Journal of Strength and Conditioning Research, 2007, 21(2), 378–384

Reliability and Detecting Change Following Short-Term Creatine Supplementation: Comparison of Two-Component Body Composition Methods

Liam P. Kilduff, Sarah Lewis, Mike I.C. Kingsley, Nick J. Owen, and Rebecca E. Dietzig

Department of Sports Science, Vivian Tower, University of Wales Swansea, Swansea, UK.

Abstract. Kilduff, L.P., S. Lewis, M.I.C. Kingsley, N.J. Owen, and R.E. Dietzig. Reliability and detecting change following short-term creatine supplementation: Comparison of two-component body composition methods. *J. Strength Cond. Res. 21(2): 378–384. 2007.*—The purpose of the present study was twofold: firstly, to assess the reliability of various body composition methods, and secondly, to determine the ability of the methods to estimate changes in fat-free mass (FFM) following creatine (Cr) supplementation. Fifty-five healthy male athletes (weight 78.3 ± 10.3 kg, age 21 ± 1 years) gave informed consent to participate in this study. Subjects' FFM was estimated by hydrostatic weighing (HW), air-displacement plethysmography (ADP), bioelectrical impedance analysis (BIA), near-infrared spectroscopy (NIR), and anthropometric measurements (ANTHRO). Measurements were taken on 2 occasions separated by 7 days to assess the reliability of the methods. Following this, 30 subjects returned to the laboratory for an additional test day following 7 days of Cr supplementation (20 g·d^{-1} Cr + 140 g·d^{-1} dextrose) to assess each method's ability to detect acute changes in FFM. In terms of reliability, we found excellent test-retest correlations for all 5 methods, ranging from 0.983 to 0.998 ($p < 0.001$). The mean biases for the 5 methods were close to 0 (range −0.1 to 0.3 kg) and their 95% limits of agreement (LOAs) were within acceptable limits (HW = −1.1 to 1.7 kg; ADP = −1.1 to 1.2 kg; BIA = −1.0 to 1.0 kg; NIR = −1.4 to 1.4 kg); however, the 95% LOAs were slightly wider for ANTHRO (−2.4 to 2.6 kg). Following Cr supplementation there was a significant increase in body mass (from 77.9 ± 10.1 kg to 78.9 ± 10.3 kg, $p = 0.000$). In addition, all 5 body composition techniques detected the change in FFM to a similar degree (mean change: HW = 0.9 ± 0.6 kg; ADP = 0.9 ± 0.6 kg; BIA = 0.9 ± 0.6 kg; NIR = 0.8 ± 0.5 kg; ANTHRO = 1.0 ± 0.7 kg; intraclass correlation coefficient = 0.962). We conclude that between-day differences in FFM estimation were within acceptable limits, with the possible exception of ANTHRO. In addition, all 5 methods provided similar measures of FFM change during acute Cr supplementation.

Key Words. Bland-Altman analysis, across-day reliability, creatine monohydrate

Introduction

Estimation of fat-free mass (FFM) has traditionally been performed using a wide range of methods, each differing with respect to assumptions, technical expertise, and cost (8). Many previous researchers in the area of body composition have been concerned with the validation of techniques, comparing relatively inexpensive techniques (e.g., bioelectrical impedance analysis [BIA]), with more complex criterion methods (e.g., hydrostatic weighing [HW]), and despite some researchers' inappropriate use of statistics, recent validation studies have extensively examined the between-method limits of agreement (11, 19).

However, a more fundamental question relating to whether techniques are reliable and sensitive to change in body composition remains unanswered. Recently, a number of papers have examined the ability of various body composition methods to detect decreases in subjects' fat mass (FM) following weight loss (5, 7, 25, 30). The primary stimulus for the majority of these papers would seem to come from 2 sources: (a) the increased prevalence of obesity and related diseases, which have risen dramatically in the last few years, and (b) increased pressures on athletes to make weight in various weight-dependent sports.

Similarly, there has been an increase in the prevalence of various diseases, such as chronic obstructive pulmonary disease, that directly or indirectly affect skeletal muscle mass. Importantly, the loss of skeletal muscle mass, for which FFM is a widely-used surrogate, has been associated with progressive disability, increased utilization of health care resources, and even mortality in many patient populations (24). Wide-ranging strategies have been employed in an attempt to restore muscle mass in various diseases (e.g., nutritional supplements). A recent study by Kristiansen and colleagues (22) reported that 98.6% of varsity athletes and 94.3% of controls used nutritional supplements, with a large proportion of these supplements (e.g., protein and creatine [Cr]) aimed at increasing FFM. Recently numerous body composition techniques have been used to quantify the changes in FFM in both athletes and disease populations; however, comparison across these studies is difficult because of the lack of data on the agreement between the various techniques' ability to detect changes in FFM.

Therefore, in light of the above, we were interested in determining the reliability and the ability of various body compositional techniques to detect acute increases in FFM following Cr supplementation. More specifically, we wished to: (a) establish the limits of agreement between 2 test days, 7 days apart, for all 5 techniques; and (b) evaluate whether all 5 techniques detect similar changes in FFM following acute Cr supplementation.

