Journal of Preventive Medicine and Public Health
November 2010, Vol. 43, No. 6, 479-485

doi: 10.3961/jpmph.2010.43.6.479

Reliability and Data Integration of Duplicated Test Results Using Two Bioelectrical Impedence Analysis Machines in the Korean Genome and Epidemiology Study

Boyoung Park[1,6], Jae Jeong Yang[1,6], Ji Hyun Yang[1,6], Jimin Kim[1,2], Lisa Y. Cho[1,6], Daehee Kang[1,6,7], Chol Shin[3], Young Seoub Hong[4], Bo Youl Choi[5], Sung Soo Kim[2], Man Suck Park[2], Sue K. Park[1,6,7]

[1]Department of Preventive Medicine, Seoul National University College of Medicine; [2]Center for Genome Science, National Health Institute, Korea Centers for Disease Control and Prevention; [3]Department of Internal Medicine, Korea University College of Medicine; [4]Department of Preventive Medicine, Dong-A University College of Medicine; [5]Department of Preventive Medicine, Hanyang University College of Medicine; [6]Seoul National University Cancer Research Institute; [7]Department of Biomedical Science, Seoul National University College of Medicine

Objectives: The Korean Genome and Epidemiology Study (KoGES), a multicenter-based multi-cohort study, has collected information on body composition using two different bioelectrical impedence analysis (BIA) machines. The aim of the study was to evaluate the possibility of whether the test values measured from different BIA machines can be integrated through statistical adjustment algorithm under excellent inter-rater reliability.
Methods: We selected two centers to measure inter-rater reliability of the two BIA machines. We set up the two machines side by side and measured subjects' body compositions between October and December 2007. Duplicated test values of 848 subjects were collected. Pearson and intra-class correlation coefficients for inter-rater reliability were estimated using results from the two machines. To detect the feasibility for data integration, we constructed statistical compensation models using linear regression models with residual analysis and R-square values.
Results: All correlation coefficients indicated excellent reliability except mineral mass. However, models using only duplicated body composition values for data integration were not feasible due to relatively low R^2 values of 0.8 for mineral mass and target weight. To integrate body composition data, models adjusted for four empirical variables that were age, sex, weight and height were most ideal (all R^2>0.9).
Conclusions: The test values measured with the two BIA machines in the KoGES have excellent reliability for the nine body composition values. Based on reliability, values can be integrated through algorithmic statistical adjustment using regression equations that includes age, sex, weight, and height.

Key words: Electric impedance, Statistical model, Regression
J Prev Med Public Health 2010;43(6):479-485

INTRODUCTION

The Korean Genome Epidemiologic Study (KoGES), supported by the Korea Centers for Disease Control and Prevention (KCDC), was launched to examine the etiology of common diseases that focuses on gene-environment and gene-gene interaction in the Korean population [1,2]. The KoGES is a multi-cohort study that includes cohorts of health examinees in cities, rural communities, twins and their families, migrants, subjects in Ansan, and subjects in Ansung. All cohort studies collected information on subjects' body compositions using Bioelectrical Impedence Analysis (BIA) machines. Information was collected by two BIA machines that were Inbody 330 (Biospace, Seoul, Korea) for six cohorts and Zeus 9.9 (Jawon Medical, Kyoungsan, Korea) for five other three cohorts in 2007. Body composition values produced from the two machines were calculated from different estimation equations according to the number of impedence values and empirical values. Inbody330 estimates body composition values using four impedances from both arms and both legs, while Zeus 9.9 estimates use two impedances from both arms and convergence values of four empirical variables that are age, sex, height and weight.

Our major concern in using different machines to measure body composition variables was whether the machines produce consistent values, even though both machines adopt the BIA method to measure the nine body compositions that are percent body fat (PBF), fat

Corresponding author : Sue K. Park MD, PhD, 28 Yeongeon-Dong, Jongro-Gu, Seoul, Korea, Tel : +82-2-740-8338, Fax : +82-2-747-4830, E-mail : suepark@snu.ac.kr
Received : 7 June 2010, Accepted : 9 September 2010

Table 1. General characteristics of study subjects repetitively measured with the two BIA machines

