BRI0007111 **E-Mail:** **ana.paulapires@uol.com.br**

Pedido: 101019-571

Usuário:

Med Sci Sports Exerc
1980 12(3) pags. 175-81 / Jackson AS; Pollock ML; Ward A / Generalized equations for predicting body density of women. [(identificado) MEDLINE_1966-1996 pmid: 7402053]

Fonte de referência: (identificado) MEDLINE_1966-1996 pmid: 7402053

Ana Paula Pires dos Santos
Av Roberto Silveira 187/201 Centro
25685-040 - Petrópolis - RJ
BRASIL

BRI0007111 **E-Mail:** **ana.paulapires@uol.com.br**

Pedido: 101019-571

Med Sci Sports Exerc
1980 12(3) pags. 175-81 / Jackson AS; Pollock ML; Ward A / Generalized equations for predicting body density of women. [(identificado) MEDLINE_1966-1996 pmid: 7402053]

Local: BR1.1
Opções: BR1.1 / BR14.1

Atendido / Páginas:

Rejeitado / Motivo:

85 / 123

3093

MEDICINE AND SCIENCE IN SPORTS AND EXERCISE
Vol. 12, No. 3, pp. 175-182, 1980

Generalized equations for predicting body density of women

ANDREW S. JACKSON, MICHAEL L. POLLOCK, and ANN WARD

University of Houston, 77004; and
University of Wisconsin
Milwaukee, WI 53233

ABSTRACT

JACKSON, ANDREW S., MICHAEL L. POLLOCK and ANN WARD. Generalized equations for predicting body density of women. *Med. Sci. Sports Exercise.* Vol. 12, No. 3, pp. 175-182, 1980. Previous research with women has shown that body composition regression equations derived from anthropometric variables were population specific. This study sought to derive generalized equations for women differing in age and body composition. The hydrostatic method was used to determine body density (BD) and percent fat (%F) on 249 women 18 to 55 years ($\bar{X}$ = 31.4 ± 10.8 yrs) and 4 to 44 %F ($\bar{X}$ = 24.1 ± 7.2 %F). Skinfold fat (S), gluteal circumference (C) and age were independent variables. The quadratic form of the sum of three, four and seven S in combination with age and gluteal C produced multiple correlations that ranged from 0.842 to 0.867 with standard errors of 3.6 to 3.8 %F. The equations were cross-validated on a different sample of 82 women with similar age and %F characteristics. The correlations between predicted and hydrostatically determined %F ranged from 0.815 to 0.820 with standard errors of 3.7 to 4.0 %F. This study showed that valid generalized body composition equations could be derived for women varying in age and body composition, but care need to be exercised with women over an age of forty.

BODY COMPOSITION, BODY DENSITY, PERCENT FAT, ANTHROPOMETRIC VARIABLES, REGRESSION ANALYSIS

Anthropometry is a common field method for measuring body density (3). Multiple regression analysis has been used extensively to derive equations with functions of estimating laboratory determined body density (BD), lean body weight (LBW) or total body volume from combinations of skinfold fat (S), body circumference (C) and bone diameter (D) measurements. It has been shown that the distributions of body composition and anthropometric measures vary with age and sex (5,9-12,23,26,27,30,33, 39,41) and have resulted in the publication of regression equations for men (4,10,16,20,28,29,31,35,38,40,42) and women (10,19,20,30,32,36,38,41,43,45,46).

It has been shown with samples of men (10,29,31) and women (10,12,30) that body composition equations tend to be population specific. When these equations are applied to samples who differ in age and body composition, the prediction errors are inflated. This can be attributed to both the lack of homogeneity of regression slopes and intercepts (10,12,30,31). It appears that the population specificity of regression equations may be largely due to three factors. First, linear regression models have been used to derive prediction equations when research has shown that relationship between skinfold fat and BD is curvilinear (1,6,7,10,16). Second, age has been shown to account for BD variation beyond that accounted for by skinfold fat (10,16). And finally, sample sizes of most studies have tended to be quite small, normally less than 75 subjects. Cooley and Lohnes (8:56-57) have warned that regression equations derived on sample sizes of under 200 subjects need to be viewed with caution. The smaller the sample size in relation to the number of variables used in the regression analysis, the more likely the true multivariate relationship will be overestimated (8,14,29).

It was shown that generalized BD equations could be developed for men varying greatly in age and body composition (16). Using quadratic and logarithmically transformed sum of skinfolds in combination with age and selected C measurements, multiple correlations above 0.90 (SE = 0.0078 g/ml) were found. When these equations were cross-validated on a second sample, the correlation between predicted and hydrostatically determined BD exceeded 0.90. The purpose of this study was to derive generalized regression equations that would provide unbiased BD estimates for women varying in age and body composition. The research concentrated on the curvilinearity of the relationship and the function of age on BD.