Methods

Experimental Approach to the Problem

Prior to entering the supplement phase of the study, subjects visited the laboratory on 2 occasions separated by 7 days (test day 1 [T1] and test day 2 [T2]) so that we could examine the reliability of all the techniques. Following the reliability period, 30 subjects performed an additional test day (test day 3 [T3]) following 7 days of Cr supplementation. The supplementation period for the groups

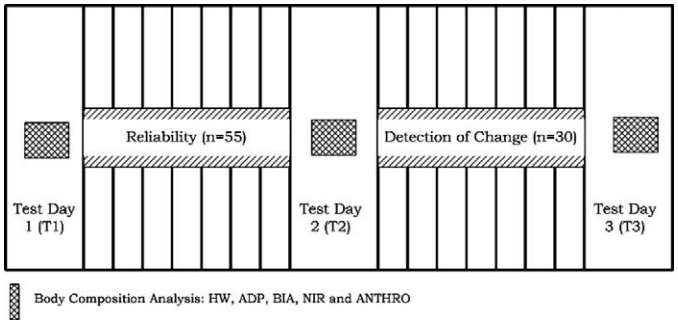

FIGURE 1. Experimental design. See text for details. T1 = test day 1; T2 = test day 2; HW = hydrostatic weighing; ADP = air-displacement plethysmography; BIA = bioelectrical impedance; NIR = near-infrared spectroscopy; ANTHRO = anthropometry.

started on the day after T2 and finished the day before T3. The experimental design is shown in Figure 1. The Cr supplementation consisted of 22.8 $g \cdot d^{-1}$ $Cr \cdot H_2O$ (equivalent to 5 g Cr 4 times daily) and 35 g of glucose polymer made up in 500 ml of warm to hot water for 7 days, taken at equal intervals throughout the day. Dissolving Cr in warm to hot water helps prevent any detectable formation of creatinine, and also helps minimize the parts of the supplement that might remain undissolved. Subjects were instructed to follow their normal diet, apart from the extra carbohydrate contained in the experimental drinks, and subjects were also requested to eliminate caffeine and caffeine-containing foods from their diets to minimize the possible inhibitory effects of caffeine on the ergogenic effect of Cr. Throughout the duration of the study subjects were encouraged to maintain their normal training habits. At the end of the study all subjects gave verbal assurance that they had complied with these instructions; in addition, all subjects were required to return the empty containers. Before any measurements were made, subjects were provided with written instructions outlining the testing protocol. The testing criteria included the following: (a) a 12-hour fast, (b) no intensive exercise within 24 hours of testing, (c) no consumption of alcohol 12 hours before testing, and (d) adequate hydration.

Subjects

Fifty-five healthy male university standard athletes (Table 1) gave informed consent to take part in the present study. The experimental procedures were in accordance with the policy statement of the American College of Sports Medicine and were approved by a University of Wales Swansea ethics committee. Subject eligibility was initially assessed by interview, whereby subjects were screened for good health and were excluded if they had a history of cardiovascular or respiratory disease or evidence of musculoskeletal injury. Subjects were recruited on the basis that they had not supplemented with Cr or with any other supplement that might alter body composition for 3 months prior to the beginning of the study.

Procedures

Subjects reported to the laboratory on the morning of testing after a standardized meal and after having refrained from alcohol intake, caffeine intake, and strenuous exercise the day before. Height was measured (to the nearest 1 cm) using a stadiometer (Holtain Ltd., Crymych, Dyfed, UK) with subjects standing barefoot. Body mass was assessed (to the nearest 0.1 kg) with subjects wearing only a swimsuit. Following the measurement of height and body mass, FFM (kg) was measured using 5 body compositional techniques: HW, air-displacement plethysmography (ADP), BIA, anthropometry (ANTHRO), and near-infrared spectroscopy (NIR). All body composition measurements were performed by the same investigator. Room temperature was maintained between 20 and 24°C.

TABLE 1. Physical characteristics of subjects at baseline (n = 55).

Variable	Mean ± *SD*
Weight (kg)	78.3 ± 10.3
Height (cm)	177.9 ± 7.7
Age (y)	20.6 ± 1.4
Body fat (%)	13.2 ± 3.9
Residual volume (L)	1.4 ± 0.3

Measurements

ANTHRO. Skinfold thickness was measured on the right side at appropriately marked sites (to the nearest 0.1 mm) using a Harpenden calliper (British Indicators Ltd., St. Albans, UK). Skinfold thickness was assessed at the chest, axilla, triceps, subscapular, abdomen, suprailium, and thigh, according to the standardized anatomical locations and methods reported by International Standards for Anthropometric Assessment. Three skinfold measurements were performed at each location, with a difference of no greater than 5% allowed between acceptable measures; the median of these 3 measurements was then used in all addition calculations. All skinfold measurements were performed by a Level 2 accredited anthropometrist with a technical error of measurement of 1.8%.

The appropriate anthropometric and demographic data (sum of 7 skinfolds, age, and sex) were entered into the Jackson and Pollock regression equation to determine body density (D_b) (18). Measurement of FFM by ANTHRO (FFM_{ANTHRO}) was calculated from D_b using Siri's equation (27).

