Br. J. Nutr. (1978), 40, 497

Generalized equations for predicting body density of men

By A. S. JACKSON* and M. L. POLLOCK†

Wake Forest University, Winston-Salem, North Carolina and Institute of Aerobics Research, Dallas, Texas, USA

(*Received* 3 *August* 1977 – *Accepted* 28 *February* 1978)

1. Skinfold thickness, body circumferences and body density were measured in samples of 308 and ninety-five adult men ranging in age from 18 to 61 years.

2. Using the sample of 308 men, multiple regression equations were calculated to estimate body density using either the quadratic or log form of the sum of skinfolds, in combination with age, waist and forearm circumference.

3. The multiple correlations for the equations exceeded 0·90 with standard errors of approximately ±0·0073 g/ml.

4. The regression equations were cross validated on the second sample of ninety-five men. The correlations between predicted and laboratory-determined body density exceeded 0·90 with standard errors of approximately 0·0077 g/ml.

5. The regression equations were shown to be valid for adult men varying in age and fatness.

Anthropometry is a common field method for measuring body density (Behnke & Wilmore, 1974). Brožek & Keys (1951) were the first to publish regression equations with functions of predicting body density with anthropometric variables. Subsequently, numerous investigators have published equations using various combinations of skinfolds and body circumferences.

The development of generalized equations for predicting body density from anthropometric equations has been found to have certain limitations. First, equations have been shown to be population specific and different equations were needed for samples of men varying in age and body fatness. It was shown that with samples of men differing in age, the slopes of the regression lines were homogeneous, but the intercepts were significantly different (Durnin & Womersley, 1974; Pollock, Hickman, Kendrick, Jackson, Linnerud & Dawson, 1976). It was further shown that the slopes of the regression lines of young adult men and extremely lean world class distance runners were not parallel (Pollock, Jackson, Ayres, Ward, Linnerud & Gettman, 1976). The differences of either slopes or intercepts resulted in bias body density estimates. A related problem has been that linear regression models have been used to derive prediction equations, when research has shown that a curvilinear relationship exists between skinfold fat and body density (Allen, Peng, Chen, Huang, Chang & Fang, 1956; Chen, Peng, Chen, Huang, Chang & Fang, 1975; Durnin & Womersley, 1974). This non-linear relationship may be the reason for the differences in slopes and intercepts.

Durnin & Womersley (1974) logarithmically transformed the sum of skinfolds to create a linear relationship with body density, but still needed different intercepts to account for age differences. The purpose of this investigation was to derive generalized regression equations that would provide unbiased body density estimates for men varying in age and body composition. Efforts were concentrated on the curvilinearity of the relationship and the function of age on body density.

* Present address: Department of Health and Physical Education, University of Houston, Houston, Texas, USA.

† Present address: Cardiovascular Disease Section, Mount Sinai Medical Center, University of Wisconsin, School of Medicine, Milwaukee, Wisconsin, USA.

Table 1. *Physical characteristics of the validations and cross-validation samples**

Variable	Validation sample (*n* 308) Mean	SD	Range	Cross-validation sample (*n* 95) Mean	SD	Range
Age (year)	32·6	10·8	18–61	33·3	11·5	18–59
Height (m)	1·792	0·065	1·63–2·01	1·784	0·059	1·66–1·91
Weight (kg)	74·8	11·8	54–123	77·6	11·7	53–102
Body density (g/ml)	1·0586	0·0181	1·0161–1·0996	1·0564	0·0188	1·0259–1·0998
Fat (%)†	17·7	8·0	1–33	18·7	8·3	1–33
Lean weight (kg)	63·9	7·4	48–100	62·4	6·7	47–81
Fat weight (kg)	14·5	7·9	1–42	15·2	7·9	1–31
Sum 7 skinfolds (mm)	122·6	52·0	32–272	124·7	53·1	31–222
Log 7 skinfolds (mm)	4·70	0·49	3·47–5·61	4·71	0·53	3·43–5·40
Sum 3 skinfolds (mm)‡	59·4	24·3	14–118	59·2	25·4	10–111
Log 3 skinfolds (mm)	3·98	0·49	2·64–4·78	3·95	0·56	2·30–4·71
Waist circumference (m)	0·871	0·097	0·67–1·25	0·874	0·1	0·68–1·14
Forearm circumference (m)	0·288	0·019	0·22–0·37	0·287	0·021	0·24–0·39

* For explanation see p. 499.
† Fat (%) = [(4·95/BD)+4·5] 100 (Siri, 1961) Fat (%).
‡ Sum of chest, abdomen and thigh skinfolds.