	Measurement with machine A then B machine at center X (n=410)	Measurement with machine B then A machine at center Y (n=438)
Age(y), mean (SD)	53.7 (9.0)	53.5 (6.9)
< 50	141 (34.4)	140 (32.0)
50 - 59	140 (34.2)	224 (51.1)
60 - 69	119 (29.0)	54 (12.3)
> 70	10 (2.4)	20 (4.6)
Sex		
Male	105 (25.6)	220 (50.2)
Female	305 (74.4)	218 (49.8)
Height (cm), mean (SD)	159.0 (7.4)	161.5 (8.3)
Weight (kg), mean (SD)	61.4 (9.2)	64.7 (9.9)

BIA: bioelectrical impedence analysis

mass (FM), soft lean body mass (SLBM), total body water (TBW), free fat mass (FFM), protein mass (PM), mineral mass (MM), target weight (TW), basal metabolic rate (BMR). Since both machines had different estimation formulas and measurement values, their absolute values may not be the same. The results of the body composition measurement analyzed by the two machines will be merged into the KoGES data. Prior to data integration, the potential inconsistency produced from the two machines should be detected and appropriate adjustments to overcome biases need to be undertaken [3].

In this study, we examined whether body composition test values from different machines can be integrated by machine statistical adjustments and aimed to develop the statistical model for correcting potential biases.

SUBJECTS AND METHODS

We selected two cohort centers that were a cohort center that used Inbody 330 and another center that used Zeus 9.9. For two months from October 15 to December 14, 2007, the body compositions of the study participants in each center were measured twice by machine which were setup side by side. In center X, the subjects were measured first with machine A and then with machine B and in center Y, subjects measured in the reverse order with machine A and then B to avoid time bias. A trained nurse measured each subjects' body compositions on the machine which was primarily used and then the second machine which was newly established for this reliability study. Both machines were connected to a computer and printer where the data was automatically stored and printed out.

Inter-rater reliability is defined as the degree to which different raters give consistent estimates of the same value [4]. To determine inter-rater reliability, we calculated the mean (standard deviation, SD) of the absolute values of the difference in the duplicated test results at each center and computed the Pearson's correlation coefficients as an index of inter-rater reliability. Fisher's Z-transformation p-values were computed to examine the difference in correlation coefficients of repeated test values obtained in both centers [5]. The correlation coefficient was interpreted as follows: $\lambda < 0.25$ = little or no reliability; $0.25 \leq \lambda < 0.50$ = fair; $0.50 \leq \lambda < 0.75$ = moderate to good; $\lambda \geq 0.75$ = good to excellent [6]. Also, intra-class correlation coefficients were calculated to estimate inter-rater reliability regardless of the centers.

To identify whether the variables- age, sex, height and weight- known as the factors that affect the body composition [7] and their interactions affect differences in body composition values between two machines and choose which variables and their interactions are to be adjusted, we conducted multivariate linear regressions with interaction terms for age, sex, height and weight and each body composition value was measured on machine B for each outcome variable of each value measured in machine A. A p-value of factors and interaction less than 0.05 was considered significant. To construct simple and accurate models that could be applied to each body composition, we chose the common variables that were significant in all of the body composition. Because the models with significant interaction terms were increased the R^2 value to less than 0.1 and included interaction variables were different according to body compositions, models without interaction term were chosen. Finally, the three types of linear regression models for data integration were as follows: Equation 1: [value measured in B machine] = α + β1 [value measured in A machine]; Equation 2: [value measured in B machine] = α+ β1 [value measured in A machine] + β2 [age] + β3 [sex]; Equation 3: [value measured in B machine] = α+ β1[value measured in A machine] + β2 [age] + β3 [sex] + β4[height] + β5[weight]. To confirm model fitting, the sums of residuals and R^2 values for each regression model were computed. Changes in R squares from Equation 1 or 2 to Equation 3 equation were calculated as percent differences according to the following equation:[R square value in Equation 3] - [R square value in Equation 1 or 2])/[R square value in Equation 3] *100. The scatter plots between the real value from machine B and the calculated value from the final equation were drawn. All statistical analyses were performed using

Table 2. Inter-rater reliability of the two repetitive values measured with the two BIA machines

