

DXA Provides a Valid Minimum Weight in Wrestlers

R. RANDALL CLARK¹, JUDE C. SULLIVAN¹, CYNTHIA J. BARTOK², and AARON L. CARREL³

¹University of Wisconsin Hospital Sports Medicine Center, University of Wisconsin, Madison, WI; ²Center for Childhood Obesity Research, The Pennsylvania State University, University Park, PA; and ³University of Wisconsin School of Medicine and Public Health, Department of Pediatrics, University of Wisconsin, Madison, WI

ABSTRACT

CLARK, R. R., J. C. SULLIVAN, C. J. BARTOK, and A. L. CARREL. DXA Provides a Valid Minimum Weight in Wrestlers. *Med. Sci. Sports Exerc.*, Vol. 39, No. 11, pp. 2069–2075, 2007. **Purpose:** To cross-validate the DXA prediction of minimum weight (MW) in high school wrestlers, using a criterion-referenced analysis. The goal was to independently evaluate whether DXA provided a MW within acceptable limits for the sport of wrestling. Secondly, the DXA prediction error was compared against the currently approved skinfold (SF) method. **Methods:** Criterion MW was calculated by hydrostatic weighing (HW) with measured residual lung volume. Whole-body scans were performed with a Norland XR-36 bone densitometer. All skinfolds were taken by the same experienced measurer. The subject's body density was computed by Lohman and was converted to percent body fat, using the equation of Brozek et al. The measured fat-free mass was used to calculate each wrestler's MW at 7% body fat. Subjects were 94 Wisconsin high school wrestlers (mean (SD): age = 15.1 yr (1.2), height = 170.3 cm (7.1), weight = 63.2 kg (9.6). **Results:** There was no significant difference in mean MW from DXA (60.6 kg (9.0)) and the HW criterion (59.8 kg (9.0)). The correlation was strong ($r = 0.98$), and the regression for the relationship between HW and DXA ($y = 0.976 \times \text{DXA} + 0.698 \text{ kg}$) did not significantly deviate from the line of identity. A low standard error of estimate (SEE) of 1.7 kg and a pure error (PE) of 1.9 kg were found, with residuals ranging from -3.94 to 2.88 kg. This PE was similar to the SF method (2.1 kg) in the sample. Bland–Altman analysis showed no systematic bias in the prediction of MW across weight classes. **Conclusion:** We conclude that DXA provided a valid prediction of MW in this sample of high school wrestlers. **Key Words:** BODY COMPOSITION METHODS, HYDROSTATIC WEIGHING, DUAL-ENERGY X-RAY ABSORPTIOMETRY, MINIMUM WRESTLING WEIGHT, SKINFOLDS

With a goal to reduce unhealthy weight loss practices and increase safe participation in the sport, the Wisconsin Interscholastic Athletic Association (WIAA) implemented a weight certification program requiring minimum weight (MW) testing for high school wrestlers in 1989. Using the guidelines of the American College of Sports Medicine (ACSM) (22) and the American Medical Association (2), the WIAA requires assessment of body composition for each wrestler before the competitive season (10). The measured fat-free mass (FFM) is used to calculate each wrestler's MW at 7% body fat (%BF).

Since then, Wisconsin's program has served as a model for several other states, and the National Collegiate Athletic Association (NCAA) program adopted in 1998. More recently, the National Federation of State High School Associations (NFHS) required that all state high school athletic associations implement a MW program covering the approximately 250,000 high school wrestlers in the United States. The new rules require measurement of body composition in all participants, but the rule does not specify which method(s) are acceptable. As a result, several methods have been used to predict body fat and MW in wrestlers. These include skinfolds (SF) (8,9,11,16,24,25,26, 28,29), bioelectrical impedance analysis (BIA) (7–9,24,31), hydrostatic weighing (HW) (8,11,16,24,25,26,29–31), near-infrared interactance (9), air displacement pletismography (30), and dual x-ray absorptiometry (DXA) (8,9). Results are equivocal, and, more importantly, conclusions regarding the acceptability for use in wrestlers vary by authors. Because of concerns regarding marketing, corporate partnerships, conflicts of interest, and company-funded research, the authors of this study urge caution and encourage states to carefully scrutinize studies and their conclusions before methods are accepted. Lesser et al. (18) suggest that industry

Address for correspondence: R. Randall Clark, University of Wisconsin Hospital Sports Medicine Center, University of Wisconsin Research Park, 621 Science Drive, Madison, WI 53711; E-mail: rr.clark@hosp.wisc.edu. Submitted for publication January 2007.

Accepted for publication July 2007.

0195-9131/07/3911-2069/0

MEDICINE & SCIENCE IN SPORTS & EXERCISE®

Copyright © 2007 by the American College of Sports Medicine

DOI: 10.1249/mss.0b013e31814fb423

funding of nutrition-related scientific articles may bias conclusions in favor of sponsors' products, with potentially significant implications for public health. Therefore, the validity must be thoroughly and independently evaluated as each state will determine an appropriate method for assessing body composition.

