

5-1-2012

Reliability and Validity of the Body Caliper to Evaluate Body Composition

Gregory L. Stalker
University of Nevada, Las Vegas

Follow this and additional works at: <https://digitalscholarship.unlv.edu/thesesdissertations>



Part of the [Kinesiology Commons](#)

Repository Citation

Stalker, Gregory L., "Reliability and Validity of the Body Caliper to Evaluate Body Composition" (2012). *UNLV Theses, Dissertations, Professional Papers, and Capstones*. 1633.
<http://dx.doi.org/10.34917/4332614>

This Thesis is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Thesis has been accepted for inclusion in UNLV Theses, Dissertations, Professional Papers, and Capstones by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

RELIABILITY AND VALIDITY OF THE BODY CALIPER
TO EVALUATE BODY COMPOSITION

by

Gregory L. Stalker

Bachelor of Arts in English
University at Albany, State University of New York
2001

A thesis submitted in partial fulfillment
of the requirements for the

Master of Science Degree in Exercise Physiology

**Department of Kinesiology and Nutrition Sciences
School of Allied Health Sciences
The Graduate College**

**University of Nevada, Las Vegas
May 2012**

Copyright by Gregory L. Stalker, 2012
All rights reserved



THE GRADUATE COLLEGE

We recommend the thesis prepared under our supervision by

Gregory L. Stalker

entitled

**Reliability and Validity of the Body Caliper to
Evaluate Body Composition**

be accepted in partial fulfillment of the requirements for the degree of

Master of Science in Exercise Physiology
Department of Kinesiology and Nutrition Sciences

Lawrence A. Golding, Ph.D., Committee Chair

John Mercer, Ph.D., Committee Member

Richard D. Tandy, Ph.D., Committee Member

Merrill Landers, D.P.T., Graduate College Representative

Ronald Smith, Ph. D., Vice President for Research and Graduate Studies
and Dean of the Graduate College

May 2012

ABSTRACT

**RELIABILITY AND VALIDITY OF THE BODY CALIPER
TO EVALUATE BODY COMPOSITION**

by

Gregory L. Stalker

Dr. Lawrence A. Golding, Examination Committee Chair
Professor of Kinesiology
University of Nevada, Las Vegas

The Body Caliper is a relatively new model of skin-fold caliper (approximately 13 years old) whose reliability and validity has yet to be investigated on an adult population. Therefore, the purpose of this study was to investigate the reliability and validity of the Body Caliper as an instrument and method for evaluating body composition in adults. One hundred adults (62 men, 38 women) aged 18 to 60 years participated in this cross-sectional study. Skin-fold measurements were taken with the Body Caliper using the YMCA protocol, and body density was calculated by the Jackson & Pollock sum-of-3 equations for men and women. Hydrostatic weighing served as the criterion method that the Body Caliper was compared to for validation. The reliability of the Body Caliper was assessed by an intraclass correlation coefficient and was proven reliable ($ICC = .997$). The validity of the Body Caliper was assessed by Bland-Altman analysis and was proven valid by 96% of the sample reaching agreement between the two methods. The results of this study have demonstrated the Body Caliper to be a reliable and valid instrument and method for evaluating body composition.

TABLE OF CONTENTS

ABSTRACT	iii
LIST OF TABLES	v
LIST OF CHARTS	vi
CHAPTER 1 INTRODUCTION.....	1
Purpose of the Study	3
Research Questions	3
Significance of the Study	3
CHAPTER 2 REVIEW OF RELATED LITERATURE.....	5
CHAPTER 3 METHODOLOGY	13
Subject characteristics	13
Instrumentation	13
Collection of the Data	14
Data analysis methods.....	19
CHAPTER 4 RESULTS	21
CHAPTER 5 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS.....	24
Discussion of Results	24
Conclusions and Recommendations for Further Study.....	25
APPENDIX A	26
APPENDIX B	35
REFERENCES.....	37
VITA.....	39

LIST OF TABLES

Table 1	Characteristics of the Participants.....	13
Table 2	Reliability of the Body Caliper	21

LIST OF CHARTS

Chart 1	Body Density Measured by Two Methods	22
Chart 2	Limits of Agreement	23

CHAPTER 1

INTRODUCTION

Body composition refers to the components of a whole body and at its basic level, is expressed as a percentage of fat mass, with the remaining percentage representing fat-free mass (Golding, 2000). This expression of body composition is known as the two-compartment model (Golding, 2000). Knowledge of body composition is important when developing weight loss, gain or maintenance programs, and exercise programs (Golding, 2000). When body composition is evaluated at the beginning of a nutrition or exercise program, it forms a baseline for comparison to future evaluations that will reveal success or failure as well as determine the efficacy of the program (Golding, 2000).

Two long-standing methods of the two-compartment model are hydrodensitometry and skin-fold assessment (Heyward & Wagner, 2004). Although both methods are indirect and may estimate body composition under the same model, they are quite opposite. Hydrodensitometry, also referred to as *hydrostatic* or *underwater weighing*, is the traditional method of reference that other methods use for validation (Heyward & Wagner, 2004). Because it is a laboratory procedure, hydrodensitometry requires a lot of time, expensive equipment, and a highly skilled staff to provide accurate measurements (Heyward & Wagner, 2004). On the other end of the spectrum is skin-fold assessment, the method of reduced cost. Primarily used in the field, skin-fold assessment requires much less time, equipment, and manpower to provide estimations of body composition (Heyward & Wagner, 2004). The popularity of skin-fold assessment is derived from its quickness and portability; the procedure usually lasts less than 10 minutes and can be performed in most settings (Heyward & Wagner, 2004). As

compared to hydrodensitometry, the level of skill required to perform a skin-fold assessment is less, however, a well-trained tester is able to provide estimations of body composition within the acceptable range of $\pm 3.5\%$ body fat (Heyward & Wagner, 2004).

An estimation of body composition requires a series of two equations (Golding, 2000). The product of the first equation is body density, which is then entered into the second equation to produce percent fat and fat-free mass (Golding, 2000). Due to the assumption of a constant density of fat-free mass for the entire population and the potential for increased error that emerges from this assumption (Golding, 2000), the equations that predict percent fat and fat-free mass will not be included in this study. Therefore, this study will use body density as the lone dependent variable in the determination of validity.

The Body Caliper is a relatively new model of skin-fold caliper (approximately 13 years old) that can measure skin-fold thickness in a range of 0 to 60 millimeters ("The Body Caliper"). With the measurement scale printed on both sides, either right- or left-handed individuals may operate the Body Caliper with ease ("The Body Caliper"). The Body Caliper houses a patented mechanism, titled the "AccuSpring," that is claimed to deliver the National Research Council's standard jaw pressure of $10\text{g}/\text{mm}^2$ (Keys, 1956) throughout its entire range of measurement, making it compatible with all skin-fold protocols ("The Body Caliper"). Another claim made by the manufacturer is that the Body Caliper is calibrated by hand at the factory and never needs recalibration, offering a lifetime of accurate measurements ("The Body Caliper"). An attractive proposition stated to consumers and an encompassing claim made by the manufacturer declared the Body Caliper to be the "first research-grade caliper at an affordable price" ("The Body

Caliper,” n.p.) Priced around \$50, affordability seems inherent, however, determining the merit of the claim of being research-grade was the primary concern of this study.