METHODS

A total of 331 adult women between 18 and 55 years of age volunteered as subjects. The sample represented a wide range of women who varied considerably in body structure, body composition, and exercise habits. The subjects were tested in one of three laboratories over a period of six years. Except for slight differences in the measurement of residual volume, methods of data collection did not vary over the six year period. The total sample was randomly divided into a validation sample consisting of 249 women and a cross-validation sample of 82 subjects. The validation sample was used to derive generalized regression equations which were cross-validated with the second sample. This procedure has been recommended by Lord and Novick (24). The physical characteristics of the two samples are presented in Table 1.

Submitted for publication February, 1979.
Accepted for publication December, 1979.

TABLE 1. Physical characteristics of the validation and cross-validation samples and correlation with percent fat.

Variable	Validation Sample (N=249) $\bar{X}\pm SD$	Range	r_{xy}	Cross-Validation Sample (N=82) $\bar{X}\pm SD$	Range	r_{xz}
Percent Fat (%)	24.09 ± 7.20	4-44	1.00	24.83 ± 6.43	5-37	1.00
Body Density (g/ml)	1.044 ± 0.016	1.002-1.091	-1.00	1.043 ± 0.014	1.016-1.088	-1.00
Age (year)	31.44 ± 10.80	18-55	0.35	29.94 ± 11.20	18-53	0.05
Height (cm)	165.02 ± 6.00	146-181	0.08	162.67 ± 6.1	148-178	0.18
Weight (kg)	57.15 ± 7.59	36-87	0.64	57.95 ± 6.97	45-77	0.55
Lean Weight (kg)	43.04 ± 4.35	30-56	-0.15	43.31 ± 4.34	36-54	-0.16
Fat Weight (kg)	14.11 ± 5.75	2-35	0.95	14.63 ± 4.82	2-26	0.94
Gluteal C (cm)	92.98 ± 6.12	79-117	0.74	93.44 ± 5.52	82-106	0.69
Skinfold (mm)						
Chest	12.66 ± 4.78	4-26	0.66	12.83 ± 5.03	3-24	0.64
Axilla	12.59 ± 5.75	4-33	0.76	13.76 ± 6.33	3-31	0.69
Triceps	17.63 ± 5.71	6-41	0.77	19.14 ± 5.67	5-33	0.80
Subscapula	13.76 ± 6.05	4-39	0.69	14.99 ± 6.64	5-41	0.63
Abdomen	23.39 ± 9.43	4-56	0.77	23.93 ± 9.01	4-51	0.67
Suprailium	13.49 ± 6.89	3-40	0.78	14.50 ± 6.75	3-33	0.70
Thigh	29.18 ± 7.89	7-53	0.73	29.87 ± 7.56	11-46	0.73
Sum of Seven	122.70 ± 40.65	35-266	0.85	129.02 ± 40.75	39-231	0.80
Log Seven	4.76 ± 0.34	3.5-5.6	0.84	4.80 ± 0.36	3.7-5.4	0.80
Sum of Four[a]	83.69 ± 26.84	20-182	0.85	87.44 ± 25.66	26-157	0.81
Log Four[a]	4.38 ± 0.33	3.0-5.2	0.84	4.42 ± 0.34	3.3-5.1	0.81
Sum of Three[b]	60.30 ± 18.50	16-126	0.84	63.51 ± 18.02	19-106	0.81
Log Three[b]	4.05 ± 0.31	2.8-4.8	0.83	4.11 ± 0.32	2.9-4.7	0.82

[a]Sum of triceps, suprailium, abdomen and thigh skinfolds.
[b]Sum of triceps, suprailium and thigh skinfolds.

The subjects reported to the laboratory at least six hours post prandial for a minimum of three hours prior to testing and refrained from smoking. To avoid the problem of fluid retention which occurs during menstruation, all subjects were evaluated at least 7 days prior to or after their cycle.

Upon arrival at the laboratory, the subjects changed into a two-piece, light weight bathing suit. Standing height was measured to the nearest 0.6 cm on a standard physician scale, and body weight recorded to the nearest 10 g. Skinfold fat was measured at the chest, axilla, triceps, subscapula, abdomen, suprailium and thigh with a Lange skinfold fat caliper. The caliper had a constant pressure of 10 g/mm and measures were taken on the right side. Recommendations published by the Committee on Nutritional Anthropometry of the Food and Nutrition Board of the National Research Council were followed in obtaining skinfold fat data (22). Gluteal C was measured to the nearest 0.1 cm with a Lufkin steel tape by the procedures recommended by Behnke and Wilmore (3). Although other C and diameter measures were collected on these women (30) because of their lower correlations with the dependent variable, they were not used in the final data analysis. Cup size was also measured on all women. For analysis, cup sizes A,B,C, and D were transformed into the numerical values of 1,2,3, and 4. One of the investigators (MLP) collected all anthropometric data and supervised the hydrostatic weighing of all subjects.