Ease of use cannot take precedence over validity. The assumption that any measure is better than no measure has inherent danger. An invalid MW is a potential health risk and, from both a health and competitive fairness standpoint, is problematic. An underestimation of MW suggests that an athlete could safely lose more weight. This could provide a false sense of safety and could potentially lead to the weight cutting behaviors that MW testing was meant to deter. An overestimation of MW would place the wrestler in an inappropriate weight class, thus removing the competitive fairness of weight classes. Both an under- and over-prediction of MW are medical (and potentially legal) concerns, because protecting the health and safety of the student athlete is the goal of the WIAA, NCAA, and NFHS rules.

DXA, considered the reference measure of bone density, is also used to assess the soft-tissue masses of the body, including nonbone lean (muscle) and fat (8,9,12,13,15,17). The technique is based on the differential attenuation of two transmitted energy levels as they pass through bone and soft tissues. Because DXA is less dependent on the assumptions critical to other methods (13,15,17), it has been embraced as an improved methodology for body composition assessment. Therefore, we hypothesized that the MW predicted by DXA would be similar to the MW provided by the HW criterion in this sample of wrestlers. To evaluate this hypothesis, we cross-validated the DXA prediction of MW in high school wrestlers, using a criterion-referenced analysis.

The DXA method would be deemed valid if the following conditions were met: the predicted mean was not significantly different from the criterion, the predicted MW was highly correlated to HW, and the regression line was similar to the line of identity. The standard error of the estimate (SEE) for the sample (19,20) would be computed to determine the 68.3 and 95.4% confidence intervals for the prediction. Systematic bias would be identified by the comparison of clinical methods suggested by Altman and Bland (1) to evaluate the method across the various weight classes, previously suggested with wrestlers (7,8,30,31). The purpose was to evaluate whether DXA provided an MW within acceptable limits for the sport of wrestling. Secondly, the DXA prediction error was compared against the currently approved SF method.

METHODS

Subjects were 94 volunteer high school wrestlers from the state of Wisconsin (Table 1). All measurements were taken during a single early morning testing session follow-

TABLE 1. Characteristics of subjects ($N = 94$).

	Mean \pm SD	Range
Age (yr)	16.1 \pm 1.2	14–18
Body weight (kg)	63.2 \pm 9.6	34.5–91.2
Height (cm)	173.3 \pm 7.1	146.1–185.4
Body density (HW)	1.0723 \pm 0.0107	1.0415–1.0964
Percent fat (HW)	11.8 \pm 4.3	3.6–24.6
Percent fat (DXA)	10.9 \pm 4.2	3.4–23.3
Percent fat (SF)	11.4 \pm 2.8	5.3–19.6
MW (kg) (HW)	59.8 \pm 9.0	33.5–85.5
MW (kg) (DXA)	60.6 \pm 9.0	31.9–86.8
MW (kg) (SF)	60.1 \pm 8.1	34.0–83.3

ing a defined preparticipation protocol. The preparticipation protocol was designed to encourage normal hydration before body composition analysis. This protocol prohibited food intake, physical activity, and ingestion of caffeine or alcohol for 12 h before the start of the body composition assessment. Subjects were encouraged to consume adequate fluids to maintain euhydration.

At 0800 h on test day, all measurements were taken under standardized conditions by the same investigators. All participants were asked to void and defecate before beginning the procedures. Urine specific gravity (USG) testing confirmed that subjects met the WIAA requirement for hydration before MW assessment (USG $< 1.025 \text{ g}\cdot\text{mL}^{-1}$). Height was measured with a wall-mounted stadiometer to the nearest 0.5 cm. Weight was measured on a calibrated beam-balance platform scale to the nearest 0.1 kg, with the subject wearing swim trunks. The procedures were approved by the clinical science center human subjects committee at the University of Wisconsin, and parental informed written consent for minors and minor assent were obtained before initiating the testing protocols.

Densitometry. Body density (BD) was determined by HW at residual volume (RV), as described by Behnke and Wilmore (5). Eight to ten trials were obtained for each subject. The mean of the three heaviest trials was used as the underwater weight for calculating BD. RV was measured by the oxygen dilution method described by Wilmore (33), using a modified Collins 13.5-L respirometer (Braintree, MA) and a Med Science model 505 N2 analyzer (St. Louis, MO). The subject's RV was measured outside the tank in a seated position simulating that used during HW. The mean of two trials within 75 mL was used as the RV. An additional correction of 100 mL, to account for gastrointestinal gas, was used in the HW calculation (5). %BF was estimated from BD, using the equation of Brozek et al. (6): $\%BF = 100 ((4.57/BD) - 4.142)$.