BIA. Measurement of FFM by BIA (FFM_{BIA}) (Bodystat Quadscan, Bodystat Ltd., Douglas, UK) was performed on the right side while the subject was supine and with the limbs slightly apart from the trunk. After the skin had been cleaned with 70% alcohol, 2 injector electrodes were placed on the dorsal surfaces of the right hand and foot (across the distal ends of the metacarpals and metatarsals, respectively), and 2 detector electrodes were placed between the radius and ulna (on the wrist between the medial and lateral styloid process and on the ankle between the medial and lateral malleoli). The impedance to current flow (5 and 200 kHz) between the injector and detector electrodes was determined. This method allows total body water (TBW; impedance measured at 200 kHz) and extracellular water (ECW; impedance measured at 5 kHz) to be estimated; from these measurements, intracellular water (ICW) can also be deduced (TBW = ICW + ECW).

Before testing, the Quadscan BIA analyzer was calibrated according to the manufacturer's instructions by testing the actual resistance obtained at 50 kHz from the analyzer current being run through a 500-Ω calibration resistor. In all cases, the resistance was within the calibration specifications (495 to 505 Ω).

ADP. The principles of measurement of FFM by ADP

(FFM_{ADP}) have been explained in detail elsewhere (6). Briefly, when a subject sits inside an enclosed chamber of fixed volume, a volume of air equal to the subject's body volume is displaced. Because, under adiabatic conditions, changes in gas volume are inversely related to pressure raised to the power 1.4 (Poisson's law), the magnitude of pressure variations relative to those for a reference chamber provide the actual air volume displaced by the body.

Initially, a standard 2-point volume calibration (0 L, 50.004 L) of the air-displacement plethysmograph (BOD POD; Life Measurement, Inc., Concord, CA) was performed according to the manufacturer's instructions. The subjects (wearing swimsuits and swim caps) then entered the chamber (volume = 450 L) and sat quietly with an erect posture, breathing normally with hands folded in their laps and feet placed on the floor. Their uncorrected body volume (Vb_{uncorr}) was then measured over a 50-second period, with a minimum of 2 tests conducted. When 2 consecutive measurements of Vb_{uncorr} (L) were within 0.2% or 150 ml (whichever was the larger), the results were averaged. The corrected body volume (Vb_{corr} [L]) was then calculated, taking into account the confounding effect of isothermal air near the skin surface (surface area artifact [SAA] [L]) and the intrathoracic gas volume (Vtg) (L) (6):

$$Vb_{corr} = Vb_{uncorr} - SAA + 40\%\ Vtg$$

Two methods for estimating Vtg were used: (a) employing the standard values provided in the system software (estimated) (Vtg_{pred}), and (b) employing Vtg as measured via an internal breathing circuit according to the methods described by Dempster and Aitkens (6) (Vtg_{meas}). The D_b was then calculated as body mass/Vb, and % FFM was estimated from Siri's equation (27).

HW. The following measurements were made before HW: weight in air (kg), height (cm), air temperature (°C), and barometric pressure (mm Hg). Residual volume was estimated by oxygen dilution as previously described by Wilmore et al. (31). Subjects were given 2 dry runs for familiarization and then duplicate measures of residual volume were obtained (providing that the values were within 100 ml). The volume of gas in the intestine was estimated as 100 ml.

Following the measurement of residual volume the subjects entered the tank and removed any air trapped in their swimsuits, in their hair, or on their skin; they were instructed to immerse themselves in the water to chin level. After the subjects were completely submerged, they exhaled maximally. A load cell (Type F256TFROKN; Novatech Measurements Ltd., East Sussex, UK) reading was obtained after the subject signalled for the end of maximal exhalation and all air bubbles from the exhalation disappeared. This procedure was repeated between 5 and 8 times per subject. The mean of the 3 trials in which the subject had the greatest mass (provided that these values were within ± 100 g) was used to determine body density.

Total body fat and measurement of FFM by HW (FFM_{HW}) were calculated from Db, using Siri's equation (27).

NIR. Measurement via NIR (Futrex-6100/XL; Futurex, Inc., Gaithersburg, MD) involves a wand, emitting infrared light, being placed over the subject's biceps. The assumption is that the degree of infrared light absorbed and reflected is related to both the composition of the tissues through which the light is being passed and the specific wavelength being emitted by the light (14). Measurements were made in accordance with the manufacturer's instructions. Each subject was seated with his right arm extended (palm up) and resting comfortably on a table. The NIR wand was placed halfway between the antecubital fossa (of the elbow) and the axilla (armpit) with the light-emitting wand held perpendicular to the measurement site. Caution was used to make sure that the shield flaps were pressed on the skin so that no external light would penetrate and affect the optical density measurements. All measurements were performed by the same investigator and care was taken to ensure that the same amount of pressure was exerted on the wand for each subject.

Statistical Analyses

Descriptive statistics (mean and *SD*) were calculated for the group. Results were reported in terms of absolute FFM expressed in kg.

