International Journal of Epidemiology
Vol. 26, No. 1(Suppl. 1)

Printed in Great Britain

Interviewer Variability in Anthropometric Measurements and Estimates of Body Composition

KERSTIN KLIPSTEIN-GROBUSCH*, THOMAS GEORG* AND HEINER BOEING*,**

Klipstein-Grobusch K (German Institute of Human Nutrition, Potsdam-Rehbrücke, Germany), Georg T and Boeing H. Interviewer variability in anthropometric measurements and estimates of body composition. *International Journal of Epidemiology* 1997; **26 (Suppl. 1):** S174–S180.

Objective. The extent of intra- and inter-interviewer variability both in anthropometric measurements and in estimates of body composition was assessed and the possibility of systematic variation due to interviewer differences investigated.

Design and subjects. Seventeen interviewers trained in the anthropometric measurement technique and 10 healthy volunteers (4 men, 6 women) participated in the study on measurement variability. To ensure participation of all interviewers the study was carried out on three different days. On each of these days interviewers got randomly allocated to the subjects being present. Each interviewer took 12 measurements—body weight, body height, sitting height, circumferences of waist, hip, and midarm, skinfolds (biceps, triceps, subscapular, and suprailiac), chest breadth and depth—per subject on two occasions. From these measurements, body mass index, waist-to-hip ratio, percentage of body fat, fat mass, fat free mass and metric index were determined. For all anthropometric variables variance components, reliability coefficients (R) and coefficients of variation (CV) were estimated and systematic differences of measurements between interviewers were assessed.

Results. Measurement reliability in basic anthropometric measures expressed as variance components, reliability coefficients and coefficients of variation was influenced to a greater extent by inter-interviewer variability (R : 0.858–0.999; CV : 0.1–20.9) than intra-interviewer variability (R : 0.979–0.999; CV : 0.0–6.4). The respective estimates of body composition exhibited comparatively higher reliability (R_{inter} : 0.975–0.999; R_{intra} : 0.995–0.999). Measurements more prone to subjectivity, e.g. skinfolds showed lower reliability (CV_{inter} : 9.3–20.9; CV_{intra} : 3.6–6.4). Although the absolute variation in measures due to interviewers was small, systematic differences among interviewers were clearly evident in all measurements and estimates except sitting height in this group of subjects.

Conclusion. Anthropometric measures and estimates of body composition obtained in the current study show the feasibility of detailed anthropometric data collection by multiple interviewers in large-scale epidemiological studies.

Keywords: anthropometry, body composition, interviewer variability, measurement error, precision, reliability, Germany, EPIC

Anthropometric measurements are essential as basic descriptive information on body composition and nutritional status. They are linked to energy intake, physical activity, energy metabolism and metabolic efficiency. The incidence of chronic disease may be related to anthropometric patterns, e.g. obesity has been identified as risk factor for high blood pressure and particular cancer sites and the ratio between circumferences of waist and hip was found to be associated with elevated risk for heart diseases.

In the Potsdam cohort study detailed anthropometric data of 30 000 men and women aged 35–65 years are to be collected during the recruitment period from 1994 to 1998. This study is part of the European Prospective Investigation into Cancer and Nutrition (EPIC), a large-scale prospective multi-centre study on the relationship of diet and the occurrence of cancer and other chronic diseases.[1] In addition to the EPIC core protocol including body weight, body height, sitting height, waist and hip circumferences, other measurements such as skinfold thickness and chest measurements are performed in the Potsdam cohort. Because of the size of the cohort, several interviewers are needed to perform data collection including the anthropometric measurements. Standardization, training in the anthropometric measurement technique, and regular quality control are therefore important prerequisites to ensure constant and high quality data collection during the 4 years of recruitment of the cohort.

* German Institute of Human Nutrition, Potsdam-Rehbrücke, Germany.
** Reprint requests: Dr H Boeing, German Institute of Human Nutrition, Potsdam-Rehbrücke, Unit of Medical Epidemiology, Arthur-Scheunert-Allee 114–116, D-14558 Bergholz-Rehbrücke, Germany.