DXA. Whole-body scans were performed using the Norland XR-36 whole-body bone densitometer (Norland Corporation, Ft. Atkinson, WI), and tissue masses were analyzed using software version 3.7.4/2.1.0. The XR-36 x-ray tube operates at 100 kV and uses dynamic samarium filtration (K-edge at 46.8 keV) to produce energy peaks at a maximum of 40 and 80 keV. The XR-36 uses dynamic filtration to minimize beam hardening (13). Dual NaI detectors measure the attenuated x-ray using a pixel size

of 6.5×13.0 mm and a scan speed of $260 \text{ mm} \cdot \text{s}^{-1}$. All subjects were positioned in the supine position and were scanned by the same investigator. Subjects removed metal objects or clothing containing metal components and wore only workout shorts and a T-shirt for the scan procedure. Each scan session was preceded by a calibration routine, using multiple quality-control phantoms that simulate soft tissue and bone. On the basis of 18 scans of six subjects, using the XR-36 whole-body procedures, the total body coefficients of variation (CV) are as follows: soft-tissue mass 0.2%, total body mass 0.2%, lean body mass 1.0%, fat mass 2.5%, percent fat 2.4%, and total BMC 0.9%. The XR-36 uses the following calibration standards for determination of bone, fat, and lean mass from x-ray attenuation: bone – hydroxyapatite, fat – steric acid, lean – 0.6% NaCl in H_2O .

Skinfold measures. SF sites were measured with a Lange caliper (Beta Technology Inc., Cambridge, MD). All measurements were taken on the right side of the body in serial fashion by the same experienced investigator. SF thickness was based on the median of three trials, according to the WIAA and NCAA protocol (21). Reliability of the investigator has been previously documented (23). Test–retest reliability of the skinfolds in this sample were $r = 0.96$, 0.94 , and 0.92 for the triceps, subscapula, and abdominal sites, respectively. BD was predicted by the Lohman equation (20). %BF was estimated from BD using the equation of Brozek et al. (6), and MW was calculated at 7%BF (fat-free body/0.93), following the WIAA and NFHS guidelines.

Statistical analysis. Comparison of methods included the criterion-referenced cross-validation approach suggested by Lohman (20), Guo and Chumlea (14), and Altman and Bland (1). Validity was evaluated by testing the hypothesis that the regression using the criterion MW (y -axis) and the DXA prediction method (x -axis) was not significantly different from the line of identity in an independent cross-validation sample (14). A t -test was used to compare means and to determine whether the slope of the regression line for MW from DXA and HW was different than 1 and whether the intercept was different than 0.

The standard error of estimate (SEE) where $\text{SEE} = \text{SDy} \sqrt{1 - r^2}$ was used to estimate the 68.3 and 95.4% confidence intervals of the prediction. To test the hypothesis that the distribution of the residual scores (criterion – prediction) form a “normal” distribution (bell-shaped curve), we performed a formal hypothesis test: H_0 : data are normally distributed versus H_A : data are not normally distributed. We used the method described by Shapiro and Wilk (25) to perform this test.

The pure error (PE), as described by Guo and Chumlea (14), where $\text{PE} = \sqrt{\sum(Y_1 - Y_2)^2/N}$ when Y_1 = the predicted value and Y_2 = the criterion value was also calculated for this independent cross-validation sample. Lohman (20) has previously described this value as total error. Significant differences in PE values were tested for using the F_{\max}

(variance comparison) test. For this test, the ratio of highest PE^2 to lowest PE^2 was calculated, and the ratio's significance was determined in reference to the F distribution (3). Statistical significance was set at $P < 0.05$ for all tests.

Systematic bias is defined as the mean difference between the two measures (i.e., mean DXA – mean HW). If this difference is not significantly different than 0 (paired t -test), it would indicate that there is no systematic bias between the two measures. If, however, it differs from 0, it would indicate the presence of a systematic bias—that is, one measure consistently produces higher (or lower) values than the other measure. To illustrate systematic bias between DXA and HW, the difference between the methods (y -axis) was plotted as a function of the mean (x -axis), according to Altman and Bland (1). This has been suggested in previous studies with wrestlers (7,8,30,31) because it allows examination of the prediction across the various weight classes represented by the study sample.

RESULTS

There was no significant difference in mean MW from DXA (60.6 ± 9.0 kg) and the HW criterion (59.8 ± 9.0 kg). The correlation was strong ($r = 0.98$), and the regression for the relationship between HW and DXA ($y = 0.976\text{DXA} + 0.698$ kg) was similar to the line of identity (Fig. 1), with the slope not significantly different than 1 ($P = 0.23$), and the intercept not significantly different than 0 ($P = 0.57$).

Bland–Altman analysis showed no bias in the prediction of MW across the weight classes (Fig. 2). A nonsignificant correlation ($r = -0.03$), with a slope not significantly different than 0 ($P = 0.78$) and an intercept not significantly different than 0 ($P = 0.73$), were found. The mean