Reliability. Pearson's product moment correlation coefficients (r) between T1 and T2 for all methods were calculated. The limits of agreement (LOAs) between T1 and T2 for ADP, HW, NIR, BIA, and ANTHRO were investigated by plotting the individual between-day differences against their respective means (Bland-Altman plots) (3). Heteroscedasticity was examined by plotting the absolute (positive) differences against the individual means and calculating the correlation coefficient (3). If the heteroscedasticity correlation was close to 0 and the differences were normally distributed (Shapiro-Wilks test), the mean bias and 95% LOAs were calculated as mean ± 1.96 *SD* of the between-method difference (3).

Detecting Change. Paired t-tests were used to compare the FFM measurements obtained from all 5 methods on T1 compared to T2. A repeated-measures 1-way analysis of variance (ANOVA) was used to compare the change (Δ) detected by each method following Cr supplementation. In addition, an intraclass correlation and Pearson's product moment correlation coefficients were calculated to assess the relationship between the changes detected in FFM (kg) between all 5 methods and the changes in body mass following Cr supplementation as previously reported by researchers in this area (7, 30). All data were analyzed using SPSS software (version 11.0; SPSS, Inc., Chicago, IL). Significance was set a priori at $p \leq 0.05$.

RESULTS

Technical Consideration

We sought to investigate whether the use of Vtg_{meas} rather than Vtg_{pred} lung volumes would be of practical relevance for FFM estimation by ADP. Mean values for FFM (kg) as calculated by using Vtg_{meas} and Vtg_{pred} were 67.2 ± 7.3 and 67.9 ± 7.8, respectively ($p = 0.014$). In addition, systematically higher values for FFM_{ADP} resulted from the use of Vtg_{pred} compared to Vtg_{meas} (mean bias 0.7 kg). However, because of the difficulty associated with the breathing technique, 14 subjects failed to get a Vtg_{meas} value. We therefore chose to use the individual Vtg_{pred} values to calculate FFM_{ADP} for the ADP-based reliability and detecting change comparisons.

Reliability

The between-method differences did not present significant heteroscedasticity for this group ($p > 0.05$). Pearson correlations for FFM (kg) estimations between T1 and T2 for all 5 methods were high, ranging from 0.983 (ANTHRO) to 0.998 (BIA) ($p < 0.001$) (Table 2). The mean

TABLE 2. Reliability of the 5 methods (n = 55).*

Method	Mean bias (kg)	95% limits of agreement (kg)	Correlation
HW	0.3	−1.1 to 1.7	0.996
ADP	−0.1	−1.1 to 1.2	0.997
BIA	0.0	−1.0 to 1.0	0.998
NIR	0.0	−1.4 to 1.4	0.995
ANTHRO	0.1	−2.4 to 2.6	0.983

* HW = hydrostatic weighing; ADP = air-displacement plethysmography; BIA = bioelectrical impedance; NIR = near-infrared spectroscopy; ANTHRO = anthropometry.

bias ranged between −0.1 and 0.3 kg for the various methods (Table 2 and Figure 2a–e). The 95% LOAs were very similar between 4 of the methods, with the 95% LOAs being wider for ANTHRO (Table 2 and Figure 2e).

Detecting Change

Following supplementation, body mass increased significantly, from 77.9 ± 10.1 kg to 78.9 ± 10.3 kg (p = 0.000). In addition, Cr supplementation led to a significant increase in FFM as measured by all 5 methods ($p < 0.05$). A 1-way ANOVA indicated that there were no significant differences between the amount of change detected by any of the 5 methods, indicating that all 5 methods measured body composition changes to a similar magnitude in response to Cr supplementation (mean changes: HW = 0.9 ± 0.6 kg; ADP = 0.9 ± 0.6 kg; BIA = 0.9 ± 0.6 kg; NIR = 0.8 ± 0.5 kg; ANTHRO = 1.0 ± 0.7 kg) (Figure 3). In support of this, the changes in FFM as detected by all methods were significantly correlated (intraclass correlation coefficient = 0.962). Table 3 presents individual correlations to allow the reader to examine relationships between the various methods' abilities to detect change.

In addition, Table 4 provides a comparison between the 4 methods and HW for a cross-method comparison.

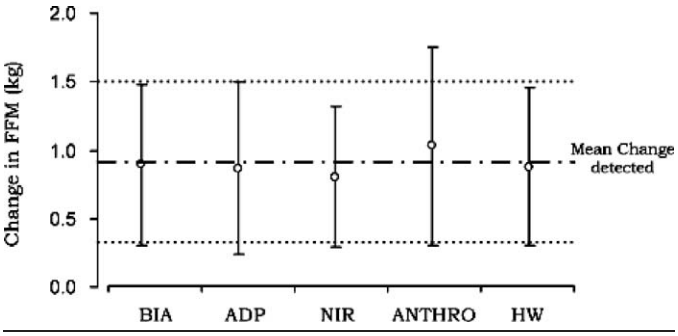

FIGURE 3. Mean difference (dots) and their standard deviation (*SD*; vertical lines) between fat-free mass (FFM) (kg) changes as calculated following creatine supplementation of the alternative methods. Horizontal line represents mean of the 5 methods. BIA = bioelectrical impedance; ADP = air-displacement plethysmography; NIR = near-infrared spectroscopy; ANTHRO = anthropometry; HW = hydrostatic weighing.