METHODS

A total of 403 adult men between 18 and 61 years of age volunteered as subjects. The sample represented a wide range of men who varied considerably in body structure, body composition, and exercise habits. The subjects were tested in one of two laboratories (Wake Forest University, Winston-Salem, North Carolina and Institute for Aerobics Research, Dallas, Texas) over a period of 4 years. The total sample was randomly divided into a validation sample consisting of 308 men and a cross-validation sample of ninety-five subjects. The validation sample was used to derive generalized regression equations and were cross-validated with the second sample. This procedure has been recommended by Lord & Novick (1968). The physical characteristics of the two samples are presented in Table 1.

Upon arrival at the laboratory, the subjects were measured for standing height to the nearest 0·01 m (0·25 in) and for body-weight to the nearest 10 g. Skinfold fat was measured at the chest, axilla, triceps, subscapula, abdomen, supra-iliac, and thigh with a Lange skinfold fat caliper, manufactured by Cambridge Scientific Industries, Cambridge, Maryland, USA.

Recommendations published by the Committee on Nutritional Anthropometry of the Food and Nutrition Board of the National Research Council were followed in obtaining values for skinfold fat (Keys, 1956). A previous study (Pollock, Hickman *et al.* 1976) showed that waist and forearm circumference accounted for body density variance beyond skinfold fat, and for this reason, were included in this study. Waist and forearm circumferences were measured to the nearest 1 mm with a Lufkin steel tape, manufactured by the Lufkin Rule Company, Apex, North Carolina, USA. The procedures and location of the anthropometric sites measured were shown and described by Behnke & Wilmore (1974).

The hydrostatic method was used to determine body density. Underwater weighing was conducted in a fibreglass 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 (Katch, 1968). Water temperature was recorded after each trial. Residual volume was determined by either the nitrogen washout or helium dilution technique. The procedure for determining body density followed the method out-

Table 2. *Regression analysis for predicting body density using the sum of seven skinfolds in adult men aged* 18–61 *years*†

Source of variance	Degrees of freedom	Sum of squares	Mean square	F, ratio for statistical significance	Standard regression certificate for full model
			Sum of seven skinfolds		
Full model	5	0·08418	0·01684	336·80*	—
Skinfold fat	(2)	0·07878	0·03939	787·20*	—
Linear	1	(0·07757)	0·07757	1551·40*	−1·18
Quadratic	1	(0·00121)	0·00121	24·20*	0·53
Age	(1)	0·00279	0·00279	55·80*	−0·14
Circumferences	(2)	0·00261	0·00261	52·20*	—
Waist	—	—	—	—	−0·32
Forearm	—	—	—	—	0·20
Residual	302	0·01612	0·00005	—	—
			Log transformation of seven skinfolds		
Full model	4	0·08425	0·02106	421·20*	—
Log skinfold fat	(1)	0·07706	0·07706	1541·20*	−0·64
Age	(1)	0·00284	0·00284	56·80*	−0·13
Circumferences	(2)	0·00435	0·00435	87·00*	—
Waist	—	—	—	—	−0·38
Forearm	—	—	—	—	0·23
Residual	303	0·01605	0·00005	—	—

* $P < 0·01$.
† For details, see Table 1.

lined by Goldman & Buskirk (1961). Body density was calculated from the formula of Brožek, Grande, Anderson & Keys (1963) and fat percentage according to Siri (1961) (see Table 1).

In a factor analysis study, it was shown (Jackson & Pollock, 1976) that skinfolds measured the same factor; therefore, the skinfolds were summed. The sum of several measurements provides a more stable estimate of subcutaneous fat. A second sum consisting of chest, abdomen and thigh skinfolds was also derived. These three skinfolds were selected because of their high intercorrelation with the sum of seven and it was thought that they would provide a more feasible field test. The sum of skinfolds were also logarithmically transformed so that they could be compared with the work of Durnin & Womersley (1974).

