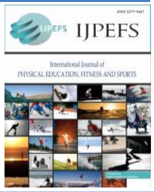
International Journal of
PHYSICAL EDUCATION, FITNESS AND SPORTS




Highlighted Article

# Psychometric evidence of body composition as a multidimensional trait in college students

Received 3rd November 2017
Accepted 30th November 2017

www.ijpefs.com

DOI:10.26524/ijpefs1741

**Peter D. Hart** [a,b,c *]

[a] Health Promotion Program, Montana State University - Northern, Havre, MT 59501, USA
[b] Kinesmetrics Lab, Montana State University - Northern, Havre, MT 59501, USA
[c] Health Demographics, Havre, MT 59501, USA
*Corresponding Author: Ph: 406.265.3719; Fax: 406.265.4129; Email: peter.hart@msun.edu

**Abstract:** Body composition (BC) assessment is often conducted using one of several different field techniques, which individually are considered valid tests. Anecdotal evidence has suggested, however, that some individuals may rank relatively high when assessed by one method and relatively low when assessed by another method. This inconsistency would indicate that BC assessments have poor convergent validity. The purpose of this study was to examine the convergent validity of common BC assessments using a norm-referenced approach. A total of 67 college students participated in this measurement study and had their BC assessed by each of three different tests: percent body fat (PBF) by skinfold technique (PBFSF), waist circumference (WC), and body mass index (BMI). Two different statistical procedures were used to evaluate convergent validity of the three BC assessments. First, Cohen's weighted kappas were calculated using quartiles of each BC measure. This analysis utilized three different 4 x 4 tables from all BC measure pairs. Second, Bland and Altman limits of agreement (LOA) plots were constructed on all pairs after T-score transformation of each measure. Mean (SD) values of PBFSF (%), WC (cm), and BMI (kg/m2) were 12.3 (5.0), 87.0 (8.3), 26.8 (3.5) and 23.3 (7.0), 77.1 (8.8), 24.8 (3.2) for males and females, respectively. Simple kappas showed poor agreement across the three pairs of BC assessments and ranged from .14 to .17. The weighted kappas improved to fair agreement and ranged from .32 to .38. None of the three LOA plots showed systematic bias toward a method. However, 95% LOA were wide for PBFSF vs. WC (± 28.9), BMI vs. PBFSF (± 25.9), and BMI vs. WC (± 12.3). Results of this measurement study indicate that common BC assessments have poor convergent validity among college students. These results further indicate that BC may be a multidimensional trait, requiring a specific test depending on the specific trait of interest.

***Key Words:*** Body composition, Measurement, Health-related fitness, Percent body fat

***Dr. Peter D. Hart*** *is an Associate Professor of Health Promotion at Montana State University - Northern. He is also Director of the Kinesmetrics Lab. He is trained in quantitative research methods including epidemiology, applied statistics, and measurement & evaluation. Dr. Hart's research interests are in applied health & kinesiology, measurement issues related to health assessments, secondary data & meta-analyses, and population health.*

## Introduction

Body composition (BC) is one of the five components of health-relate fitness, along with cardiorespiratory fitness, flexibility, muscular strength, and muscular endurance [1]. There is a growing concern placed on BC because of an international obesity pandemic. Specifically, obesity has tripled world-wide since 1975 with more than 650 million adults classified as obese in 2016 [2]. Furthermore, overweight and obesity are related to many worldwide chronic disease burdens, such as heart disease, cancer, and diabetes [3]. Due to a growing interest in both overweight and obesity, accurate assessment of BC becomes an important measurement issue.

BC is usually considered a measure of fat mass on a person's body relative to their total body mass [4]. Gold-standard methods are available to assess BC (e.g., hydrostatic weighing), but are generally restricted to clinical settings [5]. BC assessment is more commonly conducted using one of several different field techniques, which individually are considered valid tests [6]. Anecdotal evidence has suggested, however, that some individuals may rank relatively high when assessed by one method and relatively low when assessed by another method. This inconsistency would imply that BC assessments have poor convergent validity [7]. The purpose of this study was to examine the convergent validity of common BC assessments using a norm-referenced approach.