DISCUSSION

Accurate, reliable, and sensitive body composition assessment is beneficial for athletes and patients alike. The purpose of the study was to determine whether HW, ADP, BIA, NIR, and ANTHRO techniques could provide reliable measures of FFM over a 7-day period. In addition, this study aimed to examine the ability of the above-mentioned techniques to detect acute changes in FFM.

A consistent finding throughout the Cr literature is a significant increase in body mass following short-term Cr supplementation ranging from 0.6 to 1.8 kg (e.g., 12, 16, 20, 26). Considering the short time course of this increase in body weight, some investigators have attributed these increases to increases in TBW. For example, Hultman et al. (16) found a 0.6-L decline in urinary volume following acute Cr supplementation (20 g·d^{-1} for 6 days) and therefore attributed the increase in body weight to Cr-stimu-

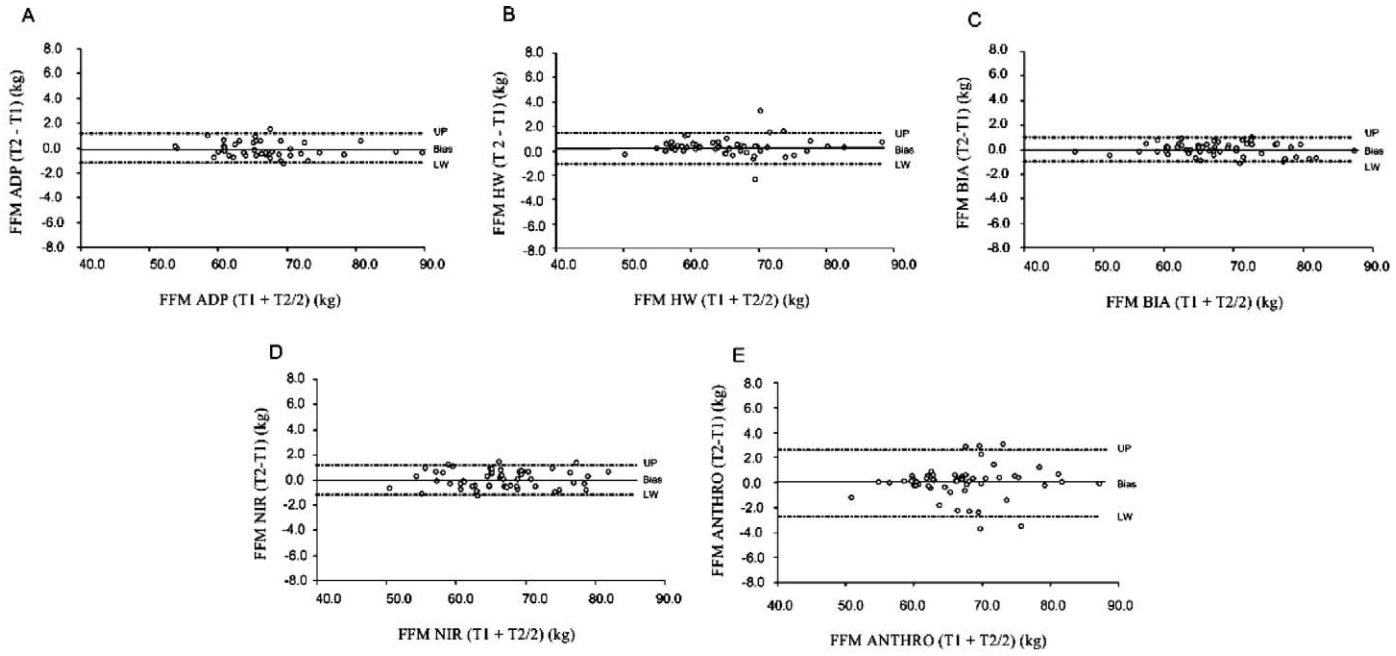

FIGURE 2. Absolute limits of agreement between test day 1 (T1) and test day 2 (T2) (separated by 7 days) for (A) fat-free mass air-displacement plethysmography (FFM ADP), (B) fat-free mass hydrostatic weighing (FFM HW), (C) fat-free mass bioelectrical impedance (FFM BIA), (D) fat-free mass near-infrared spectroscopy (FFM NIR), and (E) fat-free mass anthropometry (FFM ANTHRO). UP = upper 95% limit of agreement; LW = lower 95% limit of agreement.

TABLE 3. Correlation matrix between changes in body mass (BM) and changes detecting by all 5 body composition methods (n = 30).

	BM	BIA	ADP	NIR	ANTHRO	HW
BM	—	0.908†	0.957†	0.875†	0.930†	0.912†
BIA	0.908†	—	0.824†	0.876†	0.844†	0.915†
ADP	0.957†	0.824†	—	0.773†	0.909†	0.852†
NIR	0.875†	0.876†	0.773†	—	0.850†	0.794†
ANTHRO	0.930†	0.844†	0.909†	0.850†	—	0.835†
HW	0.912†	0.915†	0.852†	0.794†	0.835†	—

* BIA = bioelectrical impedance; ADP = air-displacement plethysmography; NIR = near-infrared spectroscopy; ANTHRO = anthropometry; HW = hydrostatic weighing.