Regression analysis (Kerlinger & Pedhazur, 1973) was used to derive the generalized equations. Polynomial models were used to test if the relationship between body-density and the sum of skinfolds was curvilinear. 'Step-down' analysis was used to determine if age, and then age in combination with the circumference measurements, accounted for additional body-density variance beyond that attributed to the sum of skinfolds. The cross-validation procedures recommended by Lord & Novick (1968) were followed to determine if the equations derived on the validation sample accurately predicted the body density of the cross-validation sample.

RESULTS

Table 1 shows that basic results derived from the validation and cross-validation samples including natural log transformations of the sum of skinfolds. The standard deviations and ranges showed that the men differed considerably in both age and body composition. Tables 2 and 3 show the regression analysis using the sum of seven and sum of three skin-

Table 3. *Regression analysis for predicting body density using the sum of three skinfolds*†

Source of variance	Degrees of freedom	Sum of squares	Mean square	F, ratio for statistical significance	Standard regression certificate for full model
			Sum of three skinfolds		
Full model	5	0·08453	0·01691	338·20*	—
Skinfold fat	(2)	0·07998	0·04000	800·00*	—
Linear	1	0·07943	0·07943	1588·60*	−1·11
Quadratic	1	0·00055	0·00055	11·00*	0·43
Age	(1)	0·00220	0·00220	44·00*	−0·12
Circumferences	(2)	0·00234	0·00117	23·40*	—
Waist	—	—	—	—	−0·31
Forearm	—	—	—	—	0·19
Residual	302	0·01577	0·00005	—	—
		Log transformation of three skinfolds			
Full model	4	0·08415	0·02104	420·80*	—
Log skinfold fat	(1)	0·07674	0·07674	1534·80*	−0·62
Age	(1)	0·00248	0·00248	49·60*	−0·11
Circumferences	(2)	0·00493	0·00493	98·60*	—
Waist	—	—	—	—	−0·41
Forearm	—	—	—	—	0·23
Residual	303	0·01626	0·00005	—	—

* $P < 0{\cdot}01$
† For details, see Table 1.

folds respectively. The correlation between the sum of three and seven skinfolds was 0·98; thus, the regression analyses for these variables were nearly identical. The full model consisted of either the linear and quadratic or the log transformed sum of skinfolds in combination with age, and body circumferences. The multiple correlations for these full models were nearly identical, ranging from 0·915 to 0·918. Regression equations for the full models may be found in Table 4.

Since the full models were significant, the step-down analysis was conducted to determine if each variable accounted for a significant proportion of body-density variance. The first analysis within the full model was to determine if the relationship between skinfold fat and body density was linear or quadratic. This was found to be quadratic which supported the findings of other investigators (Allen *et al.* 1956; Chen *et al.* 1975; Durnin & Womersley, 1974). Durnin & Womersley (1974) used a log transformation to form a linear relationship between skinfold fat and body density. For this reason, only the linear relationship with log transformed skinfolds was used.

Age was the next variable entered into the regression model and it accounted for a significant proportion of body-density variance beyond the log-transformed or quadratic form of skinfolds. Waist and forearm circumference were the last two variables entered into the full model and these measures accounted for a significant proportion of body-density variance beyond age and skinfold fat.

The standardized regression coefficients for the full model are presented in Tables 2 and 3. The magnitude of these weights represented the relative importance of each variable with the effects of the other variables held constant. These statistics showed that the linear and quadratic components accounted for most of the body density variance. The negative weighting of the sum of skinfolds and positive weighting of the squared sum of skinfolds represent the quadratic relationship between body density and the sum of skinfolds. The

Table 4. *Generalized regression equations for predicting body density (BD) of adult men ages* 18–61 *years**

