Comparison of Body Composition Measurements using the Lafayette Instrument Skinfold Caliper (model 01128), Traditional Calipers, and Other Methods

Abstract

Purpose: To determine body composition measurements using a newly developed skinfold caliper (Lafayette Instruments Skinfold II) compared to body composition measurements made using four well-validated methods: two different commercially available calipers Harpenden and Lange, hydrostatic weighing (UWW), and body plethysmography (BP). A second purpose was to determine whether body fat measures made by experienced (EX) and inexperienced (IX) technicians were similar when using different calipers. **Methods:** Skinfold measurements were performed by an EX and IX technician on 21 younger (21.2 +/- 1.5 yrs) and 20 older (59.2 +/- 4 yrs) subjects. Body fat percentage was calculated using the Jackson-Pollock seven-site formula. HW and BP tests were performed on a subset of the subjects (10 younger, 21.5 +/- 1.7 yrs; 10 older, 59.2 +/- 4.7 yrs). Statistical significance was determined a priori at alpha = 0.05. HW and BP tests were performed on a subset of the subjects (10 younger, 21.5 +/- 1.7 yrs; 10 older, 59.2 +/- 4.7 yrs). Reliability of body fat percentages for the five methods was compared using Pearson correlations. Correlations were compared using a Fishers Z-transformation and subsequently tested via Z-test. Differences between groups were determined using one-way ANOVA and Tukey HSD post hoc tests when appropriate. Effect of experience level on ability to predict HW and BP was assessed using logistic regression. **Results**: There were no significant differences between body fat measurements when comparing Lafayette Instruments EX to the other calipers (F = 2.90, p = 0.06). Furthermore, the Lafayette InstrumentsEX measurements differed by only

2.3% from Harpenden EX and Lange EX (p = .07) and were highly correlated to both Harpenden EX (r= 0.99, p<0.01) and Lange EX (r=0.99, p<0.01). There were no significant differences in the Harpenden W/BP subgroup between Lafayette Instruments EX and UWW (p = 0.111) or BP (p = 0.138) but all calipers tended to underestimate body fat compared to UWW and BP. Differences of less than 3% found between EX and IX did not help to explain additional variance in the model that was practically useful. **Conclusions:** Body fat percentages determined with Lafayette Instruments calipers by EX were similar to those from Harpenden EX and Lange EX . In a subset of subjects, body fat results from Lafayette Instruments EX were similar to UWW or BP. Performance by IX was comparable to EX and suggests that there was a similar ease of use for all three calipers.

Introduction

Body fat assessment allows for the quantification of the major body compartments: fat free mass and fat mass. Body fat assessment is an important tool for fitness professionals and can provide an indicator of health and health risk. There are several available methods for body fat estimation which include skinfold calipers, hydrostatic weighing, DEXA, and air-displacement plethysmography.

Lafayette Instruments manufactured an inexpensive skinfold caliper (Lafayette Skinfold II, Model Number 01128) that compared favorably, in preliminary measurements from our laboratory, to the more expensive Lange Skinfold calipers (Cambridge Scientific Industries, Inc, Cambridge, MD). The Lafayette Skinfold II did not compare favorably to the Harpenden Caliper (British Indicators LTD, Great Britian), but these experiments were performed on a small number of highly fit, young male and female subjects.

Valid determinations of body fat percentage are possible using skinfold calipers in men (Jackson & Pollock, 1978) and women (Jackson et al., 1980) with multiple correlations, to underwater weighing, exceeding 0.90 in men and ranging from 0.842 to 0.867 in women. For example, Jackson et al. (1980) reported that body composition determined using skinfolds was strongly correlated (r=0.82) with body fat determined using underwater weighing. Skinfold determinations can be made using several different equations, including a three-site, four-site, or seven-site skinfold formula (Jackson & Pollock, 1978).