† Correlation significant at the 0.01 level.

lated water retention. These authors also noted that the time course of urinary volume changes paralleled that of muscle Cr uptake. Furthermore, an increase in both TBW and intracellular water were demonstrated by Zeigenfuss et al. (32) with acute Cr ingestion (0.35 g·kg $FFM^{-1}\cdot d^{-1}$ for 3 days), they reported that the increase in TBW accounted for approximately 90% of the acute gain in body mass. It remains to be determined whether this increase in water is associated with an increase in protein synthesis (13). The balance of available evidence from human performance studies using Cr supplementation and more direct evidence from animal in vivo and in vitro experiments would support the notion that increasing Cr availability may indeed increase protein synthesis (17). In the present study, supplementation with Cr increased body weight by 1.0 ± 0.7 kg and significantly increased TBW and ICW, as assessed by BIA. However, the increase in body weight cannot be fully explained by the increase in TBW alone as a result of Cr-stimulated water retention, as there was no significant increase in TBW as expressed as percentage of body weight. Despite a significant increase in TBW and ICW in the Cr group expressed in absolute terms (46.2 ± 4.4 L to 47.1 ± 4.7 L, p = 0.003, and 27.1 ± 2.7 L to 28.0 ± 3.0 L, p = 0.006, respectively), if the relative volume of TBW remains constant (as in the present study), the gain in body mass may not be attributed to water retention. The increase in absolute TBW seen in this study following Cr supplementation may be indicative of intracellular water that normally accompanies dry matter growth. Previous researchers have found similar results and interpreted their findings in the same manner (20, 26).

In the present study, 4 of the techniques (HW, ADP, BIA, and NIR) were found to be highly reliable for measuring FFM (kg) on 2 separate test days separated by 7 days (Figure 2a–d); however, ANTHRO produced wider LOA compared to the other 4 methods (Figure 2e). The test-retest reliability of ADP for the measurement of FFM was good based on a high correlation, coupled with the low mean bias and narrow LOA (Table 2 and Figure 2a).

TABLE 4. Validity of the methods compared to hydrostatic weighing (HW) (n = 55).*

Method	Mean bias (kg)	95% limits of agreement (kg)	Correlation
HW-ADP	1.8	−2.8 to 6.4	0.951
HW-BIA	2.3	−3.7 to 8.3	0.918
HW-NIR	0.7	−5.2 to 6.6	0.919
HW-ANTHRO	1.5	−4.1 to 7.1	0.927

* ADP = air-displacement plethysmography; BIA = bioelectrical impedance; NIR = near-infrared spectroscopy; ANTHRO = anthropometry.

The results of the present study are in agreement with previous studies testing reliability of ADP in various populations (2, 4, 23). However, the majority of previous studies examining the reliability of ADP have focused on within-day reliability. This means that their results cannot be directly related to intervention studies in which repeat tests are performed on separate days. In order to ensure good test-retest data (reliability) during the collection of any data in a research setting, both main components of measurement error (systematic bias and random error) need to be minimized (1). Systematic bias (e.g., general learning) is minimal with the majority of body composition techniques and in terms of random error (incorporates biological or mechanical errors); the present study ensured that all external factors concerned with all the different methods were controlled to the greatest extent possible. For example, ADP has a number of factors that need to be controlled on repeat visits in order to increase ADP reliability. Because of possible effects of varying amounts of isothermal air trapped in hair or clothing, researchers should control the following factors: (a) a swim cap must be worn at all times (15); (b) minimal clothing should be worn and clothing should be kept consistent on each test day (10); (c) body temperature and air moisture need to be kept constant on repeat visits (9).

Recent studies have reported excellent test-retest correlations for the within-day reliability of HW (23). McCrory et al. (23) examined the within-day reliability of HW in 68 healthy subjects (26 female, 42 male) and reported a between-trial coefficient of variation (CV) of 2.3 ± 1.9% for HW. The results of the present study have also shown HW to have excellent reliability when measured over a period of 7 days (Table 2 and Figure 2b). Fornetti et al. (11) examined the within-day reliability of NIR and BIA in 132 female athletes. The results of this study showed excellent reliability coefficients for both multiple- and single-trial measurements (r = 0.994 for BIA and 0.957 for NIR). However, as stated previously, these high reliability coefficients were likely a function of the short interval between trials and thus of reducing random error. In light of this, their findings cannot be directly related to many intervention studies in which repeat measurements were taken over a longer time course (e.g., tests on multiple days). Data from the present study develop our knowledge with regard to the across-day reliability of HW, NIR, and BIA (Table 2 and Figure 2b–d).

In terms of the reliability of ANTHRO, although we report a high correlation coefficient of 0.975 and a mean bias of 0.1 kg, there were relatively large 95% LOAs found (Figure 2e). Klipstein-Grobusch et al. (21) examined the reliability of anthropometric measurements and reported a correlation coefficient of 0.997 and a CV of 2.27%.