Anthropometric variables	Regression equation	Equation no.	R	SE
	Sum of seven skinfolds			
S,S^2, age	$BD = 1.11200000 - 0.00043499\,(X_1) + 0.00000055\,(X_1)^2 - 0.00028826\,(X_3)$	1	0.902	0.0078
S,S^2, age, C	$BD = 1.10100000 - 0.00041150\,(X_1) + 0.00000069\,(X_1)^2 - 0.00022631\,(X_3) - 0.0059239\,(X_4) + 0.0190632\,(X_5)$	2	0.916	0.0073
log S, age	$BD = 1.21394 - 0.03101\,(\log X_1) - 0.00029\,(X_3)$	3	0.893	0.0082
log S, age, C	$BD = 1.17615 - 0.02394\,(\log X_1) - 0.00022\,(X_3) - 0.0070\,(X_4) + 0.02120\,(X_5)$	4	0.917	0.0073
	Sum of three skinfolds			
S,S^2, age (5)	$BD = 1.1093800 - 0.0008267\,(X_2) + 0.0000016\,(X_2)^2 - 0.0002574\,(X_3)$	5	0.905	0.0077
S,S^2, age, C (6)	$BD = 1.0990750 - 0.0008209\,(X_2) + 0.0000026\,(X_2)^2 - 0.0002017\,(X_3) - 0.005675\,(X_4) + 0.018586\,(X_5)$	6	0.918	0.0072
log S, age (7)	$BD = 1.18860 - 0.03049\,(\log X_2) - 0.00027\,(X_3)$	7	0.888	0.0083
log S, age, C (8)	$BD = 1.15737 - 0.02288\,(\log X_2) - 0.00019\,(X_3) - 0.0075\,(X_4) + 0.0223\,(X_5)$	8	0.915	0.0073

S, Sum of skinfolds; C, circumference; X_1, sum of chest, axilla, triceps, subscapula, abdomen, suprailium and front thigh skinfolds; X_2, sum of chest, abdomen and thigh skinfolds; X_3, age; X_4, waist circumference; X_5, forearm circumference.

* For details, see Table 1.

Table 5. *Cross-validation of generalized equations on the calibration sample* (n 95)

Variables	Equation no.*	$r_{yy'}$	SE†	Range of SE: Age‡	Range of SE: Fat§
			Sum of seven skinfolds		
S,S^2, age	1	0.915	0.0078	0.0064–0.0085	0.0066–0.0092
S, S^2, age, C	2	0.915	0.0077	0.0057–0.0094	0.0067–0.0084
log S, age	3	0.914	0.0078	0.0055–0.0085	0.0054–0.0091
log S, age, C	4	0.913	0.0078	0.0061–0.0098	0.0064–0.0091
			Sum of three skinfolds		
S, S^2, age	5	0.917	0.0077	0.0066–0.0083	0.0057–0.0087
S, S^2, age, C	6	0.920	0.0076	0.0066–0.0092	0.0058–0.0087
Log S, age	7	0.904	0.0085	0.0064–0.0112	0.0047–0.0102
log S, age, C	8	0.910	0.0082	0.0057–0.0100	0.0060–0.0097

S, sum of skinfolds; C, circumference; $r_{yy'}$, correlation between predicted (y') and laboratory determined (y) body density.

* For details, see Table 4.

† $SE = \sqrt{[\Sigma (y'-y)^2/n]}$.

‡ Age (years) categories; $<$19.9, 20.0–29.9, 39.0–39.9, 40.0–49.9, $>$ 50.0.

§ Fat (%) categories: $<$9.9, 10.0–14.9, 15.0–19.9, 20.0–24.9, $>$ 25.0.

positive weighting for waist and negative weighting for forearm is consistent with the results reported by Katch & McArdle (1973).

Table 4 lists selected raw score equations and the equation's multiple correlation and standard error. The high multiple correlations are due partially to the heterogeneous sample studied. However, the standard errors are low and well within the values reported by other