## Methods

### Participants

Data for this research came from a cross-sectional measurement study conducted at a rural public university. A total of N=67 male and female college students who had their BC assessed by each of three methods were included in this study. Participants were recruited by both study flyer and word-of-mouth. All study components were reviewed and approved by the university's institutional review board (IRB).

### Body composition measures

Three different BC measures were used in this study: percent body fat (PBF) by skinfold technique (PBFSF), waist circumference (WC), and body mass index (BMI). BC measures were assessed by trained clinicians and followed ACSM guidelines [8].

PBFSF (%) was measured using the Siri equation. Body density was first measured using the sum of chest, abdomen, and thigh skinfolds for males and triceps, suprailiac, and thigh skinfolds for females. WC (cm) was measured similarly for males and females and required an elastic tape placed at the narrowest point between the xyphoid process and umbilicus. BMI (kg/m2) was measured similarly for males and females and required measuring height (cm) using a wall mounted stadiometer and weight (kg) using an electronic floor scale.

### Statistical Analysis

Two different statistical procedures were used to evaluate convergent validity of the three BC assessments. First, both simple and weighted kappas [9,10] were calculated using quartiles of each BC measure. This analysis utilized three different 4x4 tables from all BC measure pairs. Second, Bland and Altman limits of agreement (LOA) plots [11] were constructed on all pairs after T-score transformation of each measure. All analyses were conducted using SAS version 9.4 [12].

## Results

Table 1 displays descriptive statistics and convergent validity correlations for the BC measures by sex. Mean (SD) values of PBFSF (%), WC (cm), and BMI (kg/m2) were 12.3 (5.0), 87.0 (8.3), 26.8 (3.5) and 23.3 (7.0), 77.1 (8.8), 24.8 (3.2) for males and females, respectively. Pearson correlations were weak to moderately strong, with stronger values seen between BMI and WC. Also, a sex difference (p=.019) in convergent validity strength was seen for PBFSF and WC, where female measures showed no significant correlation (p=.390).

**Table 1.** Descriptive statistics and correlations for BC measures (N=67).

| | Males (N=39) | | Females (N=28) | | *t-stat* |
|---|---|---|---|---|---|
| BC measure | *M* | *SD* | *M* | *SD* | *p* |
| **PBFSF** | 12.34 | 5.01 | 23.31 | 7.01 | <.001 |
| **WC** | 87.00 | 8.28 | 77.11 | 8.76 | <.001 |
| **BMI** | 26.77 | 3.45 | 24.79 | 3.16 | .019 |

| | Males (N=39) | | Females (N=28) | | *Z-stat* |
|---|---|---|---|---|---|
| BC measure pairs | *r* | *p* | *r* | *p* | *p* |
| **PBFSF & WC** | .654 | <.001 | .169 | .390 | .019 |
| **PBFSF & BMI** | .464 | .003 | .466 | .012 | .992 |
| **BMI & WC** | .801 | <.001 | .796 | <.001 | .960 |

***Note.*** *M* is mean. *SD* is standard deviation. *r* is Pearson correlation coefficient. *t-stat* is independent t statistic. *Z-stat* is test statistic for difference between correlation coefficients.

**Table 2.** Categorical agreement statistics as convergent validity evidence (N=67).

| BC measure pairs | $\chi^2$ | $\chi^2_M$ | *P* | *Phi* | *r* | K | $K_W$ |
|---|---|---|---|---|---|---|---|
| **PBFSF vs. WC** | 17.26[a] | 2.67[b] | 38.8 | .510 | .424 | .183 | .322 |
| **PBFSF vs. BMI** | 22.02[a] | 4.96[b] | 35.8 | .573 | .463 | .144 | .324 |
| **BMI vs. WC** | 29.93[a] | 4.84[b] | 37.3 | .668 | .555 | .165 | .376 |

***Note.*** Fleiss simple and weighted kappas are .164 and .341, respectively. $\chi^2$ is chi-square test statistic. $\chi^2_M$ is McNemar chi-square statistic. *P* is proportion of agreement statistic. *Phi* is an adjusted chi-square statistic measure of association. *r* is the Pearson correlation coefficient for quartile categories. K is simple kappa. $K_W$ is weighted kappa. [a] indicates significant at .05 level. [b] indicates not significant at .05 level.