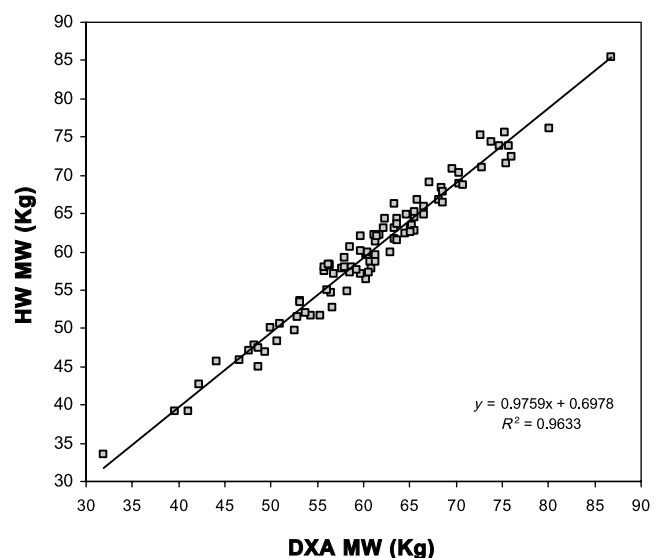


FIGURE 1—The HW criterion is the dependent variable (y -axis), and DXA is the independent variable (x -axis). The regression line between HW and DXA ($y = 0.976 \text{ DXA} + 0.698$ kg) was similar to the line of identity, with the slope not significantly different than 1 ($P = 0.23$) and the intercept not significantly different than 0 ($P = 0.57$).

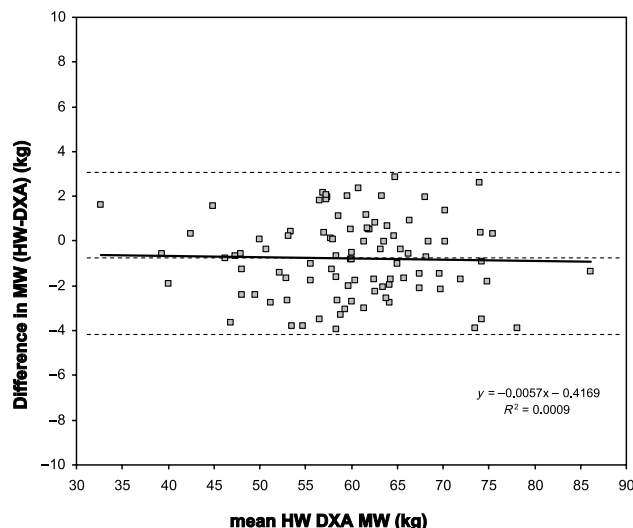


FIGURE 2—No bias across weight classes is seen in the DXA method when plotting the differences between the methods (y-axis) as a function of the mean (x-axis) according to Altman and Bland. A nonsignificant correlation ($r = -0.03$), with a slope not significantly different than 0 ($P = 0.78$) and an intercept not significantly different than 0 ($P = 0.73$), was found. The mean difference (-0.76 kg) between the methods and ± 2 SD (± 3.48 kg) are indicated.

difference (-0.76 kg) between the methods and ± 2 SD (± 3.48 kg) are indicated in Figure 2. The DXA SEE value was 1.7 kg and the PE value was 1.9 kg (Table 2), with residuals ranging from -3.94 to 2.88 kg. The distribution of the residual scores fit a normal distribution, with the P value for the Shapiro–Wilk test (25) for DXA of $P = 0.12$. Thus defining the SEE value as ± 1 SD of the residual scores (or the difference between predicted and criterion scores), DXA predicted MW within ± 1.7 kg (3.7 lb) of the criterion 68.3% of the time and 3.4 kg (7.4 lb) 95.4% of the time in this sample.

There also was no significant difference in mean MW from the currently approved Lohman SF method (20) (60.1 ± 8.1 kg) and the HW criterion (59.8 ± 9.0 kg). The correlation was strong ($r = 0.97$), and the regression for the relationship between HW and SF ($y = 1.0782\text{SF} - 5.0079$ kg) differed from the line of identity, with the slope significantly different than 1 ($P = 0.002$) and the intercept significantly different than 0 ($P = 0.002$). The SEE value for SF was 2.0 kg. The distribution of the residual scores fit a normal distribution, with the P value for the Shapiro–Wilk test (25) for SF of $P = 0.43$. Therefore, SF predicted MW within ± 2.0 kg (4.4 lb) of the criterion 68.3% of the time

TABLE 2. Mean \pm SD percent body fat (%BF), minimum weight (MW), correlation (r), standard error of estimate (SEE), and the pure error (PE).

Method	%BF	MW	r	SEE	PE
HW	11.8 ± 4.3	59.8 ± 9.0			
DXA	10.9 ± 4.2	60.6 ± 9.0	0.98	1.7	1.9
SF	11.4 ± 2.8	60.1 ± 8.1	0.97	2.0	2.1

Mean \pm SD values for percent body fat (%BF) and minimum weight (MW). The correlation, the standard error of estimate (SEE), and the pure error (PE) are used to compare the MW from the HW criterion (HW) with those from dual-energy x-ray absorptiometry (DXA) and skinfolds (SF).

and 4.0 kg (8.8 lb) 95.4% of the time in this sample. The SF method (20) yielded a PE of 2.1 kg and residuals ranging from -4.88 to 4.18 kg. This PE was not significantly different from the PE of 1.9 kg from DXA.