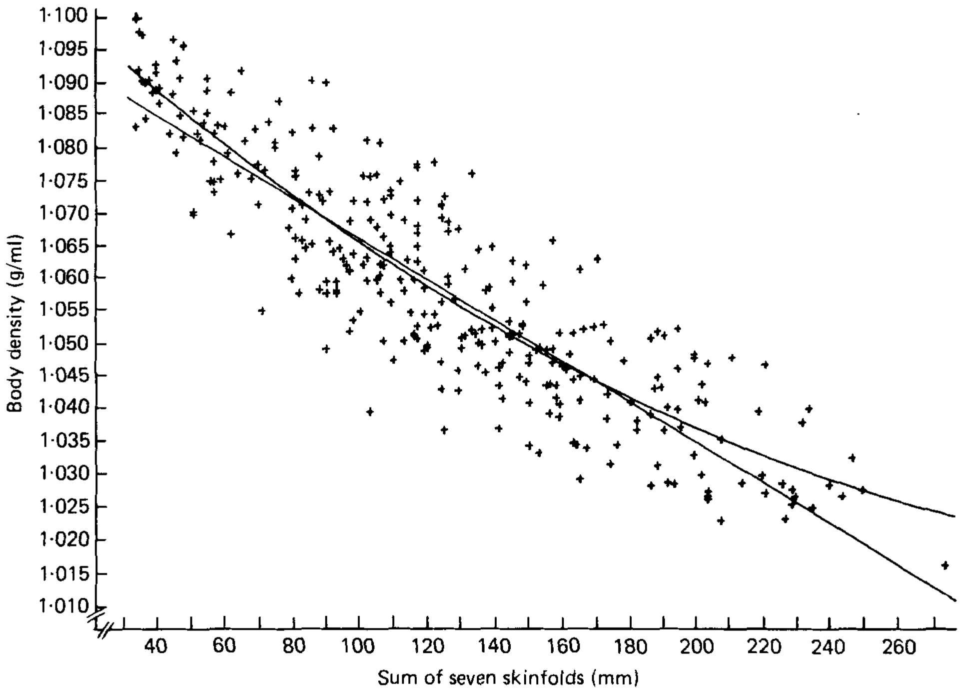

Fig. 1. Scattergram of body density and sum of seven skinfolds, with the linear quadratic regression lines, for adult men aged 18–61 years (for details, see Table 1). Details of generalized regression equations are given in Table 4.

investigators (Katch & McArdle, 1973; Pascale, Grossman, Sloane & Frankel, 1956; Pollock, Hickman *et al.* 1976; Sloan, 1967; Wilmore & Behnke, 1969; Wright & Wilmore, 1974) who used more homogeneous samples.

The 'raw score' equations were applied to the anthropometric results of the cross-validation sample. The cross-validation analysis is presented in Table 5. The product moment correlation between laboratory determined and estimated body density were all higher than 0·90, and the standard errors were within the range found with the validation sample results.

The cross-validation sample was then reduced first, to five age categories, and next, to levels of body fat content by five fat (%) categories. The ranges of standard errors for these different categories are also presented in Table 5. With the exception of the log equations, none of the standard errors exceeded 0·0100 g/ml. Since these standard error estimates were based on sample sizes that varied from ten to thirty-three cases, more variability was expected. These analyses showed that the regression equations accurately predicted body density for samples differing in age and fatness.

DISCUSSION

The findings of several studies (Durnin & Womersley, 1974; Pollock, Hickman *et al.* 1976) showed that regression equations were population specific. The application of regression equations derived on one sample, but applied to other samples that differed in age and fatness, produced biassed body density estimates. The findings of this study showed that some of this bias may be attributed to the use of linear regression models because the relationship between skinfold fat and body density was quadratic. This is shown by the 'scattergram' between the sum of seven skinfolds and body density which is presented as Fig. 1. Both linear and quadratic regression lines are provided. The differences between the two regression lines showed where the largest bias prediction errors would occur. This was at the ends of the bivariate distribution. For example, the fat (%) differences between the linear and quadratic sum of seven skinfold equations for 250 and 40 mm of skinfold fat were 2·9 and 1·3 fat (%) respectively, while the difference was only 0·5 fat (%) for 150 mm.

In a previous study (Pollock, Jackson *et al.* 1976), it was found that the slopes of the regression lines of lean world-class distance runners and young adult men were not parallel. The prediction of the body-density of the lean runner with linear equations derived on a sample of young adult men systematically underestimated the body density of these lean subjects. This source of systematic error is documented by the differences between the linear and quadratic regression lines shown in Fig. 1 and confirms the need for quadratic equations.

It has been shown that the intercepts of the regression lines of young adult men and older (+35 years) and fatter men were different (Pollock, Hickman *et al.* 1976). Since the relationship between body-density and skinfold fat was quadratic, the differences in intercepts could be partly due to the use of linear regression equations. The results reported by Durnin & Womersley (1974) showed, however, that age was also responsible for the intercept differences. Durnin & Womersley (1974) used a logarithmic transformation of the sum of four skinfolds. This transformation changed the quadratic relationship between body density and the sum of skinfolds, in the 'raw score' form, into a linear relationship. With male subjects who ranged from 16 to 59 years of age, they reported that the slopes for samples divided by 10 year intervals were parallel, but had different intercepts. This would result in biassed estimates due to age differences, thus Durnin & Womersley (1974) provided five different equations which had the same slope, but different intercepts.