Fields et al. (2002) completed a systematic review and reported that body fat determined using air-displacement plethysmography (BodPod ™) and hydrostatic weighing agreed to within determine body 1.0% for groups of adults and children (Fields et al., 2002). When wide variation was reported between these methods, the authors attributed this variation to deviations from accepted protocol, differences in equipment and researcher error. Thus, the air-displacement method is gaining acceptance as a method of body composition measurement, but there is enough variation to suggest a cautious interpretation of the results. Nevertheless, the two methods provide an adequate basis for comparison of the skinfold results.

Purpose:

This study was conducted in order to compare the Lafayette Skinfold II to other commercially available calipers (Lange and Harpenden) in a subject population that included subjects with different fitness levels and from different age groups. Further, this study attempted to determine the validity of the Lafayette Instruments II caliper in a sub-sample of subjects compared to UWW and BodPod (BP). Finally, we sought to compare the results of Lafayette Skinfold II by both experienced (EX) and inexperienced (IX) technicians.

Hypotheses:

I. There will be no differences in body fat percentages determined using the three commercially available calipers.

II. Body fat determined using the Lafayette Instruments II will be strongly correlated with hydrostatic weighing and air-displacement plethysmography.

III. There will be no differences in body fat determinations made by inexperienced or experienced technicians.

Methods

Subjects:

Forty-one apparently healthy subjects were recruited from the university student body and surrounding community into one of four groups: 18-25 years of age sedentary (YPI), 18-25 physically active (YPA), 55-70 sedentary (OPI) and 55-70 physically active (OPA). The groups were further subdivided to contain both male (M) and female (F) subjects (Table 1). The final groupings included YPIF (n = 5), YPIM (n = 5), YPAF (n = 6), YPAM (n = 5), OPIF (n = 5), OPIM (n = 5), OPAF (n = 5), and OPAM (n = 5). The Paffenbarger physical activity questionnaire (Paffenbarger, Blair et al. 1993) was used to classify subjects as sedentary or physically active. In addition, those in the physically active group self-reported a minimum of three days of endurance or resistance exercise per week during the previous six months.

Table 1. Age of Subjects by Subgroups (N = 41).

			Age (years)	Std. Dev.
Overall			39.67	20.42
	Older		59.35	3.99
		Male (n=10)	60.00	3.50
		Female (n=10	58.70	4.52
		Inactive (n=10)	58.40	3.31
		Active (n=10)	60.30	4.55
	Young			
		Male (n=10)	21.00	1.83
		Female (n=11)	21.36	1.29
		Inactive (n=10)	21.50	1.58
		Active (n=11)	20.91	1.51

Note: Age of subjects are presented in years ± standard deviation.

Procedures:

Both an experienced and inexperienced technician measured body composition using Lafayette Skinfold II, Lange and Harpenden calipers on each subject. Inexperienced technicians had not made skinfold measurements before, with the exception of minimal experience in a previous exercise physiology class lab session. Experienced technicians were required to have used skinfold measurements as part of their work or their research. Due to the volume of measurements made in this study, different technicians were employed, which kept the inexperienced technicians from completing a sufficient number of tests to be considered experienced. The inexperienced technicians were each given brief instructions on skinfold sites and techniques for measuring skinfolds.

Measurements were made on a single day between 0600-1000 h. Subjects did not eat after 2200h the night before testing, refrained from exercise for at least the previous 12 hours, were encouraged to drink water in the 24 hours preceding the test, and abstain from alcohol for the day preceding the test. Both a seven-site and three-site skinfold formula (Jackson and Pollock 1978; Jackson, Pollock et al. 1980) were applied to the test data for each of the calipers in order to determine body density. The Siri (1956) formula was used to convert density to a body fat percentage.