Determination of intra- and inter-interviewer variability in basic measurements and estimates of body composition is important to improve measurement precision and reliability of interviewers performing the measurements. This may provide insight into the type and extent of possible measurement error that may arise both as random and systematic error from inadequate training and measurement difficulties. Unreliable measurements of the exposure variable can dilute or attenuate the observed association of exposure with the disease of interest, thereby reducing the power of the study to detect a true aetiologic association.[2] The present study was undertaken to assess the extent of variability both in basic measurements and in estimates of body composition and to investigate the possibility of systematic variation due to observer differences.

SUBJECTS AND METHODS

Study Design

In the beginning 18 interviewers trained in the anthropometric measurement technique and 10 healthy volunteers (4 men, 6 women) participated in the study on measurement variability. The characteristics of these volunteers were within the following ranges: age, 28–58 years; weight, 54.7–103.1 kg; height, 1.51–1.76 m; and body mass index (BMI), 19.4–39.3. Percentage of body fat was in the range of 11.7–43.8%, and waist-hip-ratio (WHR) ranged between 0.68 and 1.04.

The study was conducted on three different days to allow all interviewers regularly involved in anthropometry at the study centre in Potsdam to participate. On each of these days interviewers got randomly allocated to all subjects present. Each interviewer took the same 12 measurements—body weight, body height, sitting height, waist, hip, and midarm circumference, skinfolds (biceps, triceps, subscapular, and suprailiac), chest breadth and depth—per subject on two occasions. To ensure independence of the two sets of measurements, interviewers were only allowed to perform one set of measurements at one occasion. Furthermore, to ensure that measurement sites on the skin could be selected independently under field circumstances no permanent marks were allowed.

One interviewer differed systematically from all other interviewers for most measures when the deviations from the mean for measured values and estimates of body composition were taken. Subsequently, this interviewer was excluded for the current analysis and also immediately suspended from the team of interviewers performing anthropometry at the Potsdam study centre. Thus, all analyses performed and results presented are based on the remaining 17 interviewers.

Anthropometric Measurements

Body weights were measured by electronic digital scales, accurate to 100 g (Soehnle, type 7720/23, Murrhardt, Germany) with subjects wearing only light underwear and after emptying the bladder. Heights were measured to the nearest 0.1 cm using a flexible anthropometer. Hip and waist circumferences were obtained in the standing position with weight distributed equally over both feet using a full-length mirror to assist the measurement process. Waist circumference was measured midway between the lower rib margin and the superior anterior iliac spine. Hip circumference was taken at the widest point over the greater trochanters. Both were taken to the nearest 0.5 cm and measured with a non-stretching tape applied horizontally. Midarm circumference was taken at the midpoint of the right upper arm between the acromion process and the tip of the olecranon. A mark was made on the skin at this position and the circumference was measured horizontally with the non-stretching tape to the nearest 0.1 cm with the arm relaxed and hanging slightly away from the side of the body. BMI was calculated as body weight (kg) divided by squared height (m^2) and WHR as waist circumference (cm) divided by hip circumference (cm).

Skinfold thickness was measured at the right side of the body at the biceps, triceps, subscapular, and suprailiac sites using standard calipers (Lange, Cambridge MD, USA) and following standard procedures.[3] On each side measurements were taken three times and the mean of these was taken to estimate skinfold thickness. The regression equations of Durnin and Womersley[4] were used to calculate body density, which was in turn used to estimate percentage of body fat (%bf) from the equation of Siri.[5] Fat free mass (FFM) was calculated as body weight minus fat mass (FM).

Chest breadth and depth were measured horizontally using a large spreading caliper to the nearest millimetre with subjects in a natural position, weight distributed equally over both feet, and arms at the sides. The mean of minimal and maximal expiration was taken, and together with height used for calculation of metric index (MI) according to the sex-specific equations of Greil.[6] Subsequently, MI was used to classify subjects into five different body builds ranging from small to large using the classification of Möhr.[7]

Approval for the procedures used in this study was obtained from the ethical committee of the state of Brandenburg and volunteers gave their informed consent.