Purpose of the Study

The purpose of this study was to investigate the reliability and validity of the Body Caliper as an instrument and method for evaluating body composition in adults.

Research Questions

Question #1. Does the Body Caliper provide reliable measurements of skin-fold thickness in adults?

Question #2. Does the Body Caliper provide valid measurements of body density in adults?

Significance of the Study

This study is significant because no publication exists that has investigated the Body Caliper with the population and methods utilized in this study. In fact, only one publication exists that has studied the Body Caliper and that forms the only support of the manufacturer's claim of offering a research-grade instrument. The referenced study was conducted at the University of New Mexico by Astorino, Orri, Lockner, Jenkins, and Heyward (1999), published more than a decade ago and was a simple comparison of skin-fold measurements made by the Body Caliper and the Harpenden caliper. Complete validation of any instrument by one study alone is troubling and the referenced study raises the following additional concerns. First, jaw pressure was not measured to determine if it conforms to the standard 10g/mm² as claimed by the manufacturer. Also, the accuracy of the Body Caliper's measurement scale went unexamined. Second, although the Harpenden caliper has been a standard caliper of research for decades, it

does not represent the criterion method of body composition. Third, since the whole sample consisted of children, the concluding recommendation of "the Body Caliper as a low cost alternative for measuring skin-fold thickness" (Astorino et al., 1999, p. S204) is applicable only to children. A gap this large in the literature, especially regarding an adult population, makes this study by the investigator absolutely necessary. Therefore, the significance of this study is derived from this necessity to fill the body of literature, which will serve the scientific community as well as consumers.

CHAPTER 2

REVIEW OF RELATED LITERATURE

According to Golding (2000), "skin-fold measurements are excellent estimates of total body fat" that "require only skin-fold calipers and skill to be reliable and valid" (p. 112). The required skill refers to the testers' ability to accurately determine the location of the sites to be measured as well as their mastery of the technique that they use to gather the skin-folds and operate the caliper (Golding, 2000).

There are many sites on the body where skin-fold thickness may be measured, and their locations are clearly defined by anatomical reference points (Hume & Marfell-Jones, 2008). Due to uneven distributions of subcutaneous fat, the tester's accuracy in locating the sites is extremely important and was proven in a study by Hume and Marfell-Jones (2008). Methods for this study included eight different skin-fold sites that were approved by the International Society for the Advancement of Kinanthropometry (Hume & Marfell-Jones, 2008). Once an exact site was determined and marked, a 1cm grid was marked around the original site, offering an additional eight points of measurement for a total of nine points for each site (Hume & Marfell-Jones, 2008). The results of this study revealed that 70% of the measurements taken at the peripheral markings were significantly different than the central grid point that represented the original approved site location (Hume & Marfell-Jones, 2008). These results clearly demonstrate the importance of accurate site location through the increase in error that was produced by a small deviation of 1cm away from the approved site location.

The YMCA Fitness Testing and Assessment Manual, which includes the YMCA protocol for skin-fold measurement, was first developed and published approximately 40

years ago and is still in use today (Golding, 2000). Lawrence A. Golding, Ph.D., was involved in this original development and remains as the editor of the current edition. As a student of Dr. Golding at the University of Nevada, Las Vegas, the investigator was instructed in the YMCA protocol and achieved proficiency. Fellow contributors to the development of the YMCA protocol were Andrew S. Jackson, Ph.D., and Michael L. Pollock, Ph.D., whose sum-of-3 equations for body density were used in this study (Golding, 2000). Also, the Body Caliper claims compatibility with the YMCA protocol (“The Body Caliper”). Together, these reasons justify the use of the YMCA protocol by investigator in this study.

With the availability of so many prediction equations that transform skin-fold measurements into body density, choosing the right one is just as important as mastering the protocol for skin-fold measurement (Heyward & Wagner, 2004). The two main types of prediction equations to choose from are population-specific and generalized equations (Heyward & Wagner, 2004). Population-specific equations were developed using homogeneous samples and linear regression models, which made them accurate but also limited their applicability to the specific characteristics of the population that they were developed from (Heyward & Wagner, 2004). Generalized equations were developed using large heterogeneous samples and quadratic regression models that utilized the curvilinear relationship found to be between body density and subcutaneous fat, therefore, they produced similar accuracy while being applicable to a large range of a population that may vary greatly in age and fat mass (Heyward & Wagner, 2004).

After the appropriate type of equation (population-specific or generalized) has been determined, the observer must adhere to the protocol and instrument used in

developing the equation while making the observations (Heyward & Wagner, 2004).

After careful review, the investigator has chosen the Jackson & Pollock sum-of-3 equations, specific for each gender, to calculate body density from skin-fold measurements for this study (Jackson & Pollock, 1978; Jackson, Pollock, & Ward, 1980). Skin-folds are measured in millimeters, and age is measured in years.

- Male Body Density= $1.10938 - (0.0008267 \times \text{sum of skin-folds}) + (0.0000016 \times \text{sum of skin-folds squared}) - (0.0002574 \times \text{age})$
- Female Body Density= $1.0994921 - (0.0009929 \times \text{sum of skin-folds}) + (0.0000023 \times \text{sum of skin-folds squared}) - (0.0001392 \times \text{age})$

These equations were chosen because the sample used in their development matched the sample that the investigator intended to recruit for this study. Also, these two equations were developed using the Lange caliper, which has a jaw surface area of 30mm^2 (Golding, 2000) that matches well with the Body Caliper, which was found in an independent study by the investigator (Appendix A) to have a jaw surface area of 28mm^2 . This study by the investigator also produced significant correlations ($r = .992 - .999$, $p < .01$) between the Body Caliper and the Lange at all the sites used in these two equations.

The accuracy of the Jackson & Pollock sum-of-3 equations was validated in a study by Scherf, Franklin, Lucas, Stevenson, and Rubenfire (1986) where the validity of four different equations for estimating body composition from skin-fold measurements were tested by comparison to hydrostatic weighing. The results of this study revealed that of the four equations tested, the Jackson & Pollock sum-of-3 produced the smallest mean difference between the two methods at 0.7% and the smallest total error at 4.4% (Scherf et al., 1986).

Hydrodensitometry has been the long-standing method of reference in body composition under the two compartment model (Heyward & Wagner, 2004). This accreditation derives from the work of Behnke, Feen, and Welham who introduced the method and model in a publication that dates back to 1942. This method measures density using the following equation: $\text{Body Density} = \text{Body Mass} / \text{Body Volume}$ (Golding, 2000). Obtaining the measurement of body mass is as easy as stepping on a scale, however, according to Behnke et al. (1942), "the essential measurement is that of body volume" (p. 495). Behnke's method of hydrostatic weighing was based off of Archimedes' principle of water displacement, which states that the volume of an object is equal to the volume of water it displaces while submerged (Golding, 2000). Rather than collecting and measuring the displaced water, Behnke et al. (1942) built their method around the concept of "equivalent volume = weight in air - weight in water" (p. 495).

