Technical error of measurement in anthropometry*

Talita Adão Perini[1], Glauber Lameira de Oliveira[1], Juliana dos Santos Ornellas[1] and Fátima Palha de Oliveira[2]

ABSTRACT

The anthropometrical measurements have been widely utilized to follow children's development, in the verification of the adaptations to the physical training in the athletes' selection, in studies of ethnic characterization, among others. The control of the precision and accuracy of the measurements will result in more reliable data. The objective of the present study was to diffuse the strategies to compute the technical error of measurement (TEM) according to Kevin Norton's and Tim Olds methodology (2000) and to analyze the laboratory' trainees performance. Three beginner observers (anthropometrists) of the Exercise Physiology Laboratory (Labofise) of the University of Brazil were analyzed. They accomplished measures of skin folds thickness (Cescorf, 0.1 mm) in nine different anthropometric points in 35 volunteers (25.45 ± 9.96 years). To accomplish the measures the International Society for Advancement in Kinanthropometry (ISAK) was adopted. For the TEM intra-evaluator verification, the measures were accomplished in the same volunteers in two different days and, to obtain the inter-observers TEM, the measures were accomplished in a same group of volunteers, in the same day by the three evaluators. The results indicated *not acceptable* TEMs only for two evaluators in the *intra-evaluator analysis*. The other TEMs reached acceptable results. *Not acceptable* TEMs demonstrated the need of technical training of evaluators in order to minimize the variability verified.

Key words: Precision. Anthropometrical measurements. Intra-evaluators. Inter-evaluators.

INTRODUCTION

The physical growth occurs according to a characteristic sequence associated to biological aspects of the development. The monitoring of the anthropometrical measurements during the growth process allows the qualification of the morphological variations as result of this process, providing data for the diagnosis of possible deficiencies[1,2].

In the sportive environment, besides expressing the athlete's physical proportionalities, the anthropometrical measurements, when performed periodically, are strong indicators of the adaptive response of the organism to stimuli from physical training[3].

In Engineering, the anthropometrical measurements are vital, once they guide the development of ergonomic projects aimed at the manufacturing of machines, tools and devices adapted to the human condition. The manufacturing of medical materials, as well, have to be based on populational anthropometrical studies, so that crutches, walking sticks, walkers and others suit perfectly to the patients' physical characteristics. Researches in Bioengineering have used anthropometrical measurements in the optimization and manufacturing of orthopedic prostheses and in equipments to test the product developed.

In clothing and shoes industries, the knowledge of the anthropometrical characteristics is required, so that products fulfill the different physical characteristics of the population. Anthropometrical studies with regard to future users of projected products become, therefore, necessary.

Despite the anthropometry applications in the daily life of individuals being so many and diverse, it is known that there is an error margin in the method. When anthropometrical measures are repeated, a variability of the measures may occur as result of the physical characteristics diversity of the population analyzed, due to the biological variation – that cannot be avoided – or due to technical variations – that can be avoided. The variability on the anthropometrical measurements caused by variations on the technique execution is responsible for the higher incidence of error. The adoption of inadequate time interval between measurements, the variation on the marking of anatomical points and the inconsistency of the measurement technique executed are some examples of technical imperfections. The improvement on the execution technique and the guaranty of higher accuracy may be obtained by the intensive training of anthropometrists.

The most common way to express the error margin in anthropometry is by means of the *technical error of measurement* (TEM), which is an accuracy index and represents the measurement quality and control dimension. The TEM index allows anthropometrists to verify the accuracy degree when performing and repeating anthropometrical measurements (intra-evaluator) and when comparing their measurement with measurements from other anthropometrists (inter-evaluator). This index is adopted by the International Society standardization Advancement in Kinanthropometry (ISAK) for the accreditation of anthropometrists in Australia.

The TEM index, which is the standard deviation between repeated measures[4], is used for the calculation of the *intra-evaluator* variability – variation of repeated measurements of the same person (or a group of persons) performed by the same anthropometrist – and *inter-evaluator* – variation of measurements performed by different anthropometrists in the same group of persons. Alterations of the anthropometrical measurements due to the *inter-evaluator* variation may be present in results of measurements performed in sportive academies or physical conditioning centers where more than one professional work in the evaluation of the public. In this situation, the periodic TEM checking is important in order to verify whether or not this index is within acceptable standards.

The TEM calculation allows the estimation if confidence intervals around the actual value of the measurement obtained that includes non-controllable possible variations – the biological variations, for example – thus enabling to verify if the alterations detected in repeated measurements before and after a training ses-

* University of Brazil (UFRJ) – Rio de Janeiro – Brazil.