DISCUSSION

The purpose of this study was to evaluate whether DXA predicted MW within acceptable limits for the sport of wrestling. Secondly, the DXA prediction error was compared against the currently approved SF prediction. DXA and SF were cross-validated against an HW criterion, using the methods suggested by Lohman (20), Guo and Chumlea (14), and Altman and Bland (1) in a sample independent from that used to construct the equations (14). This approach examines how each method performed for the group as well as for individual subjects.

The sample was consistent with previous samples of high school wrestlers from Wisconsin in terms of age, body weight, body fat, height, and MW (7,9). All weight classes were represented, with a mean \pm SD of 7.2 ± 2.8 wrestlers in each, with the exception of the heavyweight class. A breakdown by weight class for 10,581 Wisconsin high school wrestlers from the 2006–2007 season is compared with the study sample in Table 3. The number of wrestlers per weight class for the study data is based on their MW at 7% body fat, measured by the HW criterion. One limitation to this study is the percentage of wrestlers from lighter weight classes. Table 3 illustrates that 72% of the wrestlers in WI are eligible for the 152-lb weight class or lower, compared with 83% of the study sample. Conversely, 28% of state wrestlers are eligible for 160-lb class or higher, represented by 17% of the study sample.

The SEE represents the SD of the residual scores (or the difference between predicted and criterion scores). The SEE was 1.7 and 2.0 kg for DXA and SF, respectively. Whereas the SEE reflects deviations of predicted MW from criterion MW about a regression line that fits the scatter plot, the PE is reflective of the deviations around the line of identity (a perfect prediction), with an intercept of 0 and a slope of 1 (7–9,20,28,29). For this reason, the PE has been considered the best single variable for evaluating true differences

TABLE 3. Breakdown by weight class for 10,581 Wisconsin high school wrestlers from the 2006–2007 season compared with the study sample.

Weight Class	Participants	%	N	%
103	571	5.8	7	7.4
112	641	6.5	8	8.5
119	754	7.7	7	7.4
125	830	8.5	8	8.5
130	793	8.1	11	11.7
135	848	8.7	9	9.6
140	866	8.8	12	12.8
145	794	8.1	8	8.5
152	944	9.6	8	8.5
160	909	9.3	7	7.4
171	833	8.5	5	5.3
189	699	7.1	3	3.2
215	266	2.7	1	1.1
285	47	0.5		0.0

between a prediction and criterion (9,14,20,28). The smaller the PE, the better a prediction equation performs when applied to an independent sample (14). In similar samples (7,11), PE values of 2.0 kg have been considered excellent to ideal. Although a PE value that would denote a successful cross-validation has not been set (14), the 1.7-kg SEE value for DXA in this study is classified by Lohman (19) as ideal when compared with the HW criterion.

To put this into perspective for MW testing, we offer an example. In previous work with wrestlers by this author (7), the PE of leg-to-leg BIA was 3.5 kg, with a residual range of -10.4 kg to $+6.9$ kg. MW was underpredicted in one subject by 10.4 kg (22.9 lb) and overpredicted in another subject by 6.9 kg (15.2 lb). Using 2.0 kg, discussed earlier as an acceptable difference, only 23 out of 57 (40%) of the BIA values were within 2.0 kg of the criterion value. This high PE was unacceptable to the WIAA and Wisconsin wrestling coaches advisory group (Wisconsin Interscholastic Athletic Association, 7% Committee Meeting minutes, section V, paragraph A, March 30, 2005). This example illustrates that although there was no significant difference in means, and the regression line was similar to the line of identity, large individual variation was seen when estimating MW for individual wrestlers. Equality in the weight classification system and the health and safety of the wrestler depend on individual values, not average values. Therefore, evaluation of the data scatter, characterized by the SEE, the PE, and residual range (i.e., individual variation), is critical when evaluating a prediction method in wrestlers.

In contrast to this example illustrating unacceptably large individual variation, the current study indicates that DXA had a low SEE of 1.7 kg and a low PE of 1.9 kg, with residuals ranging from -3.94 to 2.88 kg. The distribution of the residual scores from the data approximate a bell-shaped curve (normal distribution) for both DXA and SF. A nonsignificant *P* value for the Shapiro–Wilk test (25) indicate that there is no violation against the normal distribution assumption; therefore, the “68%–95%” rule for the residual scores is appropriate. Thus, DXA predicted MW within ± 1.7 kg (3.7 lb) of the criterion 68.3% of the time and 3.4 kg (7.4 lb) 95.4% of the time in this sample. Although the PE of 1.9 kg from DXA was not significantly different from the PE of 2.1 from SF, the residual range for SF was higher (-4.88 to 4.18 kg). Furthermore, these data represent the lowest PE seen in methods to predict MW in wrestlers when compared with SF (2.1–2.25 kg) (8,24,29), standard BIA (2.77 kg) (9), leg-to-leg BIA (3.5–3.9 kg) (7,31), near-infrared interactance (6.03 kg) (9), and air displacement plesiomography (2.35 kg) (30).