Over the years, researchers attempted to increase accuracy and decrease error by making modifications to the methodology and the formula for body density. In 1961, Goldman and Buskirk (1961) revealed a methodology that the methods of this study were based on. The following are examples of the borrowed methodology (Goldman & Buskirk, 1961):

- They used a suspended weighing platform monitored by load cells to reduce the movement of the scale underwater and resulting fluctuation of weight during a reading, which produced a static weight underwater.
- They measured residual lung volume immediately after a hydrostatic weight measurement while the subject was still in the water.

- They recommended a water temperature of 35-36 degrees Celsius for subject comfort and recorded the water temperature prior to subject submersion to account for the different density of water at different temperatures.
- They modified the equation for density to account for the density of the water and the residual volume of air in the gastrointestinal tract.

Goldman and Buskirk's (1961) modified equation for body density was used in this study and is listed below:

- $$\text{Body Density} = \text{Mass in air} / \{ [\text{Mass in air} - \text{Mass in water} / \text{Water density}] - (\text{Residual lung volume} + \text{Residual gastrointestinal tract volume}) \}$$

Unseen is the weight of the sinker (weighted vest in this study) subtracted from the hydrostatic weight to obtain the true mass in water. Also, in a separate review of the literature and procedures by Buskirk (1961), residual volume in the gastrointestinal tract was found to vary greatly between subjects as well as within subjects on a daily basis, therefore, an assumed value of 100ml was used and this value was used in this study. This review by Buskirk (1961) also produced the method of measuring residual lung volume immediately after weight measurement. According to Buskirk (1961), "no change in lung volume should occur between the time the underwater weight is observed and the time the procedure for determining residual volume is started" (p. 100). This method was supported by observations of gradual weight increase with continued holding of the breath (Buskirk, 1961).

Of the many methods available to measure residual lung volume, the closed circuit Oxygen dilution method described by Wilmore, Vodak, Parr, Girandola, and Billing (1980) was used in this study. According to Wilmore et al. (1980), "the closed

circuit approach using the dilution principle appears to give the most reliable results and is the most efficient relative to the length of time it takes to conduct a duplicate determination of any one subject "(p. 216). This statement and new simplified method were validated by a comparison to the established method of oxygen dilution, which requires a Nitrogen analyzer and spirometer, and a Nitrogen washout method (Wilmore et al., 1980). The measured differences in volume between the three methods in this validation study were not found to be statistically significant and the new simplified method's "total procedure takes less than two minutes per test, or less than five minutes for duplicate tests."(Wilmore et al., 1980, p. 217). This method is also cost and space efficient because it "eliminates the need for a nitrogen analyzer and spirometer" (Wilmore et al., 1980, p. 218). Nitrogen content and ultimately, the volume of residual air in the lungs during hydrostatic weighing is calculated from the sum of the oxygen and carbon dioxide content measured from the rebreathing bag using the following equation (Wilmore et al., 1980):

- $RV = VO_2 \times (b - a) / (c - d)$

Where: RV = residual volume, VO₂ = volume of oxygen in the bag in the beginning of the procedure, a = percent nitrogen impurity of the original pure oxygen (assumed to be 0.0% for practical purposes), b = percent of nitrogen in the mixed air in the bag at the point of equilibrium [100 - (%O₂ + %CO₂)], c = percent nitrogen in the alveolar air at the beginning of the test (assumed to be 80%), d = percentage of nitrogen in the alveolar air during the last maximal breath (assumed to be 0.2% N₂ higher than the equilibrium percentage, i.e., b + 0.2% N₂). (p. 217)

According to Weir (2005), "reliability refers to the consistency of a test or measurement" (p. 231). In this study, a test-retest protocol was used including one instrument for measurement by a single observer on all subjects. This method eliminates the possibility of error from multiple instruments and multiple observers and places the entire focus on a single party to determine if consistency can be achieved. This test -

retest reliability of the Body Caliper as an instrument for measurement and the methods of a single observer was assessed by an intraclass correlation coefficient (ICC).

Regarding the interpretation of the results, Weir (2005) stated that "the ICC can theoretically vary between 0 and 1.0, where an ICC of 0 indicates no reliability and an ICC of 1.0 indicates perfect reliability" (p. 232).

To demonstrate the appropriateness of ICC regarding reliability determination, a study by Bedard, Martin, Krueger, and Brazil (2000) compared three methods of analysis including ICC, Pearson's r and paired t -test to determine the preferable method. In the study by Bedard et al. (2000), a hypothetical data set was created to compare the reliability estimates produced by the three methods. This data set was manipulated by the researchers to create situations that would "demonstrate good agreement, systematic bias or substantial random measurement error" (Bedard et al., 2000, p. 354). The situation of good agreement produced an estimate of reliability from all three methods (Bedard et al., 2000). The situation of systematic bias produced an estimation of reliability from Pearson's r only (Bedard et al., 2000). The situation of substantial random measurement error produced an estimation of reliability from the paired t -test only (Bedard et al., 2000). Therefore, the results of this study proved that the ICC was best able to detect bias or error, making it the preferred method for reliability determination (Bedard et al., 2000).

Validation studies that compare two methods of measurement are investigating whether one method is suitable to substitute or replace the other (Bland & Altman, 1986). According to Bland and Altman (1986), what these studies are searching for is agreement between the two methods of measurement. Statistical analysis of such studies often use

the Pearson product-moment correlation coefficient (r) to prove agreement; however, Bland and Altman (1986) stated that this method of analysis for this type of study is inappropriate:

r measures the strength of relationship between two variables, not the agreement between them. We will have perfect agreement only if the points lie along the line of equality, but we will have a perfect correlation if the points lie along any straight line. (p. 307)

In this regard, Bland and Altman (1986) followed up by stating that tests of significance are irrelevant when evaluating agreement; "the test of significance may show that the two methods are related, but it would be amazing if two methods designed to measure the same quantity were not related" (p. 308). Bland and Altman (1986) suggested the following method of analysis: First, plot the results of one method against the other, then draw a line of equality with an intercept of 0 and a rise of 1; this is where all points would rest if the two methods gave exactly the same results. Next, plot the combined mean of the two methods against the difference between them (Bland & Altman, 1986). They then indicated that one should calculate the mean and standard deviation of the differences between the two methods for the entire sample, then draw lines at ± 2 standard deviations from the mean to form the limits of agreement. According to Bland and Altman (1986), agreement occurs when a minimum of 95% of the sample rests within these limits. Due to the appropriateness of this method of analysis and the simplicity of interpreting the results, this study utilized Bland and Altman's (1986) method of analysis to determine the validity of the Body Caliper.

CHAPTER 3
METHODS

Subject Characteristics

A total of 100 adults aged 18 to 60 years volunteered to participate as subjects in this study. Of the 100 volunteers who participated, 62 were men and 38 were women.