Table 2 displays the categorical agreement statistics for convergent validity evidence. Simple kappas showed poor agreement across the three pairs of BC assessments and ranged from .14 to .18. The weighted kappas improved to fair agreement and ranged from .32 to .38.

Figure 1 shows the percentage of categorical agreement observed for all three methods by BC quartile. In this graph, the largest quartile (4th) represents individuals with the largest BMI, largest WC, and greatest PBFSF. A participant was assigned to a quartile if any of the three methods placed them in that relative position. Results here indicated that participants in the third quartile fared the worst in BC assessment accuracy, with no participants (72.5%) (That were placed in that quartile from either method) receiving the same relative standing from the other methods. As well, no participants (0%) in the third quartile saw all three BC methods place them in that relative position.

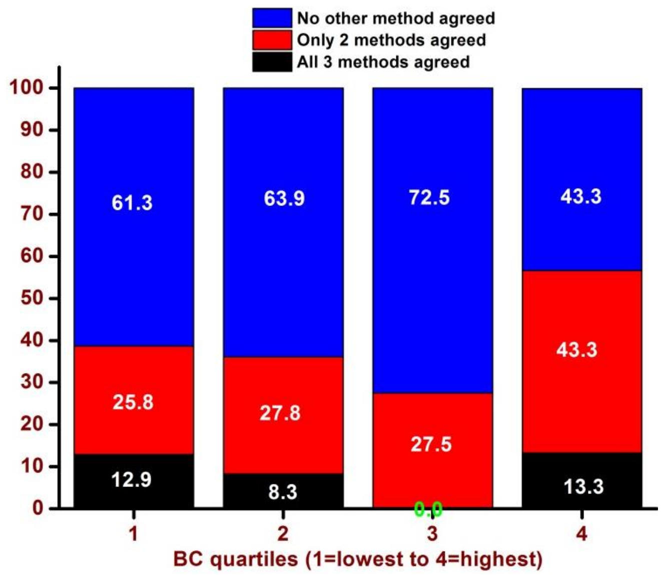



**Fig. 1.** Percentage of categorical agreement for all methods.

Figures 2 thru 4 display the Bland and Altman LOA plots for each BC assessment pair. In these plots, the vertical axes represent the difference (i.e., method 1 – method 2, etc.) in two BC methods, which were both T-score transformed. Therefore, a participant located at the zero line vertical position would indicate they were ranked the same with both BC methods. Furthermore, a participant located at the +10 vertical position would indicate the first BC method (i.e., method 1) ranked them 1 SD higher as compared to the second method (i.e., method 2). Results here show that none of the three LOA plots showed systematic bias toward a method (i.e., scatter equally distributed above and below the horizontal zero line). However, 95% LOA were wide for PBFSF vs. WC (± 28.9), BMI vs. PBFSF (± 25.9), and BMI vs. WC (± 12.3).