	Absolute value of mean (SD) in the difference between the two test values at center X (n=410)		Absolute value of mean (SD) in the difference between the two test values at center Y (n=438)		% difference in correlation coefficients	Fisher's Z-transformation	Intra-class correlation coefficient
	Mean (SD) in difference	Correlation coefficient [a]	Mean (SD) in difference	Correlation coefficient [b]	|[b] - [a]|/[b] x 100	p-value*	Coefficient (95% CI)
Percent body fat (%)	1.37 (1.21)	0.9522	1.95 (1.67)	0.9429	1.0	0.1856	0.93 (0.92-0.94)
Fat mass (kg)	0.94 (0.78)	0.9671	1.28 (1.12)	0.9539	1.4	0.0126	0.95 (0.94-0.96)
Soft lean body mass (kg)	1.49 (1.08)	0.9885	2.02 (1.32)	0.9835	0.5	0.0084	0.96 (0.95-0.97)
Total body water (%)	1.09 (0.66)	0.9880	1.00 (0.76)	0.9853	0.3	0.1385	0.98 (0.97-0.98)
Fat free mass (kg)	0.95 (0.84)	0.9880	1.28 (1.13)	0.9854	0.3	0.1523	0.98 (0.98-0.98)
Protein mass (kg)	2.37 (0.51)	0.9825	2.62 (0.61)	0.9504	3.4	<0.0001	0.37 (0.31-0.42)
Mineral mass (kg)	1.01 (0.28)	0.8886	1.08 (0.42)	0.6949	27.9	<0.0001	-0.18 (-0.25- 0.12)
Target weight (kg)	2.76 (2.62)	0.8962	3.01 (2.71)	0.9057	1.0	0.4641	0.85 (0.84-0.87)
Basal metabolic rate (kcal)	79.79 (52.19)	0.9387	87.69 (49.63)	0.9538	1.6	0.0351	0.79 (0.77-0.82)

SD: standard deviation, CI: confidence interval.
BIA: bioelectrical impedence analysis.
* Test for two correlation coefficients at centers X and Y using Fisher's Z-transformation.

Table 3. R-square values for fit of linear regression models* using the duplicated test values of the BIA machine A and B†

	Equation 1†	Equation 2§	Equation 3‖	% difference in R squares	
	R square [a] (F-value)	R square [b] (F-value)	R square [c] (F-value)	|[c] - [a]|/[c] x100	|[c] - [b]|/[c] x100
Percent body fat	0.8927 (7047.4)	0.8948 (2392.36)	0.9175 (1872.55)	2.7	2.5
Fat mass	0.9144 (9036.59)	0.9216 (3308.37)	0.9485 (3102.18)	3.6	2.8
Soft lean body mass	0.9723 (29744.0)	0.9730 (10154.6)	0.9792 (7941.09)	0.7	0.6
Total body water	0.9740 (31745.4)	0.9741 (10586.0)	0.9838 (10234.1)	1.0	1.0
Fat free mass	0.9740 (31730.2)	0.9741 (10584.1)	0.9838 (10237.6)	1.0	1.0
Protein mass	0.9302 (11279.2)	0.9380 (4255.94)	0.9388 (2582.01)	0.9	<0.1
Mineral mass	0.7995 (3373.56)	0.8577 (1695.27)	0.9930 (24028.9)	19.9	13.7
Target weight	0.8180 (3801.45)	0.8846 (2155.82)	0.9977 (73953.1)	18.0	11.3
Basal metabolic rate	0.9004 (7649.31)	0.9563 (6162.44)	0.9686 (5181.79)	7.0	1.3
Potential for compensating the bias of test values produced from the two BIA machines	Acceptable except for mineral mass and target weight	Acceptable except for mineral mass and target weight	Acceptable method for mineral mass and target weight		

Age, weight and height: continuous scale, Sex: binary scale of 1 (male) and 2 (female).
BIA: bioelectrical impedence analysis.
* Sum of residuals of all regression models were zero.
† A, B: Two values measured with the two BIA machines.
† Equation 1: [value measured in machine B] = $\alpha + \beta 1$ [value measured in machine A].
§ Equation 2: [value measured in machine B] = $\alpha + \beta 1$ [value measured in machine A] + $\beta 2$ [age] + $\beta 3$ [sex].
‖ Equation 3: [value measured in machine B] = $\alpha + \beta 1$ [value measured in machine A] + $\beta 2$ [age] + $\beta 3$ [sex] + $\beta 4$ [height] + $\beta 5$ [weight].