A sub-group of subjects also completed under-water weighing (UWW) and air displacement plethysmography (BodPod, Life Measurement Inc.). Under-water weighing uses water displacement to determine body density and is considered the long-standing gold standard. In a large tank, a subject sat on an underwater scale. The subjects were instructed to exhale completely, submerge their head, and stay as still as possible. Load cells interfaced with a computer allowed weight in water to be sampled over a three second period and an average to be calculated. An estimate of residual volume was made based on the subject's gender, height, and age. Body density was calculated using standardized equations to compare the underwater weight to the subject's dry weight. The Siri equation (Siri, 19556) was used to convert body density measurements to percent body fat. Several tests were performed on each subject to allow them to become acclimated to the procedure. The test was repeated 7-12 times until body composition values leveled off. The final three trials were averaged to arrive at a final percent body fat.

The BodPod™ utilizes air displacement plethysmography to estimate body fat percentages. The test was completed twice for each subject. In cases of a significant between test disparity, the test was completed a third time. The instrument estimated tidal volume during the last stage of each test and this value was employed by the software to calculate body fat. The BodPod software uses standard equations to calculate body density from the volume of air displaced and the subject's body mass taken from an attached scale. The Siri equation was also used to calculate percent body fat from body density.

Data Analysis:

Analysis was performed using SPSS 15.0 for Windows statistical analysis package. An a priori significance level of $\alpha = .05$ was set for all statistical tests. An omnibus multivariate ANOVA was performed to determine whether there were any significant differences between calipers. Validation of the Lafayette Instruments caliper was determined via correlation analysis of body fat measured by all three calipers. Ease of use was determined via correlation analysis of experienced versus inexperienced practitioners within each caliper determination of percent body fat. Further analysis included logistic regression to examine whether addition of alternate experience level improved the regression prediction of body fat determined by hydrostatic weighing. Reliability was examined in the three calipers by comparing correlations of each caliper with percent body fat determined by underwater weighing and air displacement plethysmography.

To find differences in correlation coefficients, a Fishers Z-transformation and subsequent Z-test were performed. Highly correlation coefficients are often not normally distributed (Meng et al., 1992), but transforming the distribution to a Z distribution allows for a simple Z test to determine if the coefficients differ significantly.

Results

Prior to analysis, the percent body fat determinations using the different site formulas in experienced an inexperienced practitioners were screened for missing values, accuracy and the assumptions of normality, independence and homogeneity of variance. Z-scores revealed no univariate outliers. No violations of assumptions were found in ungrouped or grouped data except for an expected violation of kurtosis in the age of subjects that is attributable to study design. Missing data for one subject accounted for 2.3 % of the sample. The missing data from an inexperienced practitioner was imputed via "hot deck" imputation at the skinfold site level. Group means changed very little with the imputation and no change in significance of ANOVA F score resulted.

Multivariate ANOVA revealed an insignificant omnibus F-test ($F = 2.90$, $p = 0.06$). Further analysis was warranted due to the trend toward significance in the omnibus test. Post-hoc pairwise comparisons using Tukey HSD show mean differences in body fat estimates ranging from 0.38% to 2.30% with a standard error of 1.03% (Table 2). Caliper means were similar for Harpenden (20.05%) and Lafayette Instruments (20.42%), but the Lange calipers provided the highest body fat estimates (22.35%). Further ANOVA was not performed using sub-groups or number of site formulas.

Table 2.

Differences in caliper means by caliper type determined by Tukey HSD.

Type 1	Type 2	Mean Difference	P value
Lafayette Instruments	Harpenden	0.38	0.93
Lafayette Instruments	Lange	-1.93	0.15
Harpenden	Lange	-2.30	0.07

Caliper means determined by three-site, four-site, and seven-site formulas with skinfolds measured by experienced and inexperienced practitioners.

Pearson Product-Moment correlational analysis revealed strong correlations (r = 0.89- 0.99) between the three caliper-determined body fat estimates in both experienced and inexperienced practitioners. Greater number of skinfold sites used in the formulae yielded more perfectly correlated body fat estimates (Tables 3a,3b, & 3c). Inexperienced practitioners did not have substantially different Pearson r values compared to experienced practitioners. However, further investigation revealed significant differences in correlations between experienced and inexperienced practitioners when performing Fisher's Z-transformation with a Z-test. Correlations for all caliper comparisons for experienced and inexperienced were significantly different when the seven-site equations were used (Table 4). However, there were no significant differences among caliper correlation between experienced and inexperienced when the three-site or four-site equations were used (Table 4).