As previously mentioned, the majority of studies ex-

amining various body composition methods' ability to detect change have focused on measuring alterations in FM as a result of various nutritional and exercise interventions (7, 25, 30). However, it is equally important for sports scientists and clinicians alike to be able to (a) know the ability of the various body composition methods to accurately assess the changes in FFM following training or nutritional intervention; and (b) quantify the level of agreement between the various methods in estimating the level of change in order to equally make comparisons between intervention studies. In the present study, all 5 methods assessed the changes in FFM to a similar level following acute Cr supplementation (Figure 3). Recently, 2 studies have examined the ability of various body compositional methods to detect increases in FFM following nutritional/pharmacological intervention (28, 29). For example, Van Marken Lichtenbelt and colleagues (28) examined the ability of various body composition methods (e.g., 4-C model, dual-energy X-ray absorptiometry, HW, and BIA) to detect changes in FM and FFM in 27 male bodybuilders following a period of exercise and androgenic-anabolic steroids. In this study the authors used the 4-C model as the gold standard for detecting changes following the intervention and concluded that only the 3-Cw model (incorporating body density as measured from HW and TBW as measured by deuterium dilution) could serve as an alternative for the 4-C model if accurate measurements of body composition change are needed.

In recent years, HW has been used as a criterion method against which to compare all other methods in terms of validation. To date, there is no agreed criterion method for assessing change following any intervention. Therefore, although all methods in the present study detected a similar change in FFM, we cannot say which method detected the change more accurately, only that all methods detected similar change.

In light of the above, the findings of this study indicate that HW, ADP, BIA, and NIR have excellent test-retest reliability, whereas ANTHRO test-retest results indicate more variability, as indicated by the wider LOA. Furthermore, this study indicates that all 5 methods incorporated in the present study detected similar changes in FFM following supplementation.

Practical Applications

The findings of the present study demonstrated that ADP, HW, BIA, and NIR are reliable methods when performed under strict standardized conditions. In addition, all techniques quantified acute changes in FFM following Cr supplementation to a similar level and therefore results from various techniques may be compared. As for coaches and athletic trainers, this study shows that, if correct measurement procedures are followed, relatively inexpensive techniques such as BIA, NIR, and ANTHRO can be as reliable and sensitive to change as more expensive techniques such as ADP and HW.