SAS version 9.2 (SAS Inc., Cary, NC, USA) and SPSS version 17 (SPSS Inc., Chicago, IL, USA)

RESULTS

A total of 848 subjects participated in this reliability study. 410 subjects at center X were measured using machine A then B, and 438 subjects at center Y were measured using machine B then A. 68.6% at center X and 83.1% at center Y was less than 60 years old, and 74.4% at center X and 49.8% at center Y were females (Table 1).

Table 2 shows inter-rater reliability of the two repetitive values measured in the two BIA machines. All correlation coefficients of the nine variables that were PBF, FM, SLBM, TBW, FFM, PM, MM, TW, and BMR measured by the two machines at center X showed excellent reliability (λ>0.75), however the MM at center Y showed fair reliability (λ=0.69). For Fisher's Z-transformation, the two correlation coefficients for MM and PM were significantly different (p<0.0001) and the differences between the two coefficients were 27.9% for MM and 3.4% for PM. Although the p-values of Fisher's test on FM, SLBM and BMR were significant (p<0.05), the difference between the two coefficients

was less than 1.6%. In the aspect of intra-class correlation coefficients (ICC), all intra-class correlation coefficients except MM and PM presented reliability over 0.79 (for MM, ICC was 0.37 and that for PM was -0.18).

The four empirical variables were significantly related to the nine body compositions (all p-values < 0.05) (Results not shown). Interaction of age and sex in each model was significantly related to the six body compositions except for TBW, MM and BMR. Age*weight interaction term was significant for seven body composition values except for TBW and TW and age*height was significant for only SLBM, PM, TW and BMR (all p-values<0.05) (Results not shown).

Table 3 shows the R^2 values in each regression models using two repeated body composition values. All residual plots of regression models which presented residuals between the actual and predicted values showed that all were randomly distributed around 0 (Figure not shown). For equation 1, most R^2 values were greater than 0.9, however MM, TW and PBF R^2 values showed lower values of 0.80, 0.82 and 0.89, respectively. For equation 2, R^2 values for MM increased (0.7995 in Equation 1; 0.8577 in Equation 2), while that on for TW and PBF changed slightly (0.8180 in Equation 1; 0.8846 in Equation 2 for TW, 0.8927 in Equation 1; 0.8948 in Equation 2 for PBF). For equation 3, R^2 values for the three body composition values were the greatest (0.9930, 0.9977 and 0.9175 respectively) and the other R^2 values also excellent R^2 values (greater than 0.9). The R^2 values for models with significant interaction terms (less than p-value 0.05) in the multivariate regression analysis such as age*sex, age*height or age*weight were increased but the degree was slight (less than 0.1) (Results not shown).

The scatter plots between the observed value from machine B (x-line) and the estimated value from the final equation 3 (y-line) and the equations were presented in Appendix 1. Between the observed and estimated values, most of the body composition components showed good correspondence, however PM and MM had each one distinct outlier, which did not show good correspondence.

DISCUSSION

This study evaluated the feasibility of integration of body composition values measured by two machines used in the KoGES by statistical compensation algorithm with good reliability. We found that most body composition values measured by the two machines had excellent inter-reliability ($\lambda > 0.75$ and ICC > 0.8) and linear regression models that included the four empirical variables that were age, sex, height and weight to compensate for the discrepancy of the test values of the two machines had excellent R^2 values greater than 0.9.

Except for PM and MM, most components of body composition showed excellent correlation. Most of the body compositions except MM and TW could be integrated in the equations using only BIA variable without adjustment of personal values. However, MM and TW could be integrated in the equations using BIA variable and the four control variables such as age, sex, height and weight, which showed excellent R-square value. Increasing variances after including height and weight might be caused by their effects and variations in TW and MM. However, considering low R square values in the equation using the BIA variable only and the degree of their improvement after including the four control variables in MM and TW, the BIA equations including age, sex, weight and height is necessary for integration of the all body compositions.