To examine systematic bias in DXA, a plot was constructed according to Altman and Bland (1) (Fig. 2). This approach examines the magnitude of the difference between the prediction and criterion for individual subjects and determines whether prediction accuracy varied with body weight. Because there was no significant mean

difference between the two measures, and the regression analysis had a nonsignificant correlation and a slope and intercept not significantly different than 0, we concluded that no systematic bias was present in DXA across the body weight range seen in these subjects (Fig. 2). The plot indicates that the direction of the difference in predicted MW was not a function of weight; if it were, this would be indicated by a significant correlation, slope, and intercept. A significant negative correlation (crossing the zero line near the middle-weight classes) would have suggested an overprediction of MW in lighter wrestlers and underprediction in heavier wrestlers. Conversely, a significant positive correlation would suggest an underprediction in lighter wrestlers and overprediction in heavier wrestlers.

The two-component (2C) HW technique has historically been used as the criterion method to evaluate body composition in wrestlers (9,16,24,25,28–31). However, Clark et al. (8,11) have previously questioned the validity of the 2C technique in wrestlers because it assumes consistency in hydration and density of the FFM component (12,15,19). Hydration assumptions may be violated during periods of weight cutting for competition and MW testing (22,27,32). In addition, because wrestlers may have more advanced muscular development than their age-matched, nonathletic counterparts, assumptions regarding the density of the FFM may be violated. Any potential changes in the inherent assumption of the 2C technique, such as changes in hydration or the density of the FFM, will contribute to error in MW prediction (19).

A more comprehensive, four-compartment model (4C) model does not rely on the assumptions present in the 2C model because independent measures of BD, bone mineral content (BMC) (used to compute body mineral) and total body water (TBW) are made (4). However, our previous studies have compared HW with 4C in wrestling samples (7,8,11), and they support the HW criterion. In a study by Friedl et al. (12), the error associated with a 4C model was estimated from a reported Chronbach’s alpha at 2.3%BF. The variation in MW (using average weight and %BF of the sample) associated with ± 1 SD in body fat was ± 1.54 kg. This value represents an estimate of the standard error in MW prediction associated with the criterion measure. The PE value from HW (1.34 kg) in the study by Clark et al. (8) was actually lower than this value. This implies that the two values, and techniques, are virtually identical, because a negative variance cannot be defined. It is likely that the smaller value reflected differences between the samples used in the Friedl et al. (12) study and our sample. Friedl et al. (12) also evaluated the reliability of body fat estimation by HW and the 4C model in three sessions in 1 wk. Reliability coefficients were 0.991 and 0.994, and within-subject SD were ± 1.0 and ± 1.1 for %BF estimations from the HW and 4C, respectively; fat mass was ± 0.8 kg with both models (12).

It is important to acknowledge that any criterion has error associated with it, and, therefore, a portion of the PE is

associated with the criterion method. This error can be estimated, and any additional error associated with the prediction can be calculated. Using the same approach as Friedl et al. (12), the error associated with HW criterion can be estimated in this sample. When applied to 7%BF, the range of values associated with ± 1 SD are 4.7 to 9.3%BF. Therefore, in the present study, the variation in MW (using average weight and %BF of this sample) associated with ± 1 SD in body fat would be ± 1.28 kg. This value represents an estimate of the standard error in MW prediction associated with the criterion measure. Using the same approach, the incremental error associated with the DXA method can be calculated by $\sqrt{(1.90^2 - 1.28^2)} = 1.40$ kg and, for SF method, $\sqrt{(2.11^2 - 1.28^2)} = 1.68$ kg. This analysis helps provide a basis for defining the acceptable range for the PE of a technique. A value close to 1.28 kg is virtually identical to the criterion value. A PE value of 1.81 kg represents a doubling of the error variance above the criterion, $\sqrt{(1.28^2 + 1.28^2)} = 1.81$ kg.

Although it is important to acknowledge the error in the criterion measure, it is equally important to identify sources of error in the prediction method. The error present in DXA analysis may be attributable to individual differences in attenuation patterns and tissue thickness, as well differences in the three manufacturers' calibration standards and algorithms used to determine bone, fat, and lean from x-ray attenuation. DXA offers advantages over more laborious methods and measurement of the three body compartments (BMC, fat, and lean) and is not dependent on the constancy of certain physical or chemical characteristics of the compartments. The most important sources of error in DXA are 1) beam hardening (the change in spectrum attributable to preferential attenuation of low energies with increasing body tissue thickness), 2) inadequacy of software algorithms to discriminate between bone and soft tissue, 3) the assumption made regarding the soft tissue in pixels containing bone (above or below the bone), and 4) inadequacy of the calibration algorithms and calibration materials selected for bone, fat, and lean (9,13,17).

The XR-36 DXA discussed this study uses a dynamic filtration system to minimize beam hardening. Gotfredsen et al. (13) evaluated the XR-36 (using dynamic filtration) and found good agreement ($r = 0.99$) between reference and measured amounts of tissue and fat percentages in plastic phantoms as well as smaller (0.5–4 kg) and larger (5–20 kg) additions of tissue (piles of oxen muscle and lard). Gotfredsen et al. (13) conclude that the Norland XR-36, using dynamic filtration, had a high degree of accuracy for body composition analysis, with the potential for becoming a reference method in the future. However, Kohrt (17) suggests that it may be premature to consider DXA the standard measure for soft tissue.