The hydrostatic method was used to determine body density. Underwater weighing was conducted in a water tank in which a chair was suspended from a Chatillon 15 kg scale. The hydrostatic weighing procedure was repeated six to ten times until three similar readings to the nearest 20 g were obtained. Water temperature was recorded after each trial. Residual volume (RV) was determined by either the nitrogen washout or helium dilution technique. Although RV and hydrostatic weighing determinations were administered separately, the same postural positions, (sitting) were used for both. The procedure for determining body density followed the method outlined by Goldman and Buskirk (13). Body density was calculated from the formula of Brožek et al. (5) and fat percentage according to the formula of Siri (32). The Siri formula is based on the assumption of a fat-free body density of 1.10 gm/cc throughout all age groups.

In a factor analysis study (15) it was shown that all S variables measured the same body fat factor; therefore, the S were summed to provide a more stable estimate of S fat. The sum of all seven S, the sum of four S (triceps, suprailium, abdomen and thigh) and sum of three S (triceps, suprailium and thigh) were used. The sum of three and four S were used to provide a more feasible field test. The sum of S measurements was also logarithmically transformed so these results could be compared to the work of Durnin and Wormersley (10). Gluteal C was shown to measure a body fat factor independent of the factor measured by S measurements (15), and for this reason was used as an independent variable.

Multiple regression analysis (21) was used to derive the generalized equations. The dependent variable was hydrostatically determined BD. The independent variables were the quadratic form of the sum of S in combination with age and gluteal C. A second set of equations was de-

TABLE 2. Means and standard deviations for body density, percent body fat and sum of seven skinfolds of women classified in age-groups.

Variable	17-19	20-29	AGE GROUPS 30-39	40-49	50-59
			Validation Sample (N=249)		
Body Density	1.043±0.013 (N=25)	1.049±0.014 (N=102)	1.047±0.015 (N=57)	1.037±0.015 (N=48)	1.027±0.016 (N=17)
Percent Fat	24.8±6.0 (N=25)	21.8±6.4 (N=102)	22.7±6.8 (N=57)	27.3±6.7 (N=48)	32.1±7.5 (N=17)
Sum of Seven	128.4±30.4 (N=25)	110.4±34.1 (N=102)	114.7±39.1 (N=57)	137.1±41.7 (N=48)	172.5±43.5 (N=17)
			Cross Validation Sample (N=82)		
Body Density	1.038±0.009 (N=13)	1.045±0.016 (N=36)	1.046±0.019 (N=11)	1.040±0.012 (N=19)	1.041±0.003 (N=3)
Percent Fat	27.0±4.2 (N=13)	23.7±7.1 (N=36)	23.4±8.3 (N=11)	26.2±5.3 (N=19)	25.5±1.2 (N=3)
Sum of Seven	142.2±34.2 (N=3)	122.8±40.9 (N=36)	114.1±44.9 (N=11)	141.1±42.6 (N=19)	123.7±13.1 (N=3)

TABLE 3. Standardized regression coefficients and associated *t*-ratios for test of significance of regression analysis.

Variable	Sum of Seven SRC	*t*	Sum of Four SRC	*t*	Sum of Three SRC	*t*
Raw Score Skinfolds						
Sum of S	-1.03	- 6.44[a]	-1.04	- 6.55[a]	-1.02	- 6.13[a]
Sum of S^2	0.39	2.44[b]	0.39	2.44[b]	0.39	2.33[b]
Age	-0.07	- 2.01[b]	-0.04	- 1.18	-0.08	- 2.25[b]
Gluteal C	-0.24	- 5.08[a]	-0.23	- 4.57[a]	-0.22	- 4.22[a]
Log Transformed Skinfolds						
Sum of Log S	-0.64	-13.89[a]	-0.64	-13.20[a]	-0.62	-12.40[a]
Age	-0.08	- 2.30[b]	-0.05	- 1.59	-0.09	- 2.48[b]
Gluteal C	-0.25	- 5.48[a]	-0.25	- 5.30[a]	-0.25	- 4.93[a]

[a]$P<0.01$
[b]$P<0.05$

rived replacing the quadratic form of S with log transformed sum of S. The multiple regression coefficients were tested to determine if each independent variable was related to the dependent variable with the effects of all other variables statistically controlled (21:66-70). The cross-validation procedures recommended by Lord and Novick (24) were followed to determine if the equations derived on the validation sample accurately predicted %F of the cross-validation sample.

RESULTS

Presented in Table 1 are descriptive statistics for the validation and cross-validation samples. The means, and standard deviations show that both samples were heterogeneous in age and body composition. The product-moment correlation between each variable and hydrostatically determined BD is provided. With the exception of age for the cross-validation sample, all anthropometric variables selected for regression analysis were significantly related to hydrostatically determined BD. Of interest, the sums of S were more highly correlated than any of the individual S measurements. The higher correlations were expected because it has been shown that S variables measure the same source of common variance (15). The means and standard deviations for skinfold fat, BD and %F classified by age groups are presented in Table 2. These means for the validation sample exhibited a non-linear trend by age-groups (P<0.01). Durnin and Wormersley (1974) have published similar data with women of the same age groups. The average %F of the Durnin and Wormersley sample was higher for all age groups, with the largest differences in the four older age groups. An inspection of the %F means for the Durnin-Wormersley data showed a near perfect linear trend with age which is in variance with the trend found with these data of the validation sample.