References

1. ATKINSON, G., AND A.M. NEVILL. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med.* 26:217–238. 1998.
2. BALLARD, T.P., L. FAFARA, AND M.D. VUKOVICH. Comparison of Bod Pod and DXA in female collegiate athletes. *Med. Sci. Sports Exerc.* 36:731–735. 2004.
3. BLAND, J.M., AND D.G. ALTMAN. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–310. 1986.
4. COLLINS, M.A., M.L. MILLARD-STAFFORD, P.B. SPARLING, T.K. SNOW, L.B. ROSSKOPT, S.A. WEBB, AND J. OMER. Evaluation of the BOD POD for assessing body fat in collegiate football players. *Med. Sci. Sports Exerc.* 31:1350–1356. 1999.
5. COX-REIJVEN, P.L.M., B. VAN KREEL, AND P.B. SOETERS. Accuracy of bioelectrical impedance spectroscopy in measuring changes in body composition during severe weight loss. *JPEN J. Parenter. Enteral Nutr.* 26: 120–127. 2001.
6. DEMPSTER, P., AND S. AITKENS. A new air displacement method for the determination of human body composition. *Med. Sci. Sports Exerc.* 27: 1692–1697. 1995.
7. ELBERG, J., J.R. MCDUFFLE, N.G. SEBRING, C. SALAITA, M. KEIL, D. ROBOTHAM, J.C. REYNOLDS, AND J.A. YANOVSKI. Comparison of methods to assess change in children's body composition. *Am. J. Clin. Nutr.* 80:64–69. 2004.
8. ELLIS, K.J. Human body composition: In vivo methods. *Physiol. Rev.* 80: 649–680. 2000.
9. FIELDS, D.A., P.B. HIGGINS, AND G.R. HUNTER. Assessment of body composition by air-displacement plethysmography: Influence of body temperature and moisture. *Dyn Med.* 3:3. 2004.
10. FIELDS, D.A., G.R. HUNTER, AND M.I. GORAN. Validation of the BOD POD with hydrostatic weighing: Influence of body clothing. *Int. J. Obes. Relat. Metab. Disord.* 24:200–205. 2000.
11. FORNETTI, W.C., J.M. PIVARNIK, J.M. FOLEY, AND J.J. FIECHTNER. Reliability and validity of body composition measures in female athletes. *J. Appl. Physiol.* 87:1114–1122. 1999.
12. GREENHAFF, P.L., K. BODIN, K. SODERLUND, AND E. HULTMAN. Effect of oral creatine supplementation on skeletal muscle phosphocreatine resynthesis. *Am. J. Physiol.* 266:E725–E730. 1994.
13. HAUSSINGER, D., E.R. ROTH, F. LANG, AND W. GEROK. Cellular hydration state: An important determinant of protein catabolism in health and disease. *Lancet* 341:1330–1332. 1993.
14. HEYWARD, V.H., K.A. JENKINS, K.L. COOK, V.L. HICKS, J.A. QUATROCHI, W.L. WILSON, AND S.B. GOING. Validity of single-site and multi-site models for estimating body composition of women using near-infrared interactance. *Am. J. Hum. Biol.* 4:579–593. 1992.
15. HIGGINS, P.B., D.A. FIELDS, G.R. HUNTER, AND B.A. GOWER. Effect of scalp and facial hair on air displacement plethysmography estimates of percentage of body fat. *Obes. Res.* 9:326–330. 2001.
16. HULTMAN, E., K. SODERLUND, J.A. TIMMONS, G. CEDERBLAD, AND P.L. GREENHAFF. Muscle creatine loading in men. *J. Appl. Physiol.* 81:232–237. 1996.
17. INGWALL, J.S., C.D. WEINER, M.F. MORALES, E. DAVIES, AND F.E. STOCKDALE. Specificity of creatine in the control of muscle protein synthesis. *J. Cell Biol.* 63:145–151. 1974.
18. JACKSON, A.S., AND M.L. POLLOCK. Generalized equations for predicting body density of men. *Br. J. Nutr.* 40:497–504. 1978.
19. KILDUFF, L.P., J.P. FULD, J.A. NEDER, Y.P. PITSILADIS, R. CARTER, R. STEVENSON, AND S.A. WARD. Clinical relevance of inter-method differences in fat-free mass estimation in chronic obstructive pulmonary disease. *Respiration* 70:585–593. 2003.
20. KILDUFF, L.P., P. VIDAKOVIC, G. CONNEY, R. TWYCROSS-EWIS, P. AMUNA, M. PARKER, L. PAUL, AND Y.P. PITSILADIS. Effects of creatine on isometric bench-press performance in resistance-trained humans. *Med. Sci. Sports Exerc.* 34:1176–1183. 2002.
21. KLIPSTEIN-GROBUSCH, K., T. GEORG, AND H. BOEING. Interviewer variability in anthropometric measurements and estimates of body composition. *Int. J. Epidemiol.* 26:S174–S180. 1997.
22. KRISTIANSEN, M., R. LEVY-MILNE, S. BARR, AND A. FLINT. Dietary supplement use by varsity athletes at a Canadian university. *Int. J. Sport Nutr. Exerc. Metab.* 15:195–210. 2005.
23. MCCRORY, M.A., T.D. GOMEZ, E.M. BERNAUER, AND P.A. MOLE. Evaluation of a new air displacement plethysmograph for measuring human body composition. *Med. Sci. Sports Exerc.* 27:1686–1691. 1995.
24. MOSTERT, R., A. GORIS, C. WELING-SCHEEPERS, E.F. WOUTERS, AND A.M. SCHOLS. Tissue depletion and health related quality of life in patients with chronic obstructive pulmonary disease. *Respir. Med.* 94:859–867. 2000.
25. NELSON, M.E., M.A. FIATARONE, J.E. LAYNE, I. TRICE, C.D. ECONOMOS, R.A. FIELDING, R. MA, R.N. PIERSON, AND W.J. EVANS. Analysis of body-composition techniques and models for detecting change in soft tissue with strength training. *Am. J. Clin Nutr.* 63:678–686. 1996.
26. SAAB, G., G.D. MARSH, M.A. CASSELMAN, AND R.T. THOMPSON. Changes in human muscle transverse relaxation following short-term creatine supplementation. *Exp. Physiol.* 87:383–389. 2002.
27. SIRI, W.E. Body composition from fluid spaces and density: Analysis of methods. In: *Techniques for Measuring Body Composition.* J. Brozek and A. Henschel, eds. Washington, DC: NAS/NRC, 1961. pp. 223–224.
28. VAN MARKEN LICHTENBELT, W.D., F. HARTGENS, N.B.J. VOLLAARD, S. EBBING, AND H. KUIPERS. Body composition changes in bodybuilders: A method comparison. *Med. Sci. Sports Exerc.* 36:490–497. 2004.

29. VUKOVICH, M.D., AND B.M. PEETERS. Reliability of air-displacement plethysmography in detecting body composition changes after water ingestion and after creatine supplementation. *J. Exerc. Physiol. Online* 6: 115–122. 2003.
30. WEYERS, A.M., S.A. MAZZETTI, D.M. LOVE, A.L. GOMEZ, W.J. KRAEMER, AND J.S. VOLEK. Comparison of methods for assessing body composition changes during weight loss. *Med. Sci. Sports Exerc.* 34:497–502. 2002.
31. WILMORE, J.H., P.A. VODAK, R.B. PARR, R.N. GIRANDOLA, AND J.E. BILLING. Further simplification of a method for determination of residual lung volume. *Med. Sci. Sports Exerc.* 12:216–218. 1980.
32. ZEIGENFUSS, T.N., L.M. LOWERY, AND P.W.R. LEMON. Acute fluid volume changes in men during three days of creatine supplementation. *J. Exerc. Physiol. Online* 3:1–8. 1998.

Acknowledgments

The authors acknowledge Sarah Pearce, Nerys Renowden, Emma Ozolins, and David Young for their excellent technical assistance. The cooperation of the subjects is greatly appreciated.

Address correspondence to Dr. Liam P. Kilduff, L.Kilduff@swansea.ac.uk.