Finally, the authors recommend that future studies standardize on the outcome variable of MW when cross-validating prediction methods for wrestlers. Previous studies have summarized several different variables, including %BF

(24,30,32), FFM (29,31), BD (16), and MW (7–9,11). This inconsistency makes it difficult to compare the statistical analyses and the validity of the prediction. In addition, authors have drawn different conclusions from similar results (7,31). For instance, in our example illustrated earlier, the authors (7), the WIAA, and the Wisconsin coach's advisory group found leg-to-leg BIA to be unacceptable for predicting MW in wrestlers (Wisconsin Interscholastic Athletic Association, 7% Committee Meeting minutes, section V, paragraph A, March 30, 2005). This conclusion was based on the high PE of 3.5 kg and a large residual range of -10.4 kg to $+6.9$ kg that spanned multiple weight classes. However, in a company-funded study that yielded an even larger PE of 3.9 kg in FFM, a larger residual range (with one subject underpredicted by 9.1 kg (20.0 lb) and another overpredicted by 11.8 kg (26.0 lb)), the method was described as an attractive alternative if trained SF testers were not available (31).

As state associations are faced with the choice of methods to measure body fat and calculate MW, convenience can never take precedence over validity. Providing an invalid MW for a minor is a potential risk from both a medical and a legal standpoint. The validity must be thoroughly and independently evaluated as each state is faced with choosing an appropriate method for assessing body composition. We need to consider that most wrestlers elect to wrestle at the lowest-possible weight class to achieve a perceived competitive advantage. Therefore, when predicting MW, we are essentially guiding the wrestler to a safe amount of weight that can be lost. Fairness in the wrestling weight classification system and, more importantly, the wrestler's health and safety, depend on individual values, not group means. For this reason, it is important to critically evaluate the data scatter, individual variation, and residual range when choosing an appropriate method to estimate MW.

CONCLUSION

Because there was no significant difference in means, the regression for DXA versus HW was not significantly different from the line of identity, a strong correlation ($r = 0.98$) was found, a low SEE (1.7 kg) and PE (1.9 kg) were found, and no systematic bias was found across weight classes, the prediction was considered valid. Furthermore, this study indicates that DXA provided the lowest PE (1.9 kg) seen in methods to predict MW when compared against SF (8,24,29), standard BIA (9), leg-to-leg BIA (7,31), near-infrared interactance (9), and air displacement plethysmography (30) in wrestlers. In summary, the results indicate that DXA provided a valid prediction of MW in this sample of wrestlers under the conditions of the study.

The authors have no professional relationships with, or funding from, the companies or manufacturers in the present study, and the results do not constitute endorsement of any product by the authors, the WIAA, the University of Wisconsin Hospital and Clinics, or ACSM.