Table 1

Characteristics of the Participants

	<i>Total</i> (<i>N</i> = 100)	<i>Female</i> (<i>n</i> = 38)	<i>Male</i> (<i>n</i> = 62)
Age (years)	23.64	21.66	24.85
	±8.52	±4.78	±10.005
Height (cm)	172.79	164.12	178.1
	±10.13	±8.15	±7.12
Weight (kg)	73.46	60.59	81.34
	±18.33	±12.63	±16.8

Values are means ±standard deviation

Instrumentation

Height was measured by a wall-mounted stadiometer, produced by Novell Products, Inc. Weight in air was measured by an electronic floor scale, produced by Toledo, model 2184 (floor scale) and model 8142 (digital display). The electronic floor scale was calibrated by a 10kg calibration weight. Skin-fold thickness was measured by the Body Caliper, produced by The Caliper Company, Inc. The manufacturer's calibration of the Body Caliper was examined by an independent study performed by the

investigator (Appendix A) and Body Caliper #8 was selected for use in this study. Hydrostatic weighing occurred in a fiberglass tank that is 6 feet long, 3 feet wide and 3 feet deep. Water temperature was measured by a thermometer. Two aluminum beams were mounted to the top of the tank, each were secured approximately 1 foot from either end lengthwise, from which a weighing platform with handles mounted on either side was suspended by three points of contact. Hydrostatic weight was measured by three load cells (one load cell at each point of contact) produced by Artech, model 20210-100. The weight was displayed on a digital display, produced by Toledo, model 8142 and recorded by a printer produced by Mettler Toledo, model APR310. A weighted vest was used as a sinker to keep the subject from floating off the weighing platform. Residual lung volume was measured using pure aviator-grade oxygen and a 5-liter anesthesia bag with a 2-way valve. A vacuum was used to remove all air from the anesthesia bag. The snorkel was made from a rubber mouthpiece that fit into a rubber tube that was 100cm long with a 3cm diameter and a 3-way "T" valve that allowed the subject to breathe either ambient air or from the anesthesia bag. A rubber nose-clip was used to close the nostrils off from air and water. Oxygen content was measured by a Servomex oxygen analyzer, model 570A. Carbon dioxide content was measured by an Anarad carbon dioxide analyzer, model 400. Pure nitrogen calibration gas was used to zero the oxygen and carbon dioxide analyzers and a calibration gas containing 13% oxygen and 9.9% carbon dioxide was used to finish calibration of the oxygen and carbon dioxide analyzers.

Collection of the Data

All subjects reported to the Exercise Physiology laboratory for testing. Upon completion of the informed consent form, each subject was issued a number for

identification and to protect their privacy, then their gender and age in years were recorded.

Height: wearing no shoes, the subject stood fully erect with their scapulae and heels against the wall and their eyes on a horizontal plane (Keys, 1956). The investigator lowered the arm of the stadiometer until it made contact with the top of the subject's head. The subject's height was recorded to the nearest 0.5cm.

Weight in air: the investigator calibrated the electronic floor scale at the beginning of each day of data collection prior to subjects' arrival. Wearing only a swimsuit, the subject approached, gently stepped up then stood still in the middle of the electronic floor scale with his or her arms relaxed at his or her sides. The subject's weight was recorded to the nearest 0.01kg.

Skin-folds: wearing only a swimsuit, the subject stood still while skin-fold thickness was measured at three sites on his or her body. The sites for men include the chest, abdomen and thigh (Jackson & Pollock, 1978). The sites for women include the suprailiac, triceps and thigh (Jackson et al., 1980). The YMCA protocol including site locations are listed below (Golding, 2000):

- Chest: A diagonal fold on the pectoral line midway between the axillary fold and the nipple.
- Abdomen: A vertical fold approximately 1.5 inches (3.81cm) to the right of the umbilicus.
- Suprailiac: A diagonal fold just above the crest of the ilium.
- Triceps: A vertical fold on the back of the upper arm midway between the shoulder and elbow joints.

- Thigh: A vertical fold on the front of the thigh midway between the groin line and the top of the patella.

All measurements were taken on the right side of the subject's body. Skin-folds were made by grasping the skin around the measurement site with both hands between the thumb and four fingers, then lifted away from the underlying muscle. The left hand remained holding the skin-fold while the right hand grabbed and operated the Body Caliper. The jaws were opened then perpendicularly placed at the base of the skin-fold close to (approximately 1cm) the left hand's thumb and index finger and allowed to close. When the needle on the measurement scale produced a static state (approximately 1-2 seconds), the skin-fold thickness was immediately read and recorded to the nearest millimeter. The jaws of the Body Caliper were removed and the skin-fold was released. Two trials with a maximum of three sets of measurements per trial were performed. Each trial was separated by 30 minutes. Sets of measurements were performed consecutively and cycled the skin-folds in the following order: chest, abdomen and thigh (men) or suprailiac, triceps and thigh (women). Two measurements of identical thickness at a specific site represented the value of that site for that trial. If three differing measurements of thickness were produced at a specific site, the median was taken to represent the value of that site for that trial. The values from the three sites for each trial were summed. Only the sum of skin-folds from trial 1 was entered into the Jackson & Pollock sum-of-3 equation to calculate body density.

Hydrostatic weighing & residual lung volume: The investigator calibrated the oxygen and carbon dioxide analyzers at the beginning of each day of data collection prior to subjects' arrival. The subject removed all body jewelry and wore only a swimsuit.

Prior to entering the tank, the subject showered to saturate their clothing, hair and skin, remove body oils and air trapped in the clothing, hair, and on the skin. While the subject showered, the investigator recorded the water temperature to the nearest degree Celsius, vacuumed then filled 3 anesthesia bags each with 5 liters of pure oxygen, zeroed the suspended weighing platform in still water, then immersed the weighted vest by placing it on the platform. Once all the air had released from the vest, its weight was recorded to the nearest 0.01kg. The investigator hung an anesthesia bag from the ceiling then attached it to its proper end of the 3-way T-valve. The subject was then allowed to enter the tank and sit on the platform. The investigator helped the subject place the weighted vest over his or her shoulders and buckle it across his or her chest and abdomen. The investigator helped the subject find the handles on either side of the platform that would help tether him or her to the platform. The investigator explained the subject's duties in the procedures for hydrostatic weighing and residual lung volume measurement then clarified any remaining questions. When the subject was ready, the nose-clip was placed on his or her nose and the mouthpiece was inserted into his or her mouth. The investigator checked that all of the subject's body was on the weighing platform, that the T-valve was open to ambient air, and asked the subject if he or she could breathe easily. The investigator instructed the subject to grab the handles then slowly lean forward to submerge their entire body under the water. Once submerged, the investigator coached the subject to remain submerged, continue breathing and move as little as possible while he waited for the water to become still and any remaining air bubbles from the vest, subject's swimsuit or hair to escape to the surface. When these conditions were met, the investigator instructed the subject to perform a maximal expiration and hold still in the

submerged position. Following the maximal expiration, the investigator directed the 3-way T-valve towards the anesthesia bag, which closed access to ambient air. The 2-way valve on the anesthesia bag remained closed, therefore, the subject was unable to breathe. The investigator observed the digital display and when the weight stopped fluctuating, recorded the subject's hydrostatic weight to the nearest 0.01kg. The investigator opened the 2-way valve on the anesthesia bag and instructed the subject to begin breathing from the bag. The subject was instructed to take 8-10 full breaths without breaking the seal on the mouthpiece or nose clip. The investigator counted the breaths then at the end of the final expiration, closed the 2-way valve on the anesthesia bag to capture the mixed air. The investigator removed the nose-clip and mouthpiece from the subject, detached the anesthesia bag from the snorkel then attached it to the oxygen and carbon dioxide analyzers to measure the content of oxygen and carbon dioxide of the mixed air. The air was forced through the oxygen analyzer by the investigator and the oxygen content was recorded to the nearest 0.1%. The air was pumped through the carbon dioxide analyzer and the carbon dioxide content was recorded to the nearest 0.1%. The subject's weight in air and hydrostatic weight, oxygen and carbon dioxide content of the mixed air, water temperature and density (5 decimals) at the recorded temperature, hydrostatic weight of the weighted vest and volume of pure oxygen in the anesthesia bag were entered into a computer program that calculated residual lung volume (3 decimals) and ultimately, body density (4 decimals) using the equations referenced in the literature review. Two more trials followed this same procedure and a mean was calculated from the three trials to serve as the value of body density of that specific subject.