Statistics

Since the number of male and female subjects was small statistical procedures were generally conducted for the total group of subjects. We assumed independence between components of variation and between repeated measures. Reliability was assessed by estimating the reliability coefficients (or intra-class correlation coefficients) from variance components. According to the design of the study these were obtained by analysis of a hierarchical variance model, implicitly nesting the interviewer variable as the effect of interest into the subjects variable. From this model between-person variance σ_P^2, between-interviewer variance σ_I^2, and error variance σ_E^2 components were estimated. Inter-interviewer variability (R_{inter}) and intra-interviewer variability (R_{intra}) were then determined as $R_{intra} = 1 - \sigma_E^2/\sigma_P^2$ and $R_{inter} = 1 - \sigma_I^2/\sigma_P^2$.

Additional information concerning measurement precision independent of subject variability was obtained by estimation of inter-interviewer (CV_{inter}) and intra-interviewer (CV_{intra}) coefficient of variation:

$$CV_{inter} = \frac{\sqrt{SS_I/(n-1)}}{\bar{x}} \text{ and } CV_{intra} = \frac{\sqrt{SS_E/(n-1)}}{\bar{x}}.$$

SS_I = sum of squares interviewer
SS_E = sum of squares error
$\bar{x}$ = mean

To test for systematic differences of measurements between interviewers by means of analysis of variance, skewed data (body weight, body height, biceps skinfold, sum of skinfolds, chest breadth, BMI) were logarithmically transformed to approximate normality. Results were considered statistically different at the two-sided 0.05 alpha level. Statistical analysis was performed using SAS (Release 6.10; SAS Institute, Cary, NC, USA).

RESULTS

Variance components, reliability coefficients, and coefficients of variation for all measures and estimates of body composition are given in Table 1. As expected reliability coefficients were larger for estimates of body composition than for independent anthropometric measures. Furthermore, R_{intra} was generally larger than R_{inter} for all measures and estimates and was close to 1.00 for most measures. R_{inter} was greater than 0.98 for body weight, heights, waist and hip circumferences, and all estimates of body composition except MI. Chest measurements and the resulting MI were in the range of 0.96–0.98, and skinfolds were 0.86–0.97.

When variability of measurements was expressed as coefficient of variation comparable results were obtained with CV_{inter} generally exceeding CV_{intra} and less variation in estimates of body composition than in anthropometric measures. CV_{inter} ranged from 0.1 to 2.7 for all measures except for skinfolds, sum of skinfolds and the subsequently calculated %bf and FM. For these measures CV_{inter} was less than 5% for %bf and FM, 5–10% for triceps and sum of skinfolds, 10–15% for subscapular and suprailiac skinfolds and greater than 20% for biceps skinfold. For CV_{intra} the values were also small (<1.5%) with the exception of the skinfold measures. Still, a relatively good reproducibility for skinfold measures was revealed, since they were within a range of 5% (except biceps skinfold).

The significant differences in each basic measurement across all interviewers demonstrates that each basic measurement (except sitting height) was subject to systematic differences between observers. This was even the case for the measurement of body weight using digital scales. However, the absolute differences in measurement of body weight, heights and the resulting BMI were quantitatively very small.

DISCUSSION

This study measured the extent of interviewer variability in anthropometric measures and estimates of body composition in 17 interviewers regularly involved in anthropometry at the EPIC Potsdam study centre. Measurement reliability expressed as variance components, reliability coefficients, and coefficients of variation, was influenced to a greater extent by inter-interviewer variability than by intra-interviewer variability and was more pronounced in the anthropometric measures than in the resulting estimates of body composition. The effect due to different observers was evaluated and systematic differences among observers, e.g. inter-interviewer variability were clearly found to be evident in all measurements and estimates except sitting height in this group of subjects.

Generally, errors occurring in nutritional anthropometry can be attributed to three major effects: measurement errors, alterations in the composition and physical properties of certain tissues, and use of invalid assumptions in the derivation of body composition from anthropometric measurements. Since in epidemiological studies improvement and control of precision of measurements remains the most important means to enhance reliability the current study concentrated on the components of measurement error arising from interviewer error. Specifically, effects emerging from employment of multiple interviewers in a large epidemiological study on the components of reliability of measurement, intra- and inter-interviewer precision were of interest.