Presented in Table 3 are the standardized regression coefficients and the associated t-ratios which test if the coefficient is significantly different from zero. BD was the dependent variable for this analysis. The coefficients are orthogonal to each other and the magnitude of the regression coefficient reflects the independent variable's relative contribution with the effects of all other variables statistically removed. With the exception of age in combination with the sum of four S and gluteal C, all coefficients were significantly different from zero. The zero order correlation between age and the sum of four skinfolds was slightly higher ($r = 0.37$) than the correlation found between age and the sum of seven S ($r = 0.33$) and sum of three S ($r = 0.32$). It appeared that this higher relationship produced the non-significant t-ratio for the analysis of the sum of four S. An examination of the standardized regression coefficients shows that age was the least important independent variable studied.

Presented in Tables 4 are raw score equations with functions to predict BD. The standard errors are presented as both BD and %F. The intercorrelations among the sum of three, four and seven S were high ($r<0.966$); therefore, the multiple correlations for the BD equations ranged from 0.838 to 0.867 with standard errors that varied from 3.6 to 4.0%F. The multiple correlations for the log-transformed equations tended to be slightly lower than the quadratic equations. The regression analysis in Table 3 showed that gluteal C had a higher standardized regression coefficient than age. For this reason, equations were formed with all possible S, age and gluteal C combinations.

TABLE 4. Generalized regression equations for predicting body density of adult women.

Variables	Regression Equation*	R	Standard Error BD	Standard Error %F
	Sum of Seven Skinfolds			
S,S², Age	$BD(1)=1.0970-0.00046971\ (X_1)+0.00000056\ (X_1)^2 -0.00012828\ (X_4)$	0.852	0.0083	3.8
Log S, Age	$BD(2)=1.23173-0.03841\ (\log X_1)-0.00015\ (X_4)$	0.850	0.0084	3.8
S,S², C	$BD(3)=1.1470-0.00042359\ (X_1)+0.00000061\ (X_1)^2 -0.00065200\ (X_5)$	0.865	0.0080	3.6
log S, C	$BD(4)=1.25475-0.03100\ (\log X_1)-0.00068\ (X_5)$	0.864	0.0080	3.6
S,S², Age, C	$BD(5)=1.1470-0.00042930\ (X_1)+0.00000065\ (X_1)^2 -0.00009975\ (X_4)-0.00062415\ (X_5)$	0.867	0.0079	3.6
log S, Age, C	$BD(6)=1.25186-0.03048\ (\log X_1)-0.00011\ (X_4) -0.00064\ (X_5)$	0.867	0.0079	3.6
	Sum of Four Skinfolds			
S,S², Age	$BD(7)=1.0960950-0.0006952\ (X_2)+0.0000011\ (X_2)^2 -0.0000714\ (X_4)$	0.849	0.0084	3.8
Log S, Age	$BD(8)=1.21993-0.03936\ (\log X_2)-0.00011\ (X_4)$	0.845	0.0085	3.9
S,S², C	$BD(9)=1.1443913-0.0006523\ (X_2)+0.0000014\ (X_2)^2 -0.0006053\ (X_5)$	0.861	0.0081	3.7
log S, C	$BD(10)=1.24374-0.03162\ (\log X_2)-0.00066\ (X_5)$	0.859	0.0081	3.7
S,S², Age, C	$BD(11)=1.1454464-0.0006558\ (X_2)+0.0000015\ (X_2)^2 -0.0000604\ (X_4)-0.0005981\ (X_5)$	0.862	0.0081	3.7
log S, age, C	$BD(12)=1.241721-0.031069\ (\log X_2)-0.000077\ (X_4) -0.000635\ (X_5)$	0.861	0.0081	3.7
	Sum of Three Skinfolds			
S,S², Age	$BD(13)=1.0994921-0.0009929\ (X_3)+0.0000023\ (X_3)^2 -0.0001392\ (X_4)$	0.842	0.0086	3.9
Log S, Age	$BD(14)=1.21389-0.04057\ (\log X_3)-0.00016\ (X_4)$	0.838	0.0087	4.0
S,S², C	$BD(15)=1.1466399-0.0009300\ (X_3)+0.0000028\ (X_3)^2 -0.0006171\ (X_5)$	0.851	0.0084	3.8
Log S, C	$BD(16)=1.23824-0.03248\ (\log X_3)-0.00067\ (X_5)$	0.849	0.0084	3.8
S,S², Age, C	$BD(17)=1.1470292-0.0009376\ (X_3)+0.0000030\ (X_3)^2 -0.0001156\ (X_4)-0.0005839\ (X_5)$	0.854	0.0083	3.8
Log S, Age, C	$BD(18)=1.23530-0.03192\ (\log X_3)-0.00013\ (X_4) -0.00062\ (X_5)$	0.853	0.0083	3.8

*Key: X_1=Sum of all seven skinfolds, mm; X_2=Sum of triceps, abdomen, suprailium and thigh skinfolds, mm; X_3=sum of triceps, thigh and suprailium skinfolds, mm; X_4=age, yrs; X_5=gluteal circumference, cm.