Data Analysis Methods

A spreadsheet was made to organize the collected data and perform methods of analysis. Each row represented an individual subject and each column represented a separate category of data. The sum of skin-folds from trial 1 and trial 2 were used to determine the reliability of the Body Caliper and the investigator's methods of skin-fold measurement. Each trial was designated a column on a spreadsheet and the values for each trial's sum of skin-folds were entered into the proper column for all 100 subjects. The mean and standard deviation was calculated from all 100 subjects for each trial. All the sums of skin-folds from both trials were entered into a computer program that calculated the intraclass correlation coefficient (3 decimals). The mean and standard deviation of each trial was reported in a table. The intraclass correlation coefficient was also reported and will be the ultimate determination of reliability for this study.

The values for body density calculated from skin-fold measurement and hydrostatic weighing were assigned their own columns on the spreadsheet. The Bland-Altman analysis was used to assess agreement and determine the validity of the Body Caliper. First, the values of body density from each method for all subjects were plotted on a chart with hydrostatic weighing on the y-axis and skin-fold measurement by the Body Caliper on the x-axis. A line of equality with an equivalent intercept of 0 and a rise of 1 was drawn to display where all points would rest if there were perfect agreement between the two methods. Second, for each subject, body density measured by hydrostatic weighing was subtracted from body density measured by the Body Caliper. The remainder represented the difference between the two methods and these values were assigned their own column on the spreadsheet. The mean and standard deviation were

calculated from these values using all subjects, then a z -score was calculated for each subject and these values were assigned their own column on the spreadsheet. Using the z -scores, a chart was made to display the limits of agreement and assess agreement between the two methods. The y-axis of this chart represents the mean and surrounding units of standard deviation, the x-axis represents the number of the subject, and each plot represents the subject's z -score. The limits of agreement are represented by the solid lines drawn at the recommended ± 2 standard deviations around the mean. A minimum of 95% of the sample must rest between these limits to produce agreement and determine the Body Caliper to be valid.

CHAPTER 4

RESULTS

The investigation into the reliability of the Body Caliper produced the following results:

Table 2

Reliability of the Body Caliper

	<i>Total</i> (<i>N</i> = 100)	<i>Female</i> (<i>n</i> = 38)	<i>Male</i> (<i>n</i> = 62)
Sum of Skin-folds Trial 1 (mm)	49.71	55.21	46.34
	±22.68	±17.47	±24.88
Sum of Skin-folds Trial 2 (mm)	49.74	55.5	46.21
	±22.82	±16.92	±25.27

Note. Values are means ± standard deviation

The Body Caliper in conjunction with the YMCA protocol produced a high intraclass correlation coefficient: ICC= .997

The investigation into the validity of the Body Caliper produced the following results:

Body Density by 2 Methods

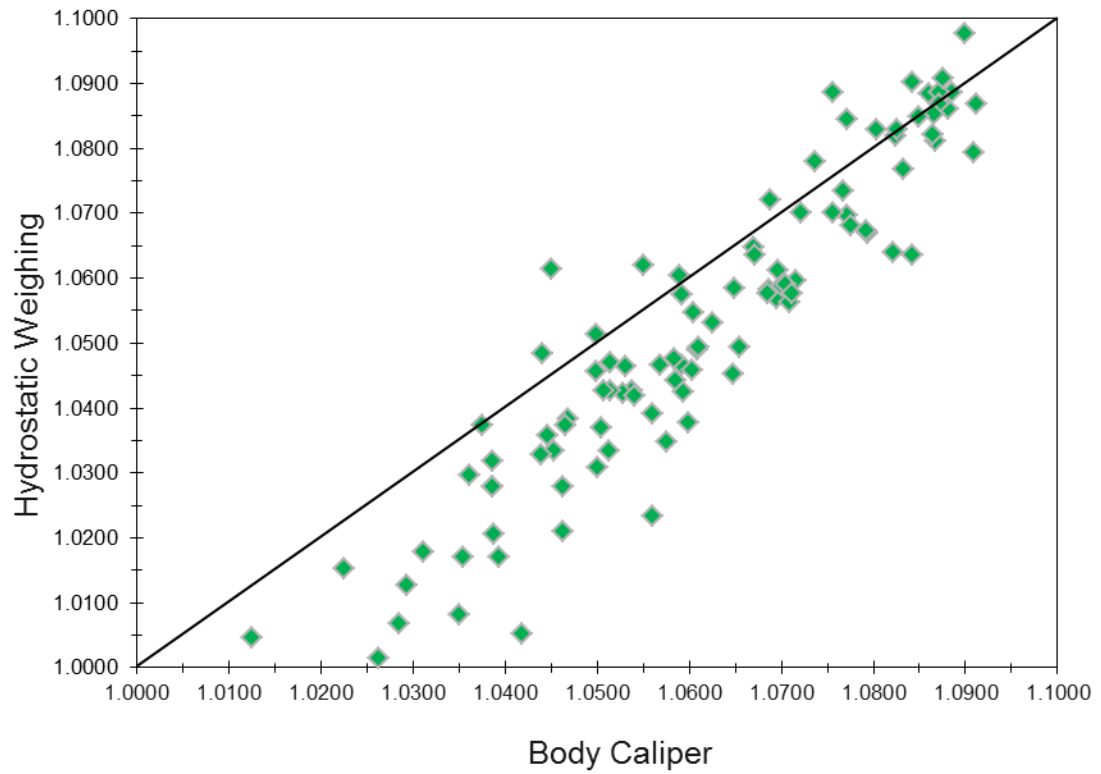


Chart 1. Body density measured by two methods.

Note. Values for body density are g/cc

As compared with hydrostatic weighing, 83% of the sample's body density was overestimated by the Body Caliper.

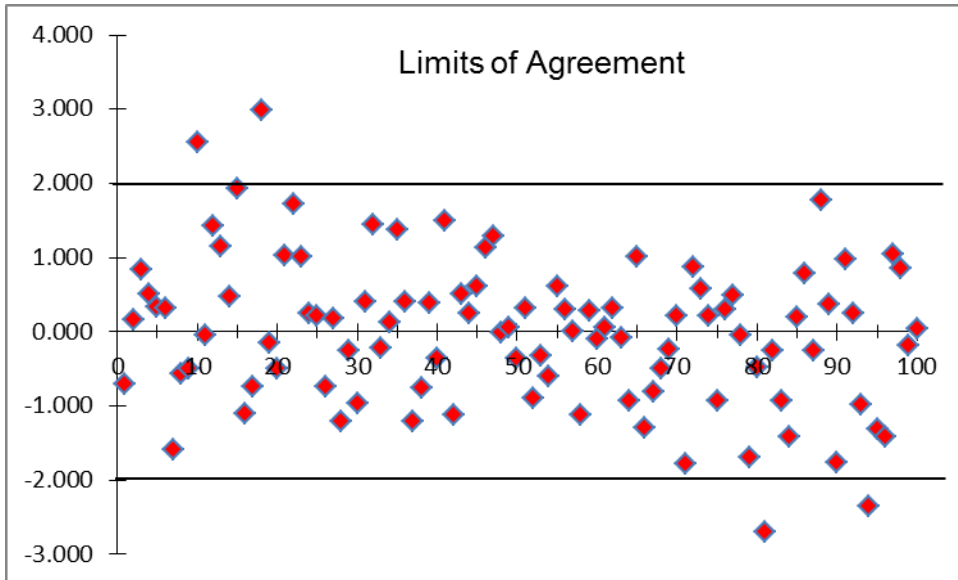


Chart 2. Limits of agreement.