Reliability is commonly defined as the extent to which a measure is reproducible over time and can be

TABLE 1 *Variance components, reliability coefficients, and coefficients of variation of basic anthropometric measurements and estimates of body composition in EPIC-Potsdam*

Variable	No.	Mean	SEM	Sum of squares		Variance components			Reliability coefficients		Coefficient of variation		Interviewer differences[b]	
				SS_I	SS_E	σ^2_P	σ^2_I	σ^2_E	R_{inter}	R_{intra}	CV_{inter}	CV_{intra}	F-Test	Pr > F
I. Measurements														
Body weight (kg)[a]	130	68.6	4.1	0.39	0.05	157.3098	0.0032	0.0008	0.999	0.999	0.08	0.03	4.86	0.0001
Body height (cm)[a]	130	164.5	2.6	12.23	7.51	59.8406	0.0555	0.1155	0.999	0.998	0.19	0.15	2.13	0.0122
Sitting height (cm)	130	86.9	1.3	18.61	16.36	15.0910	0.0465	0.2517	0.997	0.983	0.44	0.43	1.30	0.2108
Hip circumference (cm)	130	98.4	2.8	167.49	31.87	71.7248	1.3056	0.4904	0.982	0.993	1.16	0.51	8.46	0.0001
Waist circumference (cm)	130	82.6	4.9	245.39	65.38	224.7049	1.7693	1.0058	0.992	0.995	1.67	0.86	4.14	0.0001
Midarm circumference (cm)	130	28.9	1.4	60.22	9.42	17.1340	0.4852	0.1449	0.972	0.991	2.36	0.93	4.08	0.0001
Biceps skinfold (mm)[a]	130	9.5	1.8	510.42	48.00	30.6839	4.3568	0.7385	0.858	0.979	20.92	6.42	7.60	0.0001
Triceps skinfold (mm)	130	17.7	2.9	348.83	76.50	78.5880	2.6415	1.1769	0.966	0.986	9.28	4.11	3.28	0.0001
Subscapular skinfold (mm)	130	17.0	3.1	604.46	48.00	90.5054	5.2276	0.7385	0.942	0.992	12.71	3.58	9.36	0.0001
Suprailiac skinfold (mm)	130	20.6	4.3	967.15	114.00	166.7759	8.0781	1.7538	0.952	0.990	13.26	4.55	2.06	0.0155
Sum of skinfolds (mm)[a]	130	64.8	11.4	4337.02	278.50	1198.8368	38.0153	4.2846	0.968	0.997	8.95	2.27	9.26	0.0001
Chest breadth (cm)[a]	130	26.9	0.9	37.04	6.27	7.6610	0.2948	0.0964	0.962	0.988	1.99	0.82	7.46	0.0001
Chest width (cm)	130	19.5	1.2	36.52	6.68	12.5981	0.2868	0.1027	0.977	0.992	2.74	1.17	10.65	0.0001
II. Estimates of Body Composition														
Body mass index (BMI)[a]	130	25.2	1.8	1.22	0.57	28.4272	0.0068	0.0088	0.999	0.999	0.38	0.26	3.46	0.0033
Waist-hip-ratio (WHR)	130	0.8	0.0	0.03	0.01	0.01366	0.00019	0.00017	0.986	0.988	1.80	1.10	2.30	0.0063
Body fat (%)	130	27.5	3.3	157.55	16.95	99.1769	1.3284	0.0261	0.987	0.999	4.02	1.32	9.18	0.0001
Fat mass (FM)	130	19.5	3.2	76.39	7.33	95.7101	0.6509	0.1127	0.993	0.997	3.94	1.22	8.81	0.0001
Fat free mass (FFM)	130	49.1	2.2	75.65	7.54	42.8739	0.6425	0.1160	0.985	0.997	1.56	0.49	8.67	0.0001
Metric Index (MI)	130	–0.58	0.3	2.40	0.28	0.8808	0.0214	0.0044	0.975	0.995	–	–	14.35	0.0001

[a] Logarithmically transformed for approximation of normality of skewed values.
[b] Adjusted for subjects.

divided into two parts, the precision referring to the freedom from measuring error and short-term random fluctuation.[8] Reliability was estimated by means of reliability coefficients and coefficients of variation for intra- and inter-interviewer effects. Reliability coefficients are greatly affected by between-person variance which in turn is affected by age, sex, and composition of the sample of the subjects. The high proportion of between-subject variance compared to between- and within-interviewer variance reflects the heterogeneity of the sample of subjects in the study of variability. The coefficients of variation (CV) are by definition characterized by the measurement error as a proportion of sample mean and thus reflect measurement variability independent of between-subject variance.