TABLE 5. Cross-validation of generalized equations on cross-validation sample (N=82).

Variables	Equation Number	X ± SD	$r_{yy'}$[a]	SE[b]	Range of SE Age[c]	Range of SE %Fat[d]
Sum of Seven Skinfolds						
S,S², Age	BD(1)	25.0 ± 6.1	.803	3.9	3.4 - 4.0	3.4 - 4.3
Log S, Age	BD(2)	24.8 ± 6.3	.799	4.0	3.4 - 4.5	3.3 - 4.9
S, S², C	BD(3)	25.0 ± 6.2	.817	3.8	3.6 - 4.4	3.5 - 4.0
Log S, C	BD(4)	24.9 ± 6.4	.814	3.9	3.5 - 4.6	3.4 - 4.5
S, S², Age, C	BD(5)	24.9 ± 6.1	.815	3.8	3.5 - 4.5	3.5 - 4.0
Log S, Age, C	BD(6)	24.8 ± 6.3	.811	3.9	3.5 - 4.6	3.5 - 4.4
Sum of Four Skinfolds						
S, S², Age	BD(7)	25.0 ± 5.9	.813	3.8	3.3 - 4.6	3.7 - 4.3
Log S, Age	BD(8)	24.8 ± 6.1	.801	3.9	3.3 - 4.9	3.3 - 4.7
S, S², C	BD(9)	24.9 ± 6.1	.821	3.7	3.4 - 4.3	3.3 - 4.1
Log S, C	BD(10)	25.0 ± 6.0	.814	3.8	3.4 - 4.5	3.3 - 4.6
S, S², Age, C	BD(11)	24.9 ± 6.0	.819	3.7	3.4 - 4.4	3.4 - 4.1
Log S, Age, C	BD(12)	24.8 ± 6.1	.811	3.7	3.4 - 4.5	3.5 - 4.5
Sum of Three Skinfolds						
S, S², Age	BD(13)	25.1 ± 5.8	.820	3.7	3.3 - 4.6	3.5 - 4.1
Log S, Age	BD(14)	25.0 ± 6.0	.814	3.8	3.2 - 4.8	3.5 - 4.6
S, S², C	BD(15)	25.1 ± 6.0	.827	3.7	3.4 - 4.4	3.4 - 3.9
Log S, C	BD(16)	25.0 ± 6.1	.824	3.7	3.9 - 4.4	3.4 - 4.1
S, S², Age, C	BD(17)	25.0 ± 5.9	.825	3.7	3.4 - 4.4	3.5 - 3.8
Log S, Age, C	BD(18)	24.9 ± 6.0	.820	3.8	3.4 - 4.5	3.6 - 4.1

[a]zero order correlation between laboratory determined %F and %F predicted from the regression equation in Table 4

[b]$SE = \sqrt{\Sigma(Y' - Y)^2/N}$

[c]Age Categories: 29.9, 30-39.9 40.

[d]%F Categories: 19.9, 20.0-27.9 28.0

The BD equations were applied to the data of the cross-validation sample and this analysis is presented in Table 5. The product-moment correlations between laboratory determined and estimated BD were slightly lower than the multiple coefficients and ranged from 0.799 to 0.827 with standard errors expressed as %F that ranged from 3.7 to 4.0%. The means and standard deviations for %F estimated by the BD equations were similar to the laboratory determined %F values (i.e., $\bar{X} = 24.8 \pm 6.4$). The largest mean difference between predicted and laboratory determined %F was only 0.3 %F, and most standard deviations for the distributions of estimated %F exceeded 6.0 %F.

The cross-validation sample was then reduced first, to three age categories and next, to levels of body fat content by three percent fat categories. The range of standard errors (%F) for these different categories are also presented in Table 5. Since these standard errors were based on smaller sample sizes that ranged from 11 to 49 cases, more random variation was expected. With six categories and 18 equations, 108 standard errors were calculated and only six standard errors exceeded 4.4 percent fat. The analysis of categories did provide some interesting trends; the highest standard errors were found for the age category of 40 years and above, and the percent category of below 20 percent body fat. The variance in standard errors suggests that the equations might be more accurate when applied to women who are under 40 years of age and above 20 percent body fat. In all instances, the log transformed equations produced higher standard errors. Because of the smaller sizes used for this analysis, however, these results need to be viewed with caution.