Note. y-axis= mean and \pm units of standard deviation, x-axis= subject number

Ninety-six percent of the sample is located inside the limits of agreement.

Therefore, Bland-Altman analysis has generated an agreement between the two methods for measuring body density in adults.

CHAPTER 5
SUMMARY, CONCLUSIONS,
AND RECOMMENDATIONS

Discussion of Results

The results of the reliability investigation represent evidence that suggests that the Body Caliper in conjunction with the YMCA protocol is a reliable instrument and method for measuring skin-fold thickness in adults. This determination is supported by the small differences observed in the mean and standard deviation between the sum of skin-folds for trial 1 and trial 2. Females differed by only 0.29mm and 0.55mm, males differed by only 0.13mm and 0.39mm, and the whole sample differed by only 0.03mm and 0.14mm. The primary support of this reliability determination is the intraclass correlation coefficient, which was calculated from the entire sample and produced an ICC equal to 0.997. With 1.0 representing the top value and perfect reliability, an ICC equal to 0.997 serves as evidence of Body Caliper's reliability.

The results of the validity investigation represent evidence that suggests that the Body Caliper in conjunction with the YMCA protocol and the Jackson & Pollock sum-of-3 equations is a valid instrument and method for measuring body density in adults. Heterogeneity of the sample is supported by the wide range of body density measured by hydrostatic weighing (1.0014g/cc to 1.0976g/cc) and the Body Caliper (1.0126g/cc to 1.0912g/cc). Differences between the two methods ranged from 0.0001g/cc to 0.0366g/cc. The differences between the two methods produced a mean of 0.0089g/cc and standard deviation of ± 0.0093 g/cc. Although the majority (83%) of the sample's body density was overestimated by the Body Caliper, an agreement between the two

methods was produced and demonstrated by 96% of the sample's differences between the two methods being located within the required ± 2 standard deviations from the mean. This observed agreement between the two methods serves as evidence of the Body Caliper's validity.

Conclusions and Recommendations for Further Study

According to the results of this study, the investigator is able to conclude that the Body Caliper is a reliable and valid instrument for evaluating body composition in adults. However, the results of an independent study (Appendix A) by the investigator found the Body Caliper's jaw pressure to conform to the standard pressure of 10g/mm² up to 50mm of its 60mm range. Although no skin-fold measurement for any subject at any site in this study exceeded 50mm, the instrument's range of measurement was limited, and this limited the sample to participants with skin-folds equal to or less than 50mm. Therefore, a recommendation for further study would be to acquire a Body Caliper that exhibits the standard pressure throughout its range and investigate the reliability and validity of the Body Caliper at the end of its range that was not included in this study.

APPENDIX A

Mechanical Testing of the Body Caliper and Comparison to Four Popular Skin-fold Calipers

By

Gregory L. Stalker

Introduction and Purpose

The Body Caliper is a relatively new model of skin-fold caliper (approximately 13 years old) that measures skin-fold thickness to the nearest millimeter in a range of 0 – 60 millimeters (“The Body Caliper”). With the measurement scale printed on both sides, either right or left handed individuals may operate the Body Caliper with ease (“The Body Caliper”). The Body Caliper houses a patented mechanism, titled the "AccuSpring," that is claimed to deliver the National Research Council's standard jaw pressure of 10g/mm² (Keys, 1956) throughout its entire range of measurement, making it compatible with all skin-fold protocols (“The Body Caliper”). Manufacturers of the Body Caliper also claim that each caliper is calibrated by hand at the factory and never needs recalibration, offering a lifetime of accurate measurements (“The Body Caliper”).

To ensure accurate measurements, the calibration performed at the factory needs to be examined by a series of tests that will evaluate the distance and pressure between the jaws of each Body Caliper used in this study (Heyward & Wagner, 2004). There also needs to be a comparison of skin-fold measurements made by the Body Caliper to other

calipers that claim to conform to the standards of the National Research Council.

Therefore, the purpose of this study is to determine whether the Body Caliper can support its claim of being a research-grade instrument and to test its performance against other popular skin-fold calipers, two of which (Harpenden and Lange) have been proven research-grade by their use in research studies for decades (Golding, 2000).

Instrumentation

1. Body Caliper – quantity: 10
2. Vernier Caliper – brand: Westward, model: 4KU77
3. Steel Vice
4. Sling: cotton cloth and nylon cord
5. 100 gram weight – quantity: 2
6. 10 gram weight – quantity: 15
7. Rubberized-foam pad – thickness: 1cm
8. Lange Caliper
9. Harpenden Caliper
10. Lafayette Caliper
11. Generic Plastic Caliper (no mechanism) – similar to Ross Adipometer illustrated on page 62 of Heyward & Wagner, 2004.

Methods

In this cross-sectional study, 10 new (in original manufacturer's packaging) Body Caliper skin-fold calipers were purchased. Once opened, each was assigned a number (1-10) and labeled with a permanent marker, then underwent a visual inspection for damage or defects.

The surface area of each Body Caliper's jaws was measured with a vernier caliper. While a Body Caliper was secured in a work-bench mounted steel vice, the jaws of the vernier caliper were closed around all sides of the Body Caliper's jaws to measure the length and width. Both sides of the Body Caliper's jaws were measured and recorded in millimeters.

The accuracy of the measurement scale on both sides of each Body Caliper was tested with a vernier caliper (Heyward & Wagner, 2004). First, each Body Caliper underwent a visual inspection to examine if the needle read zero while the jaws were closed. Next, the vernier caliper was set at distances of 10, 20, 30, 40, 50 and 60 millimeters. The test began at 10mm and progressed to 60mm. At each distance, the jaws of each Body Caliper were opened and allowed to close over the jaws of the vernier caliper. The measurement scale on each Body Caliper was read and recorded to the nearest millimeter.

The jaw pressure of each Body Caliper was examined by a static test and a dynamic test (Heyward & Wagner, 2004). The static test began by securing a Body Caliper upright in a work-bench mounted steel vice. A sling made of cotton cloth and nylon cord was loaded with an exact weight of 280 grams, representing the National Research Council standard pressure of 10g/mm² (Keys, 1956). The jaws of the Body Caliper were opened to distances of 10, 20, 30, 40, 50 and 60 millimeters. The first measurement took place at 10mm and progressed to 60mm. While the jaws were held open at each distance, the sling was placed over the bottom jaw and released. The weight hung 50cm below the bottom jaw. Any retraction or protraction of the bottom jaw was observed. If the bottom jaw remained static, that specific Body Caliper passed for that

distance. If movement occurred, that specific Body Caliper failed for that distance. After all calipers were graded pass or fail at each distance, the calipers were reassessed at their failed distances to measure the exact jaw pressure by loading the sling with weight until the bottom jaw remained static. The weight that produced a static state was recorded in grams and the jaw pressure was calculated by dividing that weight by the jaw's surface area of 28mm².