Table 3a.

Correlation Statistics Comparing Body Fat Estimates Determined by Lafayette Instruments, Harpenden and Lange Calipers (N=41)Using Jackson-Pollock seven-site Formula.

Caliper Type	Lafayette Instruments	Harpenden	Lange
Lafayette Instruments	-	0.97	0.98
Harpenden	0.99	-	0.98
Lange	0.99	0.99	-

Note: Top half of matrix gives Pearson r for inexperienced practitioners while bottom half of matrix gives Pearson r for experienced practitioners.

Table 3b.

Correlation Statistics Comparing Body Fat Estimates Determined by Lafayette Instruments, Harpenden and Lange Calipers (N=41)Using Jackson-Pollock three-site Formula.

Caliper Type	Lafayette Instruments	Harpenden	Lange
Lafayette Instruments	-	0.97	0.98
Harpenden	0.99	-	0.98
Lange	0.99	0.99	-

Note: Top half of matrix gives Pearson r for inexperienced practitioners while bottom half of matrix gives Pearson r for experienced practitioners.

Table 3c.

Correlation Statistics Comparing Body Fat Estimates Determined by Lafayette Instruments, Harpenden and Lange Calipers (N=41) Using Modified Sinning four-site Formula.

Caliper Type	Lafayette Instruments	Harpenden	Lange
Lafayette Instruments	-	0.97	0.98
Harpenden	0.99	-	0.98
Lange	0.99	0.99	-

Note: Top half of matrix gives Pearson r for inexperienced practitioners while bottom half of matrix gives Pearson r for experienced practitioners.

The significant differences in the Z-scores can be attributable to the almost perfect correlations seen between the calipers. When variables correlate almost perfectly, an extremely small difference in Pearson's r can yield Z-test values that are significant (Meng et al., 1992). This is the case with the seven-site formula body fat determination in this population.

Table 4. Fisher's Z-transformation and Z-test Scores for Caliper Correlations Comparing Experienced and Inexperienced Practitioners.

Caliper Correlation		**seven-site**	
	Experienced Z score	Inexperienced Z score	Z test
Lafayette Instruments & Harpenden	2.65	2.09	-3.42 (p < .01)
Lafayette Instruments & Lange	2.65	2.30	-2.15 (p = .03)
Harpenden & Lange	2.65	2.30	-2.15 (p = .03)
		three-site	
	Experienced Z score	Inexperienced Z score	Z test
Lafayette Instruments & Harpenden	1.66	1.66	0.00 (p = 1.00)
Lafayette Instruments & Lange	1.59	1.59	0.00 (p = 1.00)
Harpenden & Lange	1.53	1.42	-0.65 (p = .52)
		four-site	
	Experienced Z score	Inexperienced Z score	Z test
Lafayette Instruments & Harpenden	2.09	1.83	-1.61 (p = .11)
Lafayette Instruments & Lange	2.09	1.95	-0.90 (p = .37)
Harpenden & Lange	2.09	2.09	0.00 (p = 1.00)

Estimates of body fat in the total sample when determined via BodPod and UWW were strongly correlated (r = .92, p < .01). The correlations between calipers and BodPod (r = .77 - .90) were stronger than correlations between calipers and UWW (r = .70 - .85). This suggests an underestimation of body fat percentage by all calipers. Mean BF for UWW and BP were 25.5% ± 9.5% and 25.2% ± 10.4% respectively. Caliper estimates ranged from 19.3% to 23.3% with standard deviation range from 5.6% - 7.2%. Seven-site determinations for both inexperienced (r = .83 - .85) and experienced (r = .82 - .83) technicians yielded the strongest correlations with UWW. Similar effects were seen with seven-site determinations and BodPod (Table 5.)