1. Physical Education Course underway and scientific initiation trainee – Labofise-UFRJ.
2. Assistant Professor, Coordinator of the Ergospirometry and Anthropometry sector – Exercise and Physiology Laboratory (Labofise-UFRJ).

Received in 1/7/04. 2nd version received in 6/9/04. Approved in 25/1/05.

Correspondence to: Profª Drª Fátima Palha de Oliveira, Rua das Laranjeiras, 136, aptº 203, Laranjeiras – 22240-000 – Rio de Janeiro, RJ – Brazil. Phones: home, (21) 2205-7299; mobile, 9382-3844; office, (21) 2562-6850; fax, (21) 2562-6801. E-mail: palha@ufrj.br

sion are a result of this training or a result of the method's relative variation.

Considering the aspects mentioned above, the present study has the objective of presenting the TEM calculation methodology obtained through the method of the differences, besides analyzing the performance of three beginner anthropometrists of our laboratory for skin folds measurements. In this context, there is also the objective of contributing for the diffusion of the anthropometrical measurement accuracy importance that will result in the production of more reliable populational data.

MATERIAL AND METHODS

Three anthropometrists from the Exercise and Physiology Laboratory (Labofise-UFRJ) were analyzed through the TEM calculation for their measurements results with the objective of verifying the *intra-evaluator* and *inter-evaluator* variation. This study was conducted by anthropometrists after a period of theoretical orientation and practical experimentation of the different anthropometrical measurements adopted at the Labofise.

The anthropometrical measurements were extracted from a sample of 35 volunteers (25.45 ± 9.96 years) of both genders (12 women and 23 men) who declared to be inactive. All individuals live in Rio de Janeiro and signed a free consent form that included the procedures to be adopted and the volunteers allowed the exploitation of the results found in scientific studies. The participants' privacy and anonymity were respected in the present study.

In order to obtain the *intra-evaluator* TEM, each anthropometrist measured 20 out of the 35 volunteers in two different days always in the morning. For the *inter-evaluator* TEM calculation, the results of measurements performed within a same group of 20 volunteers were analyzed in the same day by each one of the three anthropometrists.

The TEM values from each anthropometrist were calculated for the following skin fold measurements: triceps, biceps, thorax, abdomen, subscapular, suprailliac, thigh, medial-axillary and leg. The measurements were performed according to the International Society for Advancement in Kinanthropometry – ISAK[5]. For each anthropometrical point considered, three non-consecutive measurements were performed in order to extract the average and hence to compute the thickness result of the point analyzed.

The same procedure was used for measurements performed by all anthropometrists (Cescorf, 0.1 mm). All calculations were conducted in the *Excel* software (Microsoft 2000).

TEM calculation

In order to obtain *intra-evaluator* and *inter-evaluator* TEM, the following aspects must be observed:

- TEM always presents the same measurement unit (cm, mm);
- TEM is only applied to the measurement performed and to the equipment used (i.e., triceps fold or arm perimeter, performed with a given device model and label);
- TEM is only applied in similar population, in other words, it is calculated for athletes;
- In order to calculate TEM, one should consider at least 20 measurements that must be performed at the same moment (morning/afternoon).

The method of differences was adopted for the attainment of TEM, which is expressed through the standard deviation between repeated measurements[4]. This deviation is the dispersion degree of values in relation to the average. The TEM calculation was divided into four stages that will be presented below for a better understanding of the method.

➢ ***Intra-evaluator* TEM calculation**

In order to perform the *intra-evaluator* TEM calculation, the results of the skin folds measurements of 20 volunteers were considered at the first and second evaluation day. As example, table 1 shows the values obtained for the abdomen skin fold measurement and all following steps for the intra-evaluator TEM calculation with regard to this measurement.

First stage: The ***difference*** between the 1st and 2nd measurements was determined (***deviation*** between them) for each anthropometrical point considered of all volunteers measured by the ***same*** anthropometrist.

Second stage: The deviations obtained were raised to the second power.