The dynamic test measured the Body Caliper's rate of compression on rubberized-foam at uncompressed thicknesses of 10, 20, 30, 40, 50 and 60 millimeters. Due to an unknown compression ratio of the rubberized-foam used in this study, this form of testing was used only to validate static testing by assessing the caliper's uniformity in rates of compression at the distances where each caliper passed the static test. Starting at 10mm and progressing to 60mm, the jaws of each Body Caliper were opened, placed over the foam and released. When the needle on the measurement scale produced a static state, the thickness of the rubberized-foam was immediately read and recorded to the nearest millimeter. Compression rates were calculated by subtracting the thickness measured by the Body Caliper from the foam's uncompressed thickness at the start of the test. The foam's uncompressed thickness was 1cm, therefore, layers were stacked to create the varying thicknesses up to 60mm. For each test of compression, the caliper's jaws were placed over a fresh surface. At the end of the test, a 1 inch square surrounding the used surface was marked with a permanent marker and avoided during the remaining tests. Failed distances of static testing were also assessed dynamically in an attempt to validate the increased jaw pressures observed during static testing by measuring the compression rates of the rubberized-foam at those failed distances.

10 men and 10 women participated in the comparison of the Body Caliper to 4 other skin-fold calipers. Skin-fold measurements were made following the YMCA protocol and at the sites required by the Jackson & Pollock sum-of-3 equations. The sites for men include the Chest, Abdomen and Thigh (Jackson & Pollock, 1978). The sites for women include the Suprailiac, Triceps and Thigh (Jackson, Pollock, & Ward, 1980) The YMCA protocol including the site locations are described below (Golding, 2000):

Chest: A diagonal fold on the pectoral line midway between the axillary fold and the nipple.

Abdomen: A vertical fold approximately 1.5 inches (3.81cm) to the right of the umbilicus.

Suprailiac: A diagonal fold just above the iliac crest.

Triceps: A vertical fold on the back of the upper arm midway between the shoulder and elbow joints.

Thigh: A vertical fold on the front of the thigh midway between the groin line and the top of the patella.

All measurements were taken on the right side of the subject's body. Skin-folds were made by firmly grasping the skin around the measurement site with both hands between the thumb and four fingers, and then lifted away from the underlying muscle. The left hand remained holding the skin-fold while the right hand grabbed and operated the caliper. Once the caliper's jaws were opened, they were placed perpendicularly at the base of the skin-fold close to (approximately 1cm) the left hand's thumb and index finger and allowed to close. When the needle on the measurement scale produced a static state (approximately 1-2 seconds) the thickness was immediately read and recorded in whole

millimeters (no decimals), then the caliper's jaws were removed. For each site, the skin-fold was lifted and held while the measurements of all 5 different calipers were taken in the following order before it was released: Body Caliper, Lange, Harpenden, Lafayette and Generic Plastic. Three sets of measurements were taken and each set cycled the skin-folds in the following order: chest, abdomen and thigh (men) or suprailiac, triceps and thigh (women). Two measurements of identical thickness by a specific caliper at a specific site represented the value for that caliper at that site. If three differing measurements were produced by a specific caliper at a specific site, the median was taken to represent the value for that caliper at that site.

Statistical analysis of this comparison was performed using the Pearson product moment correlation coefficient (r). A test of significance with a p -value set at 0.05 was used to determine significant correlations.

Results

All 10 Body Calipers arrived in the same packaging and once opened, appeared to be new products direct from the manufacturer. Measurement of the jaws' surface area yielded a uniform 7mm of length by 4mm of width, totaling 28mm² for all calipers. The visual inspection for damage or defects found Body Caliper #5 to have arrived damaged with a 4cm crack in the housing that holds the "Accuspring" mechanism. Therefore, Body Caliper #5 was omitted from further investigation.

Measurement of the distance between the jaws to evaluate the scale accuracy of the Body Calipers with the vernier caliper produced an agreement of all Body Calipers with the vernier caliper at all distances. Also, the needle of all Body Calipers was in perfect alignment with and read 0mm on both sides while the jaws were closed.

Static testing of jaw pressure found none of the calipers to conform to the National Research Council's standard pressure of $10\text{g}/\text{mm}^2$ throughout its scale of 0 – 60mm. All failures were due to retraction of the jaws, indicating that the pressure exceeded the standard. Only 6 calipers, Body Caliper #'s 1, 2, 4, 8, 9 and 10 conformed to the standard through 50mm. Body Caliper #3 conformed through 30mm, required 320g ($11.43\text{g}/\text{mm}^2$) to remain static at 40mm and 340g ($12.14\text{g}/\text{mm}^2$) to remain static at 50mm. Body Caliper #6 conformed through 40mm and required 310g ($11.07\text{g}/\text{mm}^2$) to remain static at 50mm. Body Caliper #7 conformed through 30mm, required 320g ($11.43\text{g}/\text{mm}^2$) to remain static at 40mm and 350g ($12.5\text{g}/\text{mm}^2$) to remain static at 50mm. The heaviest weight used in this study was 350g and no caliper was able to remain static under that load at 60mm, therefore, the exact pressure at 60mm was unattainable.

Dynamic testing of jaw pressure validated the results of static testing. When a Body Caliper passed the static test by holding 280g, matching the standard pressure of $10\text{g}/\text{mm}^2$, it compressed the rubberized-foam 3mm. When a Body Caliper failed the static test, requiring a heavier load to produce a static state, the observed increase in jaw pressure resulted in increased compression of the rubberized-foam: Body Caliper #3 compressed 40mm of foam 4mm and 50mm of foam 5mm, Body Caliper #6 compressed 50mm of foam 4mm, Body Caliper #7 compressed 40mm of foam 4mm and 50mm of foam 6mm. Due to static testing being unable to measure jaw pressure at 60mm, this distance was omitted from investigation by dynamic testing.

Table 3

Comparison of the Body Caliper to Four Skin-fold Calipers

Site of Skin-fold	Body Caliper vs. Lange	Body Caliper vs. Harpenden	Body Caliper vs. Lafayette	Body Caliper vs. Generic Plastic
Chest	r= .996 p< .01	r= .998 p< .01	r= .995 p< .01	r=.995 p< .01
Abdomen	r= .992 p< .01	r= .991 p< .01	r= .991 p< .01	r= .997 p< .01
Thigh men	r= .996 p< .01	r= .996 p< .01	r= .993 p< .01	r= .997 p< .01
Suprailiac	r= .997 p< .01	r= .996 p< .01	r= .998 p< .01	r= .999 p< .01
Triceps	r= .999 p< .01	r= .991 p< .01	r= .997 p< .01	r= .995 p< .01
Thigh women	r= .997 p< .01	r= .996 p< .01	r= .996 p< .01	r= .997 p< .01

Note. Values are Pearson correlation (r) and test of significance (p-value)

Body Caliper #8 was randomly selected from the group that conformed to the standard jaw pressure of 10g/mm² through 50mm as the caliper to be used in this comparison. Statistical analysis of this comparison produced significant, high correlations for the Body Caliper with all the other calipers at all sites.