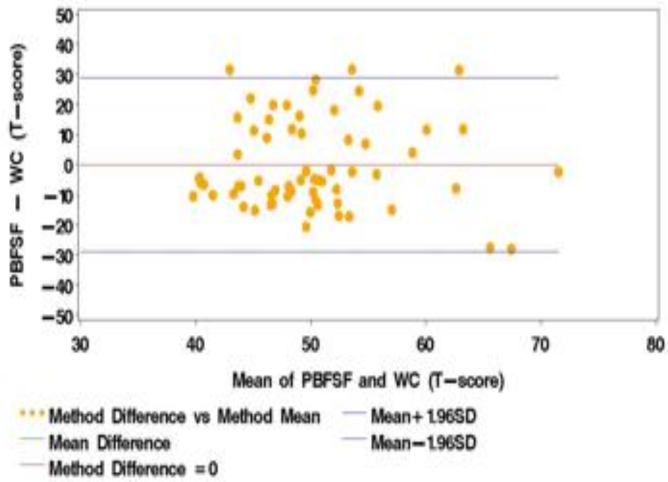



**Fig. 2.** Bland & Altman plot of PBFSF and WC.

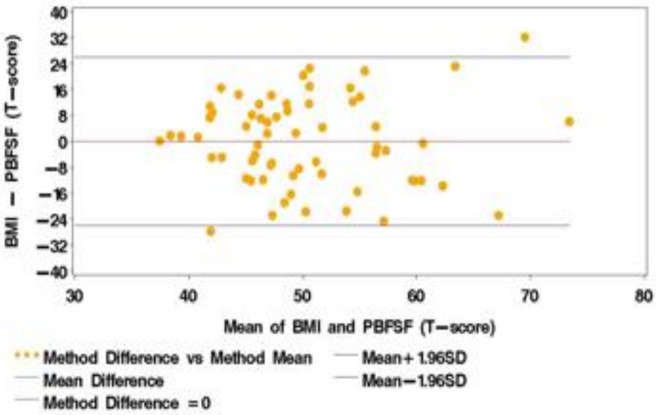



**Fig. 3.** Bland & Altman plot of BMI and PBFSF.

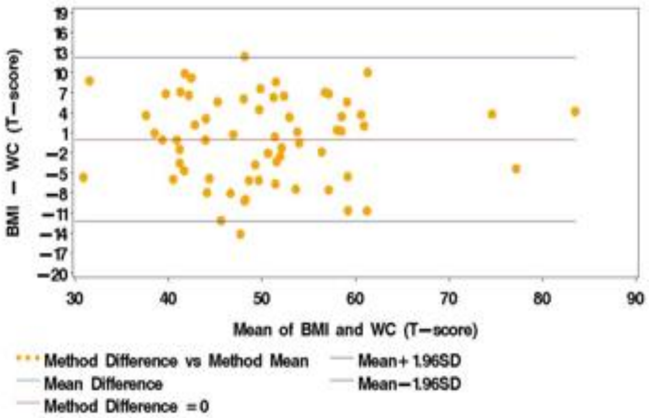



**Fig. 4.** Bland & Altman plot of BMI and WC.

## Discussion

The aim of this study was to examine the convergent validity of three commonly administered BC assessments in college students. In a norm-referenced context, individuals are ranked according to their measurement on a trait. Therefore, valid BC measures (in a norm-referenced context) would allow for accurate ranking of individuals across any BC assessment. Results of this study show, however, that an individual's relative standing in terms of BC is not always the same across the three BC assessments. These results indicate that each of the three BC tests in this study are measuring a similar yet different trait. Other similar studies have supported these findings. For example, a measurement study involving Sri Lankan premenopausal women found only a moderate correlation between BC measures of percent fat mass and BMI [13]. Another measurement study involving subjects with spinal cord injury showed that BMI was only moderately correlated with WC [14]. These findings, however, contradict those from other similar studies. For example, a recent research conducted on Sri Lankan children showed strong fat mass correlations with BMI, WC, as well as hip circumference [15]. Another BC measurement study of young Asian Indian women found strong PBF correlations with BMI, WC, as well as WC to height ratio [16].

The reasoning behind these mixed findings may simply be associated with the fitness-related obesity paradox [17]. That is, individuals who have greater fitness levels, are more likely to have greater amounts of lean body mass and hence have lower values of PBF and larger values of BMI. This paradox makes sense in particular among college-aged individuals, where both physical activity and fitness levels are generally greater than other age groups [18].