Table 5.

Experienced and Inexperienced Practitioner Caliper Correlations With Underwater Weighing and BodPod Using Three Formalae. (N = 20).

		seven-site Exp	seven-site Inexp	three-site Exp	three-site Inexp	four-site Exp	four-site Inexp
Underwater Weighing	Lafayette Instruments	0.82	0.85	0.71	0.75	0.75	0.80
	Harpenden	0.83	0.83	0.74	0.70	0.79	0.78
	Lange	0.82	0.85	0.69	0.72	0.74	0.80
BodPod	Lafayette Instruments	0.86	0.87	0.81	0.84	0.78	0.80
	Harpenden	0.87	0.90	0.85	0.84	0.82	0.86
	Lange	0.86	0.88	0.80	0.82	0.77	0.82

Note: Pearson's r values in subsample of subjects who performed additional testing where Exp = experienced and Inexp = inexperienced practitioners.

With the relationship between UWW and BodPod in mind, logistic regression analysis was performed to determine whether the addition of opposing technician skill level improved the predictive quality of the regression equation. This regression analysis was performed for the strongest predictor equation (seven-site) considering experienced and inexperienced practitioners as factors. Adding the opposing technician improved the model prediction in all cases except the Lafayette Instruments II caliper model that added inexperienced body fat estimates to the model already containing experienced body fat estimates (Table 6). However, in the models showing significant changes, the improvement in variance was too small to be considered practically important ($\Delta R^2 = 0.1 – 2.7$). This suggests that 0.1% to 2.7% of the variance could be improved by adding a second observation of the same subject to the model.

Table 6.
Logistic Regression Analysis Indicating Change in Variance by Adding a Second Observation to Each Subject. (N = 20).

		p value	ΔR^2
Lafayette Instruments	Exp	ns	4.8
	Inexp	< .02	1
Harpenden	Exp	< .01	0.1
	Inexp	< .02	0.6
Lange	Exp	< .05	2.7
	Inexp	< .05	0.1

Note: Exp = experienced practitioner and Inexp = inexperienced practitioner.

Table 7. Body fat percentage determined by caliper or body densitometry methods.

Method (Caliper Type or Densitometric Measure)	7-Site % Body Fat (presented as % ± standard deviation)	3-Site % Body Fat (presented as % ± standard deviation)
Lafayette Instruments	19.83 ± 6.59	20.51 ± 6.46
Harpenden	19.51 ± 6.56	19.99 ± 6.54
Lange	21.83 ± 6.84	22.35 ± 6.73
Under Water Weighing	25.48 ± 9.47	25.48 ± 9.47
BodPod	25.23 ± 10.41	25.23 ± 10.41

Caliper means determined by seven-site and three-site formulas with skinfolds measured by experienced and inexperienced practitioners in subset of subjects (n = 20) who performed UWW and BodPod. (P = ns)

Table 8. Body fat percentage determined by caliper or body densitometry methods.

Method (Caliper Type or Densitometric Measure)	7-Site % Body Fat (presented as % ± standard deviation)	3-Site % Body Fat (presented as % ± standard deviation)
Lafayette Instruments	19.83 ± 6.59	20.51 ± 6.46
Harpenden	19.51 ± 6.56	19.99 ± 6.54
Lange	21.83 ± 6.84	22.35 ± 6.73
Under Water Weighing	25.48 ± 9.47	25.48 ± 9.47
BodPod	25.23 ± 10.41	25.23 ± 10.41

Caliper means determined by seven-site and three-site formulas with skinfolds measured by experienced and inexperienced practitioners in subset of subjects (n = 20) who performed UWW and BodPod.