Differences in test values in the two machines were due to differences in method in calculating body composition. Though both machines use the BIA system, Inbody 330 calculates body compositions based on the impedance values produced from four limbs while Zeus 9.9 estimates are based on impedance values from two arms only and four empirical variables that are age, sex, height and weight. Therefore, this may result in significant differences in body composition value. Another reason may be due to the differences in quality control status at each center and inter-observer variation [8]. In this study, to avoid inter-observer variation, we had a trained nurse at each center manage all the processes for the survey period. However, problems related to quality control and intra-individual variations are unavoidable. The most duplicated values showed excellent inter-rater reliability ($\lambda > 0.90$) and the differences between the two correlation coefficients at centers X and Y were small (under 3.4%), while the duplicated results on MM showed fair reliability ($\lambda > 0.69$) at center Y and had excellent reliability ($\lambda = 0.89$) at center X. Most of the mean differences of body composition values except total body water was greater in center Y. It might be caused by the characteristics of each center but the differences were not significant.

Time change can lead to different results in the same individual [9]. In this study, to minimize variability according to time, two machines were set side by side in the same room and measurements were conducted

consecutively by the same nurse. At center X, subjects were measured with machine A and then B, while at center Y, subjects were measured in the reverse order. Our results demonstrate the correlation coefficients for most body composition values were similar, except for MM. Quality control issue at center Y and time variability were not thought to have resulted in the low coefficient for MM since the other coefficients regardless of center and body composition values were all high (nearly 0.9).

All random values in all the regression analyses regardless of inclusion of the four empirical variables were distributed around '0' which suggests that efficient adjustments were performed through statistical techniques. However, some regression models using only body composition values measured by machines A and B or addition of age and sex were not feasible due to relatively lower model fitting. The regression models that included age, sex, height and weight for statistical adjustment fitted very well for all body composition values. In the aspect of R square value, the models that included interaction terms showed a better fit but the degree was minimal (R^2 increased less than 0.1), thus concluded models without the interaction terms are preferable for availability.

Based on the KoGES data, this study aimed to detect the feasibility of whether body composition values produced from the two machines can be integrated through statistical compensation models using regression equations including age, sex, height and weight. This study showed that quality control and assurance issues in information collection are important in multicenter based research such as the KoGESand other multi-center cohorts. Thus, the KoGES needs to pay close attention to quality control and standardization when conducting survey processes.

ACKNOWLEDGEMENT

This study was supported by a research grant from the Korea Centers for Disease Control and Prevention (2007-E00004-00).

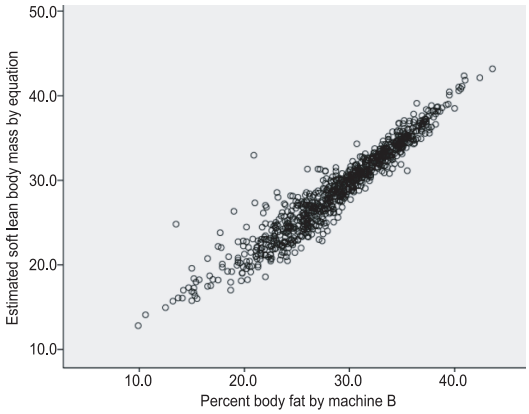
CONFLICT OF INTEREST

The authors have no conflicts of interest to declare on this study.