It has been argued that both indicators of precision, R and CV are affected by variable size, a phenomenon that is referred to as effect of scale,[8] i.e. error increases as the size of the individual measurements increases. Given our small subject sample size (n = 10) a thorough investigation of effect of scale was impossible. It seemed that imprecision was rather determined by subjectivity of measurement than effect of scale. Nevertheless, effect of scale may have been present in the measurement of biceps skinfold and could explain the relative high CV compared to the other skinfold measurements.

Many studies have dealt with the reliability of common measures such as body weight, body height, and circumferences.[9–11] Specifically, in epidemiological studies this was done in an attempt to evaluate the usefulness of self-reported values.[12–15] However, the number of interviewers included was mostly small and studies reporting on inter-interviewer variability have been scarce. In the current study inter- and intra-interviewer variability was determined for in total 17 interviewers for a variety of anthropometric measures and estimates of body composition. The reliability coefficients and coefficients of variation obtained indicated for most variables and for skinfolds and resulting estimates a moderate reliability. Deviation of individual measurements from the mean values was small for most measures, indicating a high degree of precision of measurements.

In comparison, the amount of measurement error attributed to intra-interviewer variability was generally smaller than for inter-interviewer variability. These results were in agreement with the findings of Mueller and Malina[16] and Ferrario[9] (for skinfolds only). As expected, reliability of more objective measurements such as body weight and body height was higher than for those including subjective judgement such as girth measurements or skinfold measurements (localization of skinfolds to measure skinfold thickness). However, when skinfold measurements were combined some of the random interviewer variation of the individual skinfold measurements was cancelled out. This resulted in a considerably lower variability for sum of skinfolds and estimated percentage of body fat. Consequently, use of sum of skinfolds rather than individual skinfold measurements for calculation of estimates[17,18] is advised.

To allow comparison with previous studies, inter-interviewer and intra-interviewer estimates of variance, reliability coefficients and coefficients of variations for selected studies are presented in Table 2. Intra-interviewer variation both expressed as reliability coefficient and coefficient of variation obtained in Potsdam were comparable to those in other studies.[8,9] Compared to studies reporting on inter-individual reliability[9,16] the estimates in Potsdam showed greater reliability especially for the skinfold measurements. CV for triceps and subscapular skinfold estimated by Ferrario *et al*.[9] exceeded clearly the comparable CV obtained in Potsdam. Although interviewers differed significantly from each other for all anthropometric measures and estimates of body composition (except sitting height), precision and reliability in the current study yielded more favourable results than most of the other studies referred to in Table 2. Significance of inter-interviewer error was demonstrated in several studies,[10,19] leading to the conclusion that difference between interviewers may represent the major source of error. However, during the 4-year recruitment period in EPIC-Potsdam error due to differences between interviewers may result in random rather than systematic error. Therefore, the main focus should be on high technical skills in taking anthropometric measurements of all interviewers to ensure accurate measurement values. The established standardization scheme which is conducted routinely may help to achieve this aim. Furthermore, repeated reproducibility studies will not only result in technical adequacy of anthropometric measurements but also in evaluation whether the difference between interviewers are persistent during the recruitment period.

In summary, we found relatively good precision and reliability of most anthropometric measures and estimates of body composition conducted by the 17 interviewers participating in the current study. The moderate degree of measurement error indicates that in large epidemiological studies with multiple interviewers as implemented in the EPIC Potsdam study susceptibility for misclassification of continuous anthropometric variables can be kept low through comprehensive and continuous standardization of measurement technique.