Discussion

According to the results of the mechanical testing performed on 10 new Body Calipers, the investigator is able to make the following implications:

- The manufacturer is capable of producing a uniform product, evidenced by the measured consistency of the jaw's surface area at 28mm². However, there is a 10% chance that a Body Caliper will arrive damaged or with a defect.

Considering that this product comes with a lifetime satisfaction guarantee against defects, this issue is simply a matter of inconvenience that consumers should be able to resolve.

- The printed measurement scale makes accurate readings of distance between the jaws across its entire range on both sides of the caliper.
- The claim of a consistent jaw pressure that conforms to the standard 10g/mm² throughout its range of 0 – 60mm went unsupported. Only two-thirds of the calipers tested demonstrated the standard jaw pressure up to 50mm. Therefore, it is recommended that skin-folds thicker than 50mm not be measured with the Body Caliper. Also, the factory's calibration of a new Body Caliper should be examined for accuracy and consistency prior to use.

According to the results of the comparison in measurements made by the Body Caliper and the 4 other skin-fold calipers, the investigator is able to conclude that the Body Caliper is a suitable replacement for any of the other skin-fold calipers used in this study to provide measurements of skin-fold thickness up to 50mm.

APPENDIX B



Biomedical IRB – Expedited Review Approval Notice

NOTICE TO ALL RESEARCHERS:

Please be aware that a protocol violation (e.g., failure to submit a modification for any change) of an IRB approved protocol may result in mandatory remedial education, additional audits, re-consenting subjects, researcher probation, suspension of any research protocol at issue, suspension of additional existing research protocols, invalidation of all research conducted under the research protocol at issue, and further appropriate consequences as determined by the IRB and the Institutional Officer.

DATE: December 15, 2011

TO: Dr. Lawrence Golding , Kinesiology and Nutrition Sciences

FROM: Office of Research Integrity - Human Subjects

RE: Notification of IRB Action by /Charles Rasmussen/ Dr. Charles Rasmussen, Co-Chair
Protocol Title: **Reliability and Validity of the Body Caliper to Evaluate Body Composition**
Protocol #: 1112-3986M
Expiration Date: December 14, 2012

This memorandum is notification that the project referenced above has been reviewed and approved by the UNLV Biomedical Institutional Review Board (IRB) as indicated in Federal regulatory statutes 45 CFR 46 and UNLV Human Research Policies and Procedures.

The protocol is approved for a period of one year and expires December 14, 2012. If the above-referenced project has not been completed by this date you must request renewal by submitting a Continuing Review Request form 30 days before the expiration date.

PLEASE NOTE:

Upon approval, the research team is responsible for conducting the research as stated in the protocol most recently reviewed and approved by the IRB, which shall include using the most recently submitted Informed Consent/Assent forms and recruitment materials. The official versions of these forms are indicated by footer which contains approval and expiration dates.

Should there be *any* change to the protocol, it will be necessary to submit a **Modification Form** through ORI - Human Subjects. No changes may be made to the existing protocol until modifications have been approved by the IRB. Modified versions of protocol materials must be used upon review and approval. Unanticipated problems, deviations to protocols, and adverse events must be reported to the ORI – HS within 10 days of occurrence.

If you have questions or require any assistance, please contact the Office of Research Integrity - Human Subjects at IRB@unlv.edu or call 895-2794.

Office of Research Integrity - Human Subjects
4505 Maryland Parkway • Box 451047 • Las Vegas, Nevada 89154-1047
(702) 895-2794 • FAX: (702) 895-0805

REFERENCES

- Astorino, T. A., Orri, J., Lockner, D., Jenkins, K., & Heyward, V. (1999) Comparison between the Body Caliper and Harpenden caliper for measuring skinfold thickness in children. *Medicine and Science in Sports and Exercise*, 31(5). Supplement, S204.
- Bedard, M., Martin, M. J., Krueger, P., & Brazil, K. (2000) Assessing reproducibility of data obtained with instruments based on continuous measurements. *Experimental Aging Research*, 26, 353-365.
- Behnke Jr., A. R, Feen, B. G, Welham, W. C. (1942). The specific gravity of healthy men: Body weight/volume as an index of obesity. *The Journal of the American Medical Association*, 118(7), 495-498.
- Bland, J. M., & Altman, D. G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1, 307-310.
- Buskirk, E. R. (1961). Underwater weighing and body density: A review of procedures. In J. Brozek & A. Henschel (Eds.), *Techniques for measuring body composition* (pp. 90-106) Washington, DC: National Academy of Sciences–National Research Council
- Golding, L. A. (Ed.) (2000). *YMCA fitness testing and assessment manual* (4th ed.). Champaign, IL: Human Kinetics.
- Goldman, R. F, & Buskirk, E. R. (1961). Body volume measurement by underwater weighing: Description of a method. In J. Brozek & A. Henschel (Eds.). *Techniques for measuring body composition* (pp. 78-89). Washington, DC: National Academy of Sciences–National Research Council.

- Heyward, V. H., & Wagner, D. R. (2004). *Applied body composition assessment* (2nd ed). Champaign, IL: Human Kinetics.
- Hume, P., & Marfell-Jones, M. (2008) The importance of accurate site location for skinfold measurement. *Journal of Sports Sciences*, 26(12), 1333-1340.
- Jackson, A. S., Pollock, M. L. (1978). Generalized equations for predicting body density of men. *British Journal of Nutrition*, 40, 497-504.
- Jackson, A. S, Pollock, M. L., & Ward, A. (1980). Generalized equations for predicting body density of women. *Medicine and Science in Sports and Exercise*, 12(3), 175-182.
- Keys, A. (1956). Recommendations concerning body measurements for the characterization of nutritional status. Committee on Nutritional Anthropometry, Food and Nutrition Board, National Research Council. *Human Biology*, 28(2), 111-123.
- Scherf, J., Franklin, B. A., Lucas, C. P., Stevenson, D., & Rubenfire, M. (1986). Validity of skinfold thickness measures in formerly obese adults. *The American Journal of Clinical Nutrition*, 43, 128-135.
- The body caliper. (n.d.). Retrieved from <http://www.bodycaliper.com>
- Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, 19(1), 231-240.
- Wilmore, J. H., Vodak, P. A., Parr, R.B., Girandola, R.N., & Billing, J. E. (1980). Further simplification of a method for determination of residual lung volume. *Medicine and Science in Sports and Exercise*, 12(3), 216-218.

VITA

Graduate College
University of Nevada, Las Vegas

Gregory L. Stalker

Home Address:

2609 Gilmary Ave.
Las Vegas, NV 89102

Degrees:

Bachelor of Arts, 2001
University at Albany, State University of New York

Thesis Title:

Reliability and Validity of the Body Caliper to Evaluate Body Composition

Thesis Examination Committee:

Chair, Lawrence A. Golding, Ph. D.
Committee Member, Richard D. Tandy, Ph. D.
Committee Member, John Mercer, Ph. D.
Graduate College Representative, Merrill Landers, D.P.T.