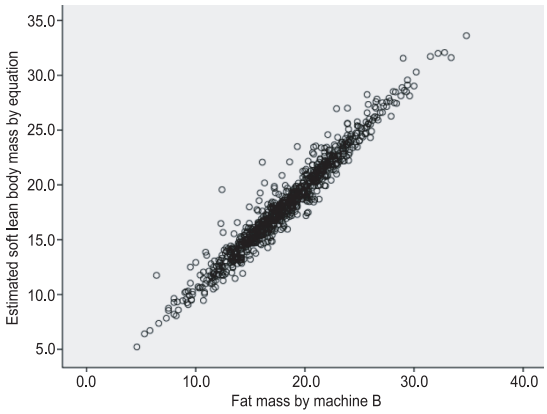
REFERENCES

1. Center for Genomic Science, National Institute of Health, KCDC. Korean Genome and Epidemiology Study (KoGES). [cited 2010 Sept 5]. Available from: http://www.nih.go.kr/bio/koges/a_a_a.jsp.
2. Yoo KY, Shin HR, Chang SH, Choi BY, Hong YC, Kim DH, et al. Genomic epidemiology cohorts in Korea: present and the future. *Asian Pac J Cancer Prev* 2005; 6(3): 238-243.
3. Gassman JJ, Owen WW, Kuntz TE, Martin JP, Amoroso WP. Data quality assurance, monitoring, and reporting. *Control Clin Trials* 1995; 16(2 Suppl): 104S-136S.
4. The Korean Society for Preventive Medicine. *Preventive Medicine and Public Health*. Seoul: Gyechuk-Munhwasa; 2010, p. 83-86. (Korean).
5. Ahn YO, Yoo KY, Park BJ. *Manual for Medical Statistics*. Seoul: Seoul National University Press; 2005, p. 105-113. (Korean).
6. Altman DG. *Practical Statistics for Medical Research*. London: Chapman & Hall.; 1992. p. 409-419.
7. Dehghan M, Merchant AT. Is bioelectrical impedance accurate for use in large epidemiological studies? *Nutr J* 2008; 7: 26.
8. Biau DJ, Halm JA, Ahmadieh H, Capello WN, Jeekel J, Boutron I, et al. Provider and center effect in multicenter randomized controlled trials of surgical specialties: an analysis on patient-level data. *Ann Surg* 2008; 247(5): 892-898.
9. Gordis L. *Epidemiology*, 3rd ed. Pennsylvenia: Elsevier Saunders; 2004. p. 87-88.
10. Bangdiwala SI, de Paula CS, Ramiro LS, Muñoz SR. Coordination of international multicenter studies: governance and administrative structure. *Salud Publica Mex* 2003; 45(1): 58-66.

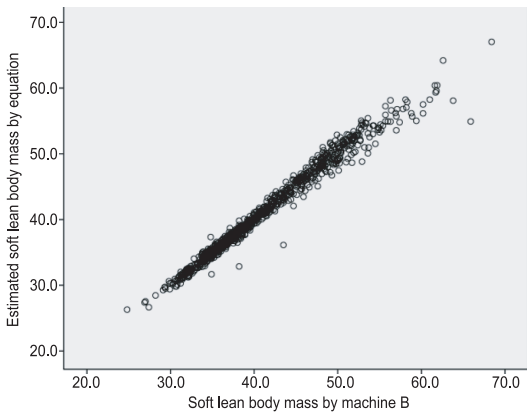
Appendix 1. Scatter plots of the test values between machine A and machine B



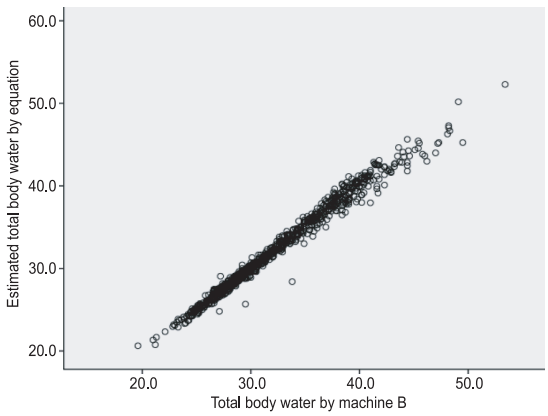
[percent body fat in machine B] = 13.266 + 0.5557[percent body fat in machine A] + 0.0421[age] + 3.1357[sex] + 0.1834[weight] -0.1187[height]



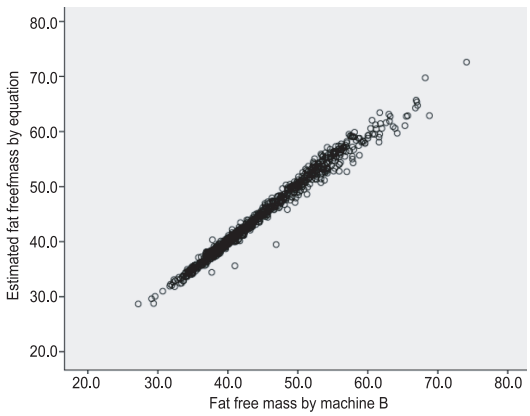
[fat mass in machine B] = -0.5038 + 0.5577[percent body fat in machine A] + 0.0252[age] + 1.9856[sex] + 0.2394[weight] -0.0682[height]