A second method used to cross-validate the generalized equation was to apply the equation to data reported in the literature. Equation BD(13) was applied to the average values reported by other investigators, and the predicted BD and %F means were compared to the reported laboratory means. These means are provided in Table 6. Katch and Michael (19) and Durnin-Wormersley (10) did not measure thigh skinfold; therefore, a second quadratic equation consisting of the sum of triceps, subscapula, and suprailium skinfolds in combination with age was derived with the validation data. This second sum of three equation exhibited concurrent validity similar to that of equation BD(13) (R = 0.832; SE = 0.0088 g/ml and 3.9%F). The Katch-Michael and Durnin-Wormersley studies did not report mean age, but did provide an age range. The mid-point of the age range was used, and this could introduce an error of about 0.5 to 1%F. The generalized equations provided accurate mean predictions. Of eleven values, seven means differed by 1.5%F or less and only one mean difference exceeded 3%F. All predicted means were within one standard error of prediction (i.e., 3.9%F) of the reported laboratory determined mean.

TABLE 6. Body density and percent fat means predicted from generalized equations and laboratory determined means reported in the literature.

Investigator	Predicted BD	Predicted %F[a]	Lab Determined BD	Lab Determined %F[a]
Katch and Michael (1968)	1.0500	21.4	1.0490	21.9
Katch and McArdle (1973)	1.0426	24.8	1.0394	26.2
Sinning (1978)	1.0612	16.4	1.0635	15.4
Sloan et al. (1962)	1.0442	24.0	1.0467	22.9
Wilmore and Behnke (1969)	1.0439	24.2	1.0406	25.7
Young et al. (1961)	1.0325	29.4	1.0342	26.6
Durnin and Wormersley (1974)[b]				
(16-19 yrs)	1.0545	23.5	1.040	26.0
(20-29 yrs)	1.0358	27.9	1.034	28.7
(30-39 yrs)	1.0254	32.7	1.025	32.9
(40-49 yrs)	1.0267	32.1	1.020	35.3
(50-68 yrs)	1.0188	35.9	1.013	38.6

[a] $\%F = [(4.95/BD) - 4.50]\ 100$.
[b] Durnin and Wormersley split their total sample into five age groups. The sample size ranged from 29 to 100 women.

DISCUSSION

The multiple correlations found for the generalized equations are high and compare favorably with the results reported in the literature. The high multiple correlations are due partially to the heterogenous subjects studied; however, the standard errors are well within the values reported by other investigators (10,19,20,30,32,36,38,41, 43,45,46) who used more homogeneous samples of women. Multiple correlations based on data from an original sample tend to overestimate the true relationship. While the cross-validation standard errors of this study (Table 5) are slightly higher than the values derived for the regression equations (Table 3), the cross-validation standard errors are well within the range reported in the literature. One of the more extensive studies with women was published by Durnin and Wormersley (9). They reported BD standard errors for log transformed S equations which ranged from 0.0082 to 0.0125 g/ml which represents a range of 3.72 to 5.65 percent body fat and demonstrates less accuracy than the cross-validation statistics presented in Table 5.

The most common statistical method used to select the variables to be included in body composition regression equations has been the forward step-wise regression model (12,15,17,20,29-31,32,40-46). The step-wise method will yield the highest multiple correlation for the data set studied, but tends to overestimate the true multivariate relationship (8:57,14,21:282). Rather than using a stepwise method, the strategy used in this study was to select the variables that research has shown were related to body composition. For the subjects studied, the testing procedure included securing seven S, thirteen C, and seven body diameter measures, age, height, weight, and breast cup size. The variables and methods are described in another source (30). To determine if any of these additional variables accounted for more variation, an equation consisting of the sum of seven S in combination with the thirteen C, seven D and cup size was derived. The multiple correlation for this 22-variable equation was 0.884 which is slightly higher than the equation with the highest multiple correlation presented in Table 3. The multiple correlation 0.884 was found not to be significantly higher than the multiple of 0.868 found for equation BD(5) Table 4 ($F_{(18,226)} = 1.59$; $P > 0.05$) which showed that the additional variance accounted for by the 22 variable equation was random.

The non-significant difference between the multiple correlations of equation BD(5), Table 4 and the 22-variable equation is supported by the results of a factor analytical study (15). A factor analysis of 25 anthropometric variables with hydrostatically determined lean and fat weight on women isolated four common factors which accounted for 80 percent of the total variance measured by the 27 variables. The best independent variables to use with regression equations are those unrelated with each other, but correlated with the dependent variable. Since anthropometric variables were shown to measure only four different sources of common variance, prediction accuracy is not increased by using a large set of anthropometric variables. The substantial contribution of gluteal C was expected because it was shown to measure a fat factor independent of the more general fat factor isolated with S measures.

The findings of several studies (10,12,17,29-31) showed that regression equations were population specific. The application of regression equations derived on one sample but applied to a second sample produces bias body composition estimates. Flint and associates (12) showed that published equations varied in their accuracy to predict laboratory determined %F with subsamples of women who varied in age, physical fitness level and body composition. Prediction errors have been shown to be due to the lack of homogeneity and regression slopes (29) and intercepts (17,31). The findings of this study showed that some of this prediction bias may be attributed to the use of linear regression models because the relationship between S fat and hydrostatistically determined %F is curvilinear (1, 6,7,10,16). The significant contribution of age beyond that accounted for by S fat is a second reason for population specific equations. The results of this study are in agreement with a similar study conducted with a sample of adult men (16).