REFERENCES

1. ALTMAN, D. G., and J. M. BLAND. Measurement in medicine: the analysis of method comparison studies. *Statistician* 32:307–317, 1983.
2. AMERICAN MEDICAL ASSOCIATION. Wrestling and weight control. *JAMA* 201:541–543, 1967.
3. BARTOK, C., D. A. SCHOELLER, R. R. CLARK, J. C. SULLIVAN, and G. L. LANDRY. The effect of dehydration on wrestling minimum weight assessment. *Med. Sci. Sports Exerc.* 36:160–167, 2004.
4. BARTOK-OLSON, C. J., D. A. SCHOELLER, J. C. SULLIVAN, and R. R. CLARK. The “B” in the Selinger four compartment body composition formula should be body mineral instead of bone mineral. *Ann. N.Y. Acad. Sci.* 904:342–344, 2000.
5. BEHNKE, A. R., and J. H. WILMORE. *Evaluation and Regulation of Body Build and Composition*. Englewood Cliffs, NJ: Prentice-Hall, 1974.
6. BROZEK, J., F. GRANDE, J. T. ANDERSON, and A. KEYS. Densitometric analysis of body composition: revision of some quantitative assumptions. *Ann. N.Y. Acad. Sci.* 110:113–140, 1963.
7. CLARK, R. R., C. BARTOK, J. C. SULLIVAN, and D. A. SCHOELLER. Is leg-to-leg BIA valid for predicting minimum weight in wrestlers? *Med. Sci. Sports Exerc.* 37:1061–1068, 2005.
8. CLARK, R. R., C. BARTOK, J. C. SULLIVAN, and D. A. SCHOELLER. Minimum weight prediction methods cross-validated by the four-component model. *Med. Sci. Sports Exerc.* 36:639–647, 2004.
9. CLARK, R. R., J. M. KUTA, J. C. SULLIVAN, W. M. BEDFORD, J. D. PENNER, and A. STUDESVILLE. A comparison of methods to estimate minimal weight in high school wrestlers. *Med. Sci. Sports Exerc.* 25:151–158, 1993.
10. CLARK, R. R., and R. A. OPPLIGER. Minimal weight standards in high school wrestling. The Wisconsin Model. *Orthop. Phys. Ther. Clin. N. Am.* 7:23–45, 1998.
11. CLARK, R. R., J. C. SULLIVAN, C. BARTOK, and D. A. SCHOELLER. Multicomponent cross-validation of minimum weight predictions for college wrestlers. *Med. Sci. Sports Exerc.* 35:342–347, 2003.
12. FRIEDL, K. E., J. P. DELUCA, L. J. MARCHITELLI, and J. A. VOGEL. Reliability of body-fat estimations from a four-compartment model by using density, body water, and bone mineral measurements. *Am. J. Clin. Nutr.* 55:764–770, 1992.
13. GOTTFREDSEN, A., L. BAEKSGAARD, and J. HILSTED. Body composition analysis by DEXA by using dynamically changing samarium filtration. *J. Appl. Physiol.* 82:1200–1209, 1997.
14. GUO, S. S., and W. C. CHUMLEA. Statistical methods for the development and testing of prediction equations. In: *Human Body Composition*, A. F. Roche, S. B. Heymsfield, and T. G. Lohman (Eds.). Champaign, IL: Human Kinetics Books, pp. 191–202, 1996.
15. HEYMSFIELD, S. B., J. WANG, S. HESKA, J. J. KEHAYIAS, and R. N. PIERSON. Dual-photon absorptiometry: comparison of bone mineral and soft tissue mass measurements in vivo with established methods. *Am. J. Clin. Nutr.* 49:1283–1289, 1989.
16. HOUSH, T. J., G. O. JOHNSON, K. B. KENNEY, et al. Validity of anthropometric estimations of body composition in high school wrestlers. *Res. Q. Exerc. Sport.* 60:239–245, 1989.
17. KOHRT, W. M. Body composition by DXA: tried and true? *Med. Sci. Sports Exerc.* 27:1349–1353, 1995.
18. LESSER, L. I., C. B. EBBELING, M. GOOZNER, D. WYPIJ, and D. S. LUDWIG. Relationship between funding source and conclusion among nutrition-related scientific articles. *PLoS Med.* 4:e5, January 2007.
19. LOHMAN, T. G. *Advances in Body Composition Assessment: Current Issues in Exercise Science*. Champaign, IL: Human Kinetics Books, 1992.
20. LOHMAN, T. G. Skinfolts and body density and their relation to body fatness: a review. *Hum. Biol.* 53:181–225, 1981.
21. NCAA Wrestling Weight Certification Forms: Skinfold Directions. Indianapolis, IN: NCAA, 1998.
22. OPPLIGER, R. A., H. S. CASE, C. A. HORSWILL, G. L. LANDRY, and A. C. SHETLER. American College of Sports Medicine position stand: weight loss in wrestlers. *Med. Sci. Sports Exerc.* 2:ix–xii, 1996.
23. OPPLIGER, R. A., R. R. CLARK, and J. M. KUTA. Efficacy of skinfold training clinics: a comparison between clinic trained and experienced testers. *Res. Q. Exerc. Sport.* 63:438–443, 1992.
24. OPPLIGER, R. A., D. H. NIELSEN, and C. G. VANCE. Wrestlers’ minimal weight: anthropometry, bioimpedance, and hydrostatic weighing compared. *Med. Sci. Sports Exerc.* 23:247–253, 1991.
25. ROYSTON, P. An extension of Shapiro and Wilk’s W test for normality to large samples. *Appl. Stat.* 31:115–124, 1982.
26. SINNING, W. E. Body composition assessment of college wrestlers. *Med. Sci. Sports Exerc.* 6:139–145, 1974.
27. STEEN, S. N., and K. D. BROWNE. Patterns of weight loss and regain in wrestlers: has the tradition changed? *Med. Sci. Sports Exerc.* 6:762–768, 1990.
28. THORLAND, W. G., G. O. JOHNSON, C. J. CISAR, and T. J. HOUSH. Estimation of minimal wrestling weight using measures of body build and body composition. *Int. J. Sports Med.* 8:365–370, 1987.
29. THORLAND, W. G., C. M. TIPTON, T. G. LOHMAN, et al. Midwest wrestling study: prediction of minimal weight for high school wrestlers. *Med. Sci. Sports Exerc.* 23:1102–1110, 1991.
30. UTTER, A. C., F. L. GOSS, P. D. SWAN, G. S. HARRIS, R. J. ROBERTSON, and G. A. TRONE. Evaluation of air displacement for assessing body composition of collegiate wrestlers. *Med. Sci. Sports Exerc.* 35:500–505, 2003.
31. UTTER, A. C., D. C. NIEMAN, G. J. MULFORD, et al. Evaluation of leg-to-leg BIA in assessing body composition of high-school wrestlers. *Med. Sci. Sports Exerc.* 37:1395–1400, 2005.
32. WEBSTER, S., R. RUTT, and A. WELTMAN. Physiological effects of a weight loss regimen practiced by college wrestlers. *Med. Sci. Sports Exerc.* 2:229–234, 1990.
33. WILMORE, J. H. A simplified method for determination of residual lung volumes. *J. Appl. Physiol.* 27:98–100, 1969.