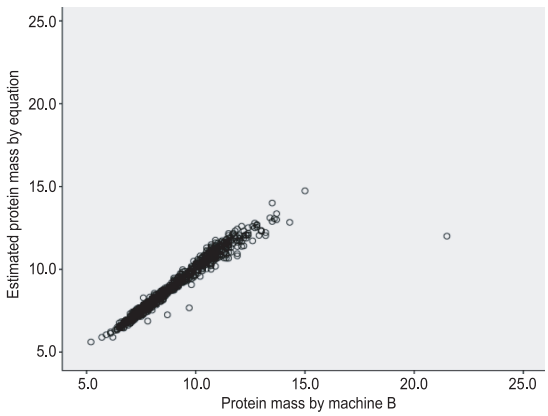
[soft lean body mass in machine B] = 1.8481 + 0.5827[soft lean body mass in machine A] - 0.027[age] - 2.0858[sex] + 0.1522[weight] + 0.06[height]



[total body water in machine B] = 1.0444 + 0.574[total body water in machine A] - 0.0187[age] - 1.522[sex] + 0.146[weight] -0.0464[height]

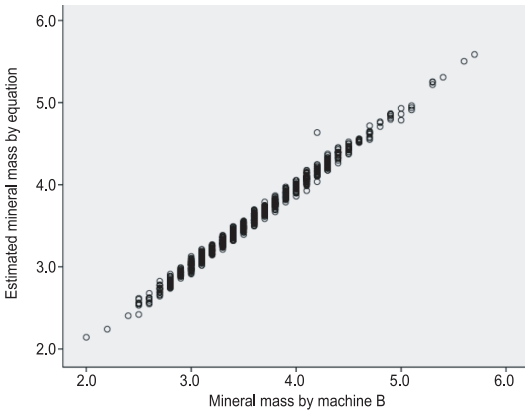


[fat free mass in machine B] = 1.1463 + 0.5588[total body water in machine A] - 0.026[age] -2.121[sex] + 0.2028[weight] -0.0643[height]

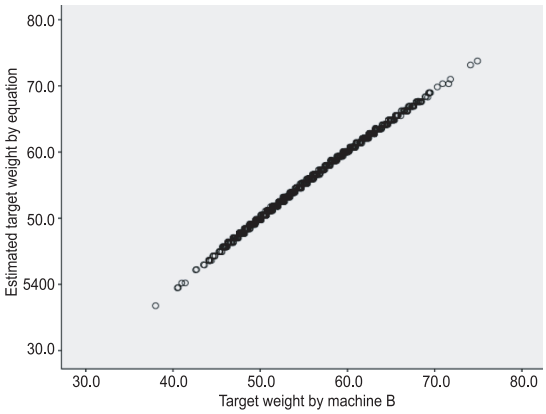


[protein mass in machine B] = 0.8442 + 0.6076[protein mass in machine A] - 0.0084[age] - 0.5671[sex] + 0.006[weight] + 0.0135[height]

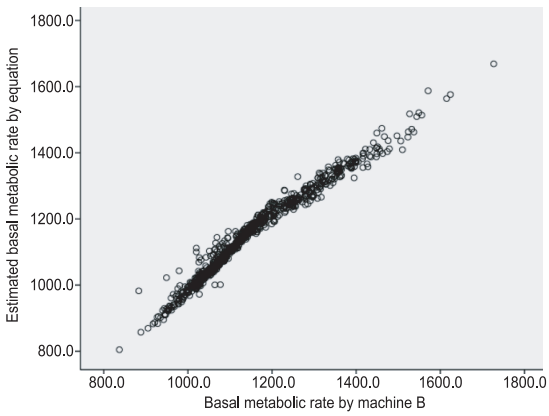
Appendix 1. Continued



[mineral mass in machine B] = -0.4285 + 0.0298[mineral mass in machine A] +0.0011[age] - 0.0318[sex] + 0.0507[weight] + 0.0045[height]



[target weight in machine B] = -49.604 + 0.0131[mineral mass in machine A] +0.0008[age] - 2.7023[sex] + 0.0007[weight] + 0.676[height]



[basal metabolic rate in machine B] = 335.11 + 0.3734[basal metabolic rate in machine A] -5.0572[age] - 35.892[sex] + 2.9011[weight] + 3.1215[height]