The generalizations made from this study should be considered along with its limitations. The main limitation in this study was the specific population in which the sample was drawn from. Since measurement properties of any test are situation specific [7], results from this study should be considered only for college students attending a rural public university.

## Conclusions

Results of this measurement study indicate that common BC assessments have poor convergent validity among college students. These results further indicate that BC may be a multidimensional trait, requiring a specific test depending on the trait of interest. Clinicians and researchers who collect BC data should be aware of their lack of convergence in college students.

## Acknowledgements

No financial assistance was used to assist with this project.

## References

[1] American College of Sports Medicine, editor. ACSM's health-related physical fitness assessment manual, Lippincott Williams & Wilkins; 2013 Jan 21.

[2] World Health Organization (WHO. Obesity and Overweight factsheet from the WHO. Health. 2017 Oct 24.

[3] J.C. Seidell, J. Halberstadt, The global burden of obesity and the challenges of prevention, *Annals of Nutrition and Metabolism*, 66 (2015) 7-12.

[4] W.J. Kraemer, S.J. Fleck, M.R. Deschenes, Exercise physiology: integrating theory and application, Lippincott Williams & Wilkins; 2011 Mar 1.

[5] W.D. McArdle, F.I. Katch, & V.L. Katch, (2010). Exercise physiology: nutrition, energy, and human performance. Lippincott Williams & Wilkins.

[6] A. Aandstad, K. Holtberget, R. Hageberg, I. Holme, & S.A. Anderssen, Validity and reliability of bioelectrical impedance analysis and skinfold thickness in predicting body fat in military personnel, *Military Medicine*, 179 (2014) 208-217.

[7] J.R. Morrow Jr, D. Mood, J. Disch, M. Kang, Measurement and Evaluation in Human Performance, 5th Edition, Human Kinetics; 2015 Aug 25.

[8] American College of Sports Medicine. (2013). ACSM's guidelines for exercise testing and prescription. Lippincott Williams & Wilkins.

[9] J.L. Fleiss, B. Levin, & M.C. Paik, (2013). Statistical methods for rates and proportions. John Wiley & Sons.

[10] J. Cohen, Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit, *Psychological Bulletin*, 70 (1968) 213-220.

[11] J.M. Bland, & D. Altman, Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327 (1986) 307-310.

[12] R.P. Cody & J.K. Smith, (2006), Applied statistics and the SAS programming language. 5th Edition. Pearson.

[13] J. Nanayakkara & S. Lekamwasam, Validity of BMI, hip and waist circumferences as surrogate measures of obesity in a cohort of Sri Lankan premenopausal women, *Ceylon Medical Journal,* 58 (2013) 72–75.

[14] I. Eriks-Hoogland, R. Hilfiker, M. Baumberger, S. Balk, G. Stucki, C. Perret, Clinical assessment of obesity in persons with spinal cord injury: validity of waist circumference, body mass index, and anthropometric index, *The Journal of Spinal Cord Medicine*, 34 (2014) 416-422.

[15] V.P. Wickramasinghe, C. Arambepola, D.M. Bandara, M. Abeysekera, S. Kuruppu, P. Dilshan, B.S. Dissanayake, Validity of newly-developed BMI and waist cut-off values for Sri Lankan children, *Annals of Human Biology*, 40 (2013) 280-285.

[16] N. Singhal, P. Mathur, R. Pathak, Validity of simple, novel measures of generalized and central obesity among young Asian Indian women, Indian *Journal of Medical Sciences* 65 (2011) 518-527.

[17] P.A. McAuley, P.F. Kokkinos, R.B. Oliveira, B.T. Emerson, J.N. Myers, Obesity paradox and cardiorespiratory fitness in 12,417 male veterans aged 40 to 70 years, *Mayo Clinic Proceedings*, 85 (2010) 115-121.

[18] P. Raven, D. Wasserman, W. Squires, T. Murray, Exercise Physiology, Nelson Education; 2012.