TABLE 2 *Inter-interviewer and intra-interviewer estimates of body composition measurements: measurement error variance, reliability, coefficient, and coefficient of variation from published results of selected studies*

Study	Description of study	Measures	No.	Mean	Intra-interviewer				Inter-interviewer			
					σ^2_B[a]	σ^2_{intra}[b]	R	CV%	σ^2_B[a]	σ^2_{inter}[c]	R	CV%
Jackson *et al.*[17]	Healthy males;	Skinfolds:										
	21–39 years;	Triceps (mm)	35	16.1			0.976					
	repeated measurements	Subscapular (mm)	35	21.0			0.977					
	on 3 different days	Suprailiac (mm)	35	19.9			0.974					
	taken by 3 different											
	interviewers	Circumference:										
		Waist (cm)	35	94.3			0.997					
Mueller and	HES cycle III Study;	Skinfolds:										
Malina[16]	non-random sample of	Triceps (mm)	77/224[d]		34.545	0.640	0.981		34.545	3.572	0.897	
	12–17 year old adolescents;	Subscapular (mm)	77/224[d]		39.094	3.349	0.914		39.094	2.341	0.940	
	repeated measures taken	Suprailiac (mm)	77/224[d]		76.431	3.497	0.954		76.431	6.002	0.921	
	2–3 weeks apart by											
	several interviewers	Circumferences:										
		Waist (cm)	77/224[d]		60.464	1.711	0.972		60.464	2.437	0.970	
		Hip (cm)	77/224[d]		62.951	1.523	0.976		62.951	1.891	0.957	
Marks *et al.*[8]	NHANES II:	Males:										
	males 40 (±14) years,	Weight (kg)	92	75.2	915.35	1.45	1.00	1.6				
	females 38 (±13) years;	Height (cm)	84	176.1	51.56	0.11	1.00	0.2				
	repeated measurements	Sitting height (cm)	93	93.3	11.72	0.29	0.98	0.6				
	taken 2 and more											
	weeks apart by 10	Skinfolds:										
	interviewers	Triceps In(mm)	93	2.46	0.23	0.03	0.88	7.2				
		Subscapular In(mm)	92	2.77	0.27	0.02	0.95	4.3				
		Females:										
		Weight (kg)	129	66.1	848.08	1.68	1.00	2.0				
		Height (cm)	127	163.2	38.18	0.16	1.00	0.2				
		Sitting height (cm)	128	87.2	11.44	0.28	0.98	0.6				
		Skinfolds:										
		Triceps In(mm)	129	3.15	0.13	0.03	0.81	5.5				
		Subscapular In(mm)	129	2.92	0.25	0.02	0.93	4.7				
Rimm *et al.*[12]	Health Professional	Males:	123									
	Study and Nurses	Weight (kg)		82.0	656.7	16.9	0.97	5.0				
	Health Study;	Hip (cm)		94.0	73.5	9.0	0.89	3.2				
	males 40–75 years,	Waist (cm)		102.6	36.1	4.5	0.88	2.1				
	females 41–65 years;	WHR		0.92	1.6	9.0	0.68	3.1				
	repeated measures taken											
	6–8 months apart by	Females:	140									
	2 interviewers	Weight (kg)		66.8	700.8	15.0	0.98	5.8				
		Hip (cm)		79.0	113.5	11.6	0.91	4.3				
		Waist (cm)		101.9	75.5	10.3	0.88	3.2				
		WHR (cm)		0.77	2.8	0.75	0.78	3.6				
Sonnenschein	New York University	Weight (kg)	1851	66.0	137.3	8.0	0.95	4.3				
et al.[11]	Women's Health Study;	Height (cm)		161.8	37.4	0.7	0.98	0.5				
	women; 36–65 years;	BMI		25.5	18.1	1.2	0.94	4.3				
	3–6 repeated measures	Waist (cm)		76.1	124.0	16.0	0.89	5.3				
	taken by one inter-	Hip (cm)		100.0	85.2	19.4	0.81	4.4				
	viewer over 3 years	WHR		0.76	4.0	1.0	0.74	4.2				

Continued

TABLE 2 *continued*

Study	Description of duty	Measures	No.	Mean	Intra-interviewer				Inter-interviewer			
					σ^2_B [a]	σ^2_{intra} [b]	R	CV%	σ^2_B [a]	σ^2_{inter} [c]	R	CV%
Ferrario *et al*.[9]	ARIC Study; adults, 45–65 years; random sample selection; one resp. two interviewers	Skinfolds:										
		Triceps (mm)	378/411[d]	24.8	111.8	1.3	0.99	4.6	127.6	13.4	0.81	15.2
		Subscapular (mm)	378/411[d]	24.0	133.9	1.2	0.99	4.6	228.2	55.1	0.91	20.6
		Circumferences:										
		Waist (cm)	212/146[d]	97.8	188.2	5.2	0.97	2.4	193.0	3.6	0.98	1.9
		Hip (cm)	212/146[d]	105.3	108.3	0.97	0.99	0.9	101.3	1.5	0.99	1.2
		WHR	212/146[d]	0.93	6.6	0.7	0.91	2.8	5.6	0.3	0.94	2.0

[a] Between-subject variance.
[b] Inter-interviewer variance.
[c] Intra-interviewer variance.
[d] No. of subjects for analysis of intra/inter variance.