It has been shown that regression equations derived from women differing greatly in age do not accurately predict body composition for subjects of different age groups (9,10,12). Durnin and Wormersley (10) adjusted for age differences by developing log transformed S equations with the same slopes by different intercepts to adjust for age differences. In the development of generalized equations for men (16), age was used as an independent variable because it was found to account for a significant proportion of BD variance beyond that accounted for by the quadratic or log form of S fat. The results of this study are in agreement with these results, but an examination of the standardized regression weights showed that age was less important for women than men. With men, the standardized regression weights ranged from −0.11 to −0.15 and all were significantly different from zero ($P < 0.01$). Whereas, with these women, the standardized regression weights were less than −0.10 and not significantly different from zero with the equation of the sum of four skinfolds. The average age of the men was one year older than the women. It is suspected that this difference in age distribution may account for some of the differences in regression coefficients.

The cross-validation analysis by age categories provided evidence that the equations may be less accurate for women above 40 years of age. The BD and %F means provided in Table 2 were found to be non-linear and it was suspected that the higher standard errors found for the older women may have been due to the use of linear regression weights in the generalized equations. Polynominal regression analysis showed that the relationship between age and %F was quadratic ($F_{(1,246)} = 19.51$; $p<0.01$); however, when the quadratic form of age was tested with the quadratic form of the sum of seven S, the quadratic component for age was found to be non-significant ($F_{(1,244)} = 0.85$; $p>0.05$). The regression analysis showed that the quadratic relationship found for age was due to the common variance between age and the sum of seven S. Durnin and Wormersley (10) published equations for women of different age groups. They reported that the standard errors for women in age categories of 40 to 49 years and 50 to 68 years were lower than the younger age group of 20 to 29 years and 30 to 39 years. The Durnin and Wormersley findings do not agree with the cross-validation findings for the older subjects of our sample. Of interest, the largest mean differences reported in Table 6 were found in the Durnin and Wormersley subjects who were over 40 years of age. It is suspected that this difference may not be due to age, but body composition. The subjects studied by Durnin and Wormersley were classified as sedentary; whereas, many of the subjects of this sample were engaged in exercise programs and were on the average, about 6 to 8%F less than the Durnin and Wormersley subjects. These findings suggest the need for further research with older subjects of different body composition.

The results of this study and a similar study with men (16) support the concept of generalized body composition regression equations for adults varying in age and body composition. During recent years, the research strategy has been to develop specific equations with application to unique populations (4,19,20,28-31,32,35,36,40-46). Specific equations will provide valid estimates with subjects representative of the specific population, but the more specific the population, the less general application an equation will have. The approach taken in this study was to use a sample representative of a heterogeneous population and develop a regression model to account for the important sources of variability of the population. The use of the heterogeneous population was the approach used by Kannel and associates (18) to develop a generalized equation for the prediction of cardiovascular disease. The generalized equation replaced several equations valid for specific age groups (37). The development of generalized equations is dependent upon having a representative sample with a sufficient number of cases. Cooley and Lohnes (8:56) maintain that regression equations based on samples of less than 200 subjects need to be viewed with caution. A major limitation of published body composition regression equations is the sample sizes tend to be small, typically with less than 75 subjects. The development of specific equations encourages the use of small samples because one has less subjects from which to select.

The accuracy of the generalized equations of this study when applied to subjects representative of unique populations can only be documented by cross-validation research. The intent of this study was to develop equations applicable to a wide range of adult females. The results of this study support the conclusion that the equations are valid within a defined degree of accuracy when applied to samples of adult women varying in age and body composition. Care needs to be taken with women over age 40. Body composition equations may be less accurate with older women and Wormersley and Durnin (1974) have suggested that the constant of 1.100 g/ml for fat-free weight may not apply to women of this age group. The strongest evidence supporting the validity of the generalized equations was provided by the cross-validation standard errors which were low and nearly identical to the standard errors found with the validation sample. The relatively small differences in the standard errors of the equations demonstrate that the equations may all be used with similar confidence. Equation BD(13), the sum of three skinfolds with age, is recommended because it would be the most feasible for mass testing and its cross-validation statistics are similar to the values found with the other 17 equations.

Data for this study were collected at: Wake Forest University, Winston—Salem, North Carolina; Institute for Aerobics Research, Dallas, Texas; and Mount Sinai Medical Center, Milwaukee, Wisconsin.

Present address for A.S. Jackson: Department of Health and Physical Education, University of Houston, Houston, Texas.

Present address for M.L. Pollock and A. Ward: Department of Medicine, University of Wisconsin Medical School—Mount Sinai Medical Center, Milwaukee, Wisconsin.

