when looking for changes in response to the training program. These results reflect the considerable effort that was put into standardizing the procedures across the four clinical centers, including identical equipment and protocols, centralized training and certification of all technical staff, and constant monitoring of the technical staff to assure compliance with the HERITAGE Family Study Manual of Procedures. The use of the traveling crew (see Methods, Quality Assurance) and the ICQC substudy also contributed to the quality of the data across the four clinical centers.

For many of the variables measured, there was better reproducibility at 60% of $\dot{V}O_{2max}$ compared with those at 50 W. This is largely the result of the fact that for these variables ($\dot{V}O_2$, $\dot{V}E$, $\dot{V}CO_2$, cardiac output, and stroke volume) the variation in response is much greater at different rates of work compared with the same rate of work. To elaborate, at 50 W on a cycle ergometer, subjects will have similar $\dot{V}O_2$ values since there is little variation between subjects in their efficiency to cycle at this absolute rate of work. Thus, intraclass correlations would be expected to be much lower for the same absolute rate of work where there is little intersubject variability in their physiological response compared with the same relative rate of work

(% $\dot{V}O_{2max}$) where the actual rate of work varies between subjects and intersubject variability in their physiological responses would be high. This would also be true for $\dot{V}E$ and $\dot{V}CO_2$. Since the major purpose of increasing cardiac output is to provide increased oxygen delivery to meet the needs for a given rate of work, it too, as well as stroke volume, would exhibit less variability between subjects for the same absolute rate of work compared with the same relative rate of work. This would result in lower intraclass correlations. These results point out important differences between using a fixed absolute rate of work and using a fixed relative rate of work for obtaining physiological responses at submaximal work rates.

In summary, the results from these early analyses of the HERITAGE Family Study data on cardiovascular, respiratory, and metabolic variables suggest good measurement reproducibility. This is critical to any subsequent analyses of this data set in which changes subsequent to endurance exercise training and the genetic basis of these changes are evaluated. Of particular importance was the close agreement across the four clinical centers, since the data from all clinical centers will be pooled for most all subsequent analyses.

The HERITAGE Family Study is supported by the National Heart, Lung and Blood Institute through the following grants: HL45670 (C. Bouchard, PI); HL47323 (A. S. Leon, PI); HL47317 (D. C. Rao, PI); HL47327 (J. S. Skinner, PI); and HL47321 (J. H. Wilmore, PI). Credit is also given to the University of Minnesota Clinical Research Center, NIH Grant MO1-RR000400. Further, Jack H. Wilmore is partially supported by the Margie Gurley Seay Centennial Professorship, and Art Leon is partially supported by the Henry L. Taylor Professorship in Exercise Science and Health Enhancement.

Thanks are expressed to all of the co-principal investigators, investigators, coinvestigators, local project coordinators, research assistants, laboratory technicians, and secretaries who have contributed to this study (see reference 2). Finally, the HERITAGE consortium is very thankful to those hardworking families whose participation has made these data possible.

Address for correspondence: Jack H. Wilmore, Department of Kinesiology & Health Education, Belmont Hall 222, The University of Texas at Austin, Austin, TX 78712.

REFERENCES

1. BECQUE, M. D., V. KATCH, C. MARKS, and R. DYER. Reliability and within subject variability of $\dot{V}E$, $\dot{V}O_2$, heart rate, and blood pressure during submaximum cycle ergometry. *Int. J. Sports Med.* 14:220-223, 1993.
2. BOUCHARD, C., A. S. LEON, D. C. RAO, J. S. SKINNER, J. H. WILMORE, and J. GAGNON. The HERITAGE Family Study: aims, design, and measurement protocol. *Med. Sci. Sports Exerc.* 27:721-729, 1995.
3. COLLIER, C. R. Determination of mixed venous CO_2 tensions by rebreathing. *J. Appl. Physiol.* 9:25-29, 1956.
4. GAGNON, J., M. A. PROVINCE, C. BOUCHARD, et al. The HERITAGE Family Study: Quality assurance and quality control. *Ann. Epidemiol.* 6:520-529, 1996.
5. HARTUNG, G. H., R. J. BLANCO, D. A. LALLY, and L. P. KROCK. Estimation of aerobic capacity from submaximal cycle ergometry in women. *Med. Sci. Sports Exerc.* 27:452-457, 1995.
6. HENRY, F. M. Individual differences in oxygen metabolism of work at two speeds of movement. *Res. Q.* 22:324-333, 1951.
7. KATCH, V. L., S. S. SADY, and P. FREEDSON. Biological variability in maximum aerobic power. *Med. Sci. Sports Exerc.* 14:21-25, 1982.
8. KIRBY, T. E. The CO_2 rebreathing technique for determination of cardiac output: Part I. *J. Card. Rehabil.* 5:97-101, 1985.
9. MELANSON, E. L., P. S. FREEDSON, D. HENDELMAN, and E. DEBOLD. Reliability and validity of a portable metabolic measurement system. *Can. J. Appl. Physiol.* 21:109-119, 1996.
10. MITCHELL, J. H., B. J. SPROULE, and C. B. CHAPMAN. The physiological meaning of the maximal oxygen intake test. *J. Clin. Invest.* 37:538-547, 1958.
11. MORGAN, D. W., P. E. MARTIN, G. S. KRAHENBUHL, and F. D. BALDINI. Variability in running economy and mechanics among trained male runners. *Med. Sci. Sports Exerc.* 23:378-383, 1991.
12. PATERSON, D. H., D. A. CUNNINGHAM, M. J. PLYLEY, C. J. R. BLIMKIE, and A. P. DONNER. The consistency of cardiac output measurement (CO_2 rebreath) in children during exercise. *Eur. J. Appl. Physiol.* 49:37-44, 1982.
13. SHROUT, P. E., and J. L. FLEISS. Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* 86:420-428, 1979.
14. TAYLOR, C. Some properties of maximal and submaximal exercise with reference to physiological variation and the measurement of exercise tolerance. *Am. J. Physiol.* 142:200-212, 1944.

15. TAYLOR, H. L., E. BUSKIRK, and A. HENSCHL. Maximal oxygen intake as an objective measure of cardio-respiratory performance. *J. Appl. Physiol.* 8: 1955.
16. VAN HERWAARDEN, C. L. A., R. A. BINKHORST, J. F. M. FENNIS, and A. VAN'T LAAR. Reliability of the cardiac output measurement with the indirect Fick-principle for CO₂ during exercise. *Pflügers Arch.* 385:21-23, 1980.
17. WILMORE, J. H. Influence of motivation on physical work capacity and performance. *J. Appl. Physiol.* 24:459-463, 1968.
18. WILMORE, J. H., P. A. FARRELL, A. C. NORTON, et al. An automated, indirect assessment of cardiac output during rest and exercise. *J. Appl. Physiol.* 52:1493-1497, 1982.
19. WOLFE, L. A., D. A. CUNNINGHAM, G. M. DAVIS, and P. A. RECHNITZER. Reliability of noninvasive methods for measuring cardiac function in exercise. *J. Appl. Physiol.* 44:55-58, 1978.
20. ZEIDIFARD, E., M. SILVERMAN, and S. GODFREY. Reproducibility of indirect (CO₂) Fick method for calculation of cardiac output. *J. Appl. Physiol.* 33:141-143, 1972.