The finding of this study, that age accounted for a significant proportion of body-density variation beyond that attributed to quadratic or logarithmic sum of skinfolds agreed with the findings reported by Durnin & Womersley (1974). They suggested that this age-relationship may be due to a higher proportion of total body fat being situated internally and a decrease in the density of fat-free mass. The decrease in fat-free mass was primarily attributed to skeletal changes (Durnin & Womersley, 1974). In the present study, the use of age as an independent variable accounted for intercept difference, and eliminated the need for several different age-adjusted equations. The cross-validation results documented the accuracy of a generalized equation for samples differing in age and fatness. The standard errors found in these analyses are within the range reported by Durnin & Womersley (1974). Using 209 men who varied in age from 16 to 72, Durnin & Womersley (1974) reported standard errors that ranged from 0·0059 to 0·0117 g/ml for prediction equations derived for similar age groups.

The multiple correlations for the generalized equations derived with the logarithmic or quadratic sum of skinfolds were nearly identical. The results of the cross-validation analysis suggested that the quadratic equations were more accurate. The standard errors tended to be lower for the total sample and less variable for the total sample and for the different age and fat (%) categories. This was expecially true for the sum of three skinfolds.

The generalized equations provided valid and accurate body-density estimates with adult men varying in age and fatness. The cross-validation of equations is important because one is not certain that equations developed with one sample will predict body density with the same accuracy when applied to the data of a different sample. The best evidence is provided by the standard error when the equation is cross-validated on the second sample. The standard errors for the cross-validation analysis were low and nearly identical to the standard errors found with the validation sample. This provided the strongest evidence that the generalized equations were accurate and valid for use with adult men varying in age and body density.

REFERENCES

Allen, T. H., Peng, M. T., Chen, K. P., Huang, T. F., Chang, C. & Fang, H. S. (1956). *Metabolism* **5**, 346.
Behnke, A. R. & Wilmore, J. H. (1974). *Evaluation and Regulation of Body Build and Composition*. Englewood Cliffs: Prentice-Hall.
Brožek, J., Grande, F., Anderson, J. T. & Keys, A. (1963). *Ann. N.Y. Acad. Sci.* **110**, 113.
Brožek, J. & Keys, A. (1951). *Br. J. Nutr.* **5**, 194.
Chen, S., Peng, M. T., Chen, K. P., Huang, T. F., Chang, C. & Fang, H. S. (1975). *J. appl. Physiol.* **39**, 825.
Durnin, J. V. G. A. & Rahaman, M. M. (1967). *Br. J. Nutr.* **21**, 681.
Goldman, R. F. & Buskirk, E. R. (1961). In *Techniques for Measuring Body Composition*, p. 78 [J. Brožek and A. Henschels, editors]. Washington, DC: National Academy of Science.
Jackson, A. S. & Pollock, M. L. (1976). *Med. Sci. Sports* **8**, 196.
Katch, F. I. (1968). *Research Quarterly* **39**, 993.
Katch, F. I. & McArdle, W. D. (1973). *Human Biol.* **45**, 445.
Kerlinger, F. N. & Pedhazur, E. S. (1973). *Multiple Regression in Behavioral Research*. New York: Holt, Rinehart and Winston.
Keys, A. (1956) *Human Biol.* **28**, 111.
Lord, F. M. & Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*, pp. 285–288. Reading, Mass.: Addison-Wesley.
Pascale, L. R., Grossman, M. I., Sloane, H. S. & Frankel, T. (1956). *Human Biol.* **28**, 165.
Pollock, M. L., Hickman, T., Kendrick, Z., Jackson, A. S., Linnerud, A. C. & Dawson, G. (1976). *J. appl. Physiol.* **40**, 300.
Pollock, M. L., Jackson, A. S., Ayres, J., Ward, A., Linnerud, A. & Gettman, L. (1976). *Ann. N.Y. Acad. Sci*, **301**, 361.
Siri, W. E. (1961). In *Techniques for Measuring Body Composition*, p. 223 [J. Brožek and A. Hanschels, editors]. Washington DC: National Academy of Science.
Sloan, A. W. (1967). *J. appl. Physiol.* **23**, 311.
Wilmore, J. H. & Behnke, A. R. (1969). *J. appl. Physiol.* **27**, 25.
Wright, H. F. & Wilmore, J. H. (1974). *Aerospace Med.* **45**, 301.

Printed in Great Britain