REFERENCES

1. Allen, T.H., M.T. Peng, K.P. Chen, T.F. Huang, C. Chang, and H.S. Fang. Prediction of total adiposity from skinfolds and the curvilinear relationship between external and internal adiposity. *Metabolism* 5:346-352, 1956.
2. Armore, S.J. *Elementary statistics and decision making.* Columbus: Charles E. Merrill Publishing Co., 1973, pp. 311-317.
3. Behnke, A.R. and J.H. Wilmore. *Evaluation and regulation of body build and composition.* Englewood Cliffs: Prentice Hall, Inc., 1974.
4. Brožek, J. and A. Keys. The evaluation of leanness-fatness in man: norms and intercorrelations. *Br. J. Nutr.* 5:194-206, 1951.
5. Brožek. J., F. Grande, J.T. Anderson, and A. Keys. Densitometric analysis of body composition; revision of some quantitive assumptions. *Ann. N.Y. Acad. Sci.* 110:113-140, 1963.
6. Chen, K.P. Report on measurement of total body fat in American women estimated on the basis of specific gravity as an evaluation of individual fatness and leanness. *J. Formosan Med. Assoc.* 52:271-276, 1953.
7. Chen, S., M. Peng, K. Chen, T. Huang, C. Chang, and H. Fang. Longitudinal studies on adipose tissue and its distribution in human subjects. *J Appl. Physiol.* 39:825-830, 1975.
8. Cooley, W.W. and P.R. Lohnes. *Multivariate data analysis.* New York, John Wiley and Sons Inc., 1971.
9. Durnin, J.V.G.A. and M.M. Rahaman. The assessment of the amount of fat in the human body from measurements of skinfold thickness. *Br. J. Nutr.* 21:681-689, 1967.
10. Durnin, J.V.G.A. and J. Wormersley. Body fat assessed from total body density and its estimation from skinfold thickness; measurements on 481 men and women aged from 16 to 72 years. *Br. J. Nutr.* 32:77-92, 1974.
11. Edwards, D.A.W. Differences in the distributions of subcutaneous fat with sex and maturity. *Clinical Sci.* 10:305-315, 1951.
12. Flint, M.M., B.L. Drinkwater, C.L. Wells, and S.M. Horvath. Validity of estimating body fat of females: effect of age and fitness. *Hum. Biol.* 49:559-572, 1977.
13. Goldman, R.F. and F.B. Buskirk. Body volume measurement by underwater weighing description of a method. In *Techniques for Measuring Body Composition.* Ed. by J. Brožek and A. Henschel. pp. 78-89, Washington, D.C., Natl. Acad. Sci., 1961.
14. Jackson, A.S. Biometric problems of estimating body composition. Paper presented at the symposium "Assessment of Body Composition: Application to Athletes." 25th Annual Meeting of the American College of Sports Medicine, Washington, D.C., 1978.
15. Jackson, A.S. and M.L. Pollock. Factor analysis and multivariate scaling of anthropometric variables for the assessment of body compositon. *Med. Sci. Sports* 8:196-203, 1976.
16. Jackson, A.S. and M.L. Pollock. Generalized equations for predicting body density of men. *Br. J. Nutr.* 40:497-504, 1978.
17. Jackson, A.S. and M.L. Pollock. Prediction Accuracy of body density, lean body weight, and total body volume equations. *Med. Sci. Sports* 9:197-201, 1977.
18. Kannel, W.B., D. McGee, and T. Gordon. A general cardiovascular risk profile: The Framingham study. *Am. J. Cardiol.* 38:46-50, 1976.
19. Katch, F.J. and E.D. Michael. In Prediction of body density from skinfold and girth measurements of college females. *J. Appl. Physiol.* 25:92-94, 1968.
20. Katch, F.I. and W.D. McArdle. Prediction of body density from simple anthropometric measurements in college age women and men. *Hum. Biol.* 45:445-454, 1973.
21. Kerlinger, F.M. and E.S. Pedhazur. *Multiple regression in behavioral research.* New York, Holt, Rinehart and Winston, Inc., 1973.
22. Keys, A. (Chairman). Recommendations concerning body measurements for the characterization of nutritional status. *Human. Biol.* 28:111-123, 1956.
23. Lesser, G.T., I. Kumar, and J.M. Steele. Changes in body composition with age. *Ann. N.Y. Acad. Sci.* 110:579-588, 1963.
24. Lord, F.M. and M.R. Novick, *Statistical theories of mental test scores.* pp. 285-288, Reading, Mass., Addison-Wesley, 1968.
25. MacMillan, M.G., C.M. Reid, D. Sterling, and P. Passmore. Body composition resting oxygen consumption, and urinary creatinine in Edinburgh students. *Lancet* 1:728-729, 1965.
26. Malina, R.M. Quantification of fat, muscle and bone in man. *Clin. Orthopaed.* 65:9-38, 1969.
27. Novak, I.P. Aging, total body potassium, fat-free mass and cell mass in males and females between ages 18 and 85 years. *J. Gerontol.* 27:438-443, 1972.