Discussion

The Lafayette Instruments II caliper performed well when compared to the Harpenden and Lange calipers. In previous research, Cyrino et al. (2003) performed a validation study on the Cescorf caliper using the Lange as the standard. These authors reported 5.2% to 6.9% differences in means for the calipers when using four different equations. Mean differences in the present study ranged from 0.4% to 2.3% between three calipers using a seven-site formula. Interestingly, we compared 20 subjects to UWW while Cyrino et al. compared 250 subjects.

There were strong correlations (0.82 and 0.87) when comparing the Lafayette Instruments caliper to UWW and BP. These correlations are weaker than the .90 reported by Jackson & Pollock (1978) when comparing body fat determination using Lange calipers with underwater weighing in male subjects. Our results were obtained using both men and women. Jackson et al. (1980) reported correlations near .85 in female subjects which strongly agree with our data. Gruber et al. (1990) predicted Lange performance with Harpenden calipers to within 90% of variance explained.

Our results agree with the comparison of BodPod underwater weighing performed by Fields et al. (2002). These authors reported 1.0% difference in body fat estimates between the two methods. We found an overall difference in means of 0.3% between BodPod and hydrostatic weighing. Fields et al. (2002) further described extremely large variation within subjects which was also seen in our study with standard deviations in BodPod of 10.4%. Skinfold caliper underestimation of body fat compared to UWW could be due to the subjects' inability to exhale fully during submersion during hydrostatic weighing. However, with similar estimates for UWW and BodPod, this may suggest that in heterogeneous populations, utilization of general body density formulae may limit predictive ability.

Ease of use was assessed by comparing determinations of body fat by inexperienced and experienced practitioners. No significant differences were found in an omnibus ANOVA. However, the Fisher's Z transformation and Z-test revealed that some of the calipers were, in fact, performing differently ($p < .05$). These differences are tempered by the limitations of the Z-transformation for highly correlated correlation coefficients (Meng et al., 1992) which were seen in our results. Further analysis was performed to investigate the differences between the statistical methods. Logistic regression was performed and also found significant. However, the variance explained by the model by adding another observation per subject, via addition of opposite level of expertise, was less than 3%. This suggests that any differences in caliper performance are practically unimportant.

Conclusions:

The Lafayette Instruments Company developed a new and relatively inexpensive caliper. This caliper performed well compared to Lange and Harpenden. Prediction of body fat compared to underwater weighing and BodPod may be somewhat underestimated by all three calipers.

References

Cyrino, E. S., Okano, A. H., Glaner, M. F., Romanzini, M., Gobbo, L. A., Makoski, A., et al. (2003). Impact of the use of different skinfold calipers for the analysis of body composition. [English Version] *Revista Brasileira de Medicina do Esporte, 9(3),* 150-153.

Fields, D. A., Goran, M. I., & McCrory, M. A. (2002). Body-composition assessment via air-displacement plethysmography in adults and children: a review. *American Journal of Clinical Nutrition, 75,* 453-467.

Gruber, J. J., Pollock, M. L., Graves, J. E., Colvin, A. B. & Braith R. W. (1990). Comparison of Harpenden and Lange calipers in predicting body composition. *Research Quarterly for Exercise and Sport 61(2),* 184-190.

Jackson, A. S. & Pollock, M. L. (1978). Generalized equations for predicting body density of men. *British Journal of Nutrition, 40(3)*, 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),* 175-181.

Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111(1),* 172-175.

Paffenbarger, R. S., Jr., Blair, S. N., Lee, I. M., & Hyde R. T. (1993). Measurement of Physical activity to assess health effects in free-living populations. *Medicine and Science in Sports and Exercise, 25(1), 60-70..*

Siri, W. E. (1956). The gross composition of the body. *Advances in biological and Medical physics,* 4, 239-280.

Lafayette Instrument
3700 Sagamore Parkway North
P.O. Box 5729 · Lafayette, IN 47903 U.S.A.
Tel: (765) 423-1505 · Fax: (765) 423-4111
E-mail: info@lafayetteinstrument.com
www.lafayetteinstrument.com