ACKNOWLEDGEMENTS

The authors thank the participants and the interviewers of the EPIC-Study Potsdam for their co-operation. The study was supported in part by the Federal Ministry of Technology, Germany and by funds of the 'Europe against cancer program' of the European community; Grant No. SOC 95 201408 05F02.

REFERENCES

1 Riboli E: Nutrition and Cancer: Background and rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann Oncol* 1992; **3:** 783–91.

2 Flegal K M, Keyl P M. Differential misclassification arising from non-differential error in exposure measurement. *Am J Epidemiol* 1991; **134:** 1233–44.

3 Lohman T G, Roche F, Reynaldo M. *Anthropometric Standardization Reference Manual*. Champaign, IL: Human Kinetics Books, 1988.

4 Durnin J V G A, Womersley J. 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* 1974; **32:** 77–97.

5 Siri W E. Body composition from fluid spaces and density: analysis of methods. In: Brozek J and Henschel A (eds). *Techniques for Measuring Body Composition*. Washington, DC: National Academy of Sciences NRC, 1961, pp. 223–44.

6 Greil H. *Der Körperbau im Erwachsenenalter—DDR repräsentative anthropologische Querschnittsstudie 1982/1984*. Habilitation. Berlin: Humboldt University, 1987. (In German).

7 Moehr M. Untersuchungen zur Beurteilung des Ernährungszustandes Erwachsener. *Zärztl Fortbild* (In German) 1976; **70:** 843–52 .

8 Marks G E, Habicht J-P, Mueller W H. Reliability, dependability, and precision of anthropometric measurements. The Second National Health and Nutrition Examination Survey 1976–1980. *Am J Epidemiol* 1989; **130:** 578–87.

9 Ferrario M, Carpenter M A, Chambless L E: Reliability of body fat distribution measurements. The ARIC study baseline cohort results. *Int J Obesity* 1995; **19:** 449–57.

10 Fuller N J, Jebb S A, Goldberg G R, Pullicino E, Adams C, Cole T J, Elia M. Inter-observer variability in the measurement of body composition. *Europ J Clin Nutr* 1991; **45:** 43–49.

11 Sonnenschein E G, Kim M Y, Pasternack S, Toniolo P G. Sources of variability in waist and hip measurements in middle-aged women. *Am J Epidemiol* 1993; **138:** 301–09.

12 Rimm E B, Stampfer M J, Colditz G A, Chute C G, Litin L B, Willett W: Validity of self-reported waist and hip circumferences in men and women. *Epidemiology* 1990; **1:** 466–73.

13 Perry G S, Byers T E, Mokdad A H, Serdula M K, Williamson D F. The validity of self-reports of past body weights by US adults. *Epidemiology* 1995; **6:** 61–66.

14 Rowland M L. Self-reported weight and height. *Am J Clin Nutr* 1990; **52:** 1125–33.

15 Troy L M, Hunter D J, Colditz G A, Stampfer M J, Willett W C. The validity of recalled weight among younger women. *Int J Obesity* 1995; **19:** 570–72.

16 Mueller W H, Malina R M. Relative reliability of circumferences and skinfolds as measures of body fat distribution. *Am J Phys Anthr* 1987; **72:** 437–39.

17 Jackson A S, Pollock M L, Gettman L R. Intertester reliability of selected skinfold and circumference measurements and percent fat estimates. *Res Q* 1978; **4:** 546–51.

18 Womersley J, Durnin J V G A. An experimental study on the variability of measurements of skinfold thickness in young adults. *Hum Biology* 1973; **45:** 281–88.

19 Housh T J, Johnson G O, Thorland W G *et al*. Validity and intertester error of anthropometric estimations of body density. *J Sports Med Phys Fitness* 1989; **29:** 149–56.