Third stage: The results of the second stage were summed (Σd^2) and applied to equation 1 in order to obtain the absolute TEM.

$$\text{AbsoluteTEM} = \sqrt{\frac{\Sigma\, d_i^2}{2n}} \qquad \text{Equation 1}$$

Where:
Σd^2 = summation of deviations raised to the second power
n = number of volunteers measured
i = the number of deviations

Fourth stage: The absolute TEM was transformed into relative TEM in order to obtain the error expressed as percentage corresponding to the total average of the variable to be analyzed. So, the equation 2 was used. In this stage, it was necessary to obtain the variable average value (VAV). To do so, the arithmetic mean of the mean between both measurements obtained (1st and 2nd measurements) of each volunteer for the same skin fold. In other words, the measurement performed at the first and second day of the same volunteer for a give skin fold was summed up and then divided by two, thus generating the average of this fold. This procedure was performed for each one of the 20 volunteers and the 20 averages obtained were summed up and divided by 20 (total of volunteers) – generating VAV (table 1).

$$\text{relative TEM} = \frac{\text{TEM}}{\text{VAV}} \times 100 \qquad \text{Equation 2}$$

Where
TEM = Technical error of measurement expressed in %
VAV = Variable average value

➢ ***Inter-evaluator* TEM calculation**

The four steps previously described for the *intra-evaluator* TEM calculation must be followed in order to perform the *inter-evaluator* TEM calculation. The procedures are the same, but the skin folds measurements to be considered should be performed by the anthropometrists to be evaluated in the same group of volunteers. In the present study, 20 volunteers (randomly selected among the 35 participants of the study) were measured by the three anthropometrists within the same day and shift and one anthropometrist could not see the measurement performed by another. The same equipments and methodological procedures for skin folds measurements were adopted by all anthropometrists. The *inter-evaluator* TEM was achieved by comparing two anthropometrists each time.

TEM classification

After the calculation of the relative TEM, for both *intra-evaluator* and *inter-evaluator* variation analysis, the next step is to classify it (table 3). It is important observing that the lower the TEM obtained, the better is the accuracy of the appraiser to perform the measurement.

RESULTS

The description of the physical characteristics of volunteers from this study is found in table 4.

TABLE 1
Results of abdomen skin folds thickness performed by the same anthropometrist in 20 volunteers for the intra-evaluator TEM calculation

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1st measurement	31.0	6.0	23.0	42.0	13.8	22.3	12.7	21.7	45.0	24.0	29.5	43.8	26.8	11.5	41.2	21.8	27.0	26.7	13.2	7.2
2nd measurement	35.0	7.7	27.0	40.3	12.3	24.0	13.3	21.3	43.3	25.3	31.7	42.8	25.5	11.0	42.5	22.7	26.0	25.5	12.7	8.2
Deviations	–4.0	–1.7	–4.0	1.7	1.5	–1.7	–0.7	0.3	1.7	–1.3	–2.2	1.0	1.3	0.5	–1.3	–0.8	1.0	1.2	0.5	–1.0
$(Deviations)^2$	16.0	2.8	16.0	2.8	2.3	2.8	0.4	0.1	2.8	1.8	4.7	1.0	1.8	0.3	1.8	0.7	1.0	1.4	0.3	1.0
$\Sigma (Deviations)^2$	61.8																			
Absolute TEM	1.24																			
VAV	24.71																			
Relative TEM %	5.02 acceptable																			

TEM = Technical error of measurement; Σ = Summation; VAV = Variable average value.

TABLE 2
Results of abdomen skin folds thickness performed by two anthropometrists in 20 volunteers for the inter-evaluator TEM calculation

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Anthropometrist 1	28.6	9.7	20.0	10.0	10.8	17.5	23.5	30.0	8.5	21.5	22.1	25.5	26.0	14.8	7.5	24.5	10.0	21.0	10.0	22.0
Anthropometrist 2	29.0	10.0	20.3	10.0	11.5	17.0	23.0	30.0	8.0	21.0	22.0	25.3	25.8	15.0	7.3	25.0	10.0	21.2	11.0	22.0
Deviations	–0.4	–0.3	–0.3	0.0	–0.7	0.5	0.5	0.0	0.5	0.5	0.1	0.2	0.2	0.2	0.2	–0.5	0.0	–0.2	–1.01	0.0
$(Deviations)^2$	0.16	0.09	0.09	0.00	0.49	0.25	0.25	0.0	0.25	0.25	0.01	0.04	0.04	004	0.04	0.25	0.0	0.04	1.00	0.0
$\Sigma (Deviations)^2$	3.29																			
Absolute TEM	0.29																			
VAV	18.20																			
Relative TEM %	1.58 acceptable																			

TEM = Technical error of measurement; Σ = Summation; VAV = Variable average value VAV = [(average 1st measurement/2nd measurement)/2].

TABLE 3
Relative TEM values considered as acceptable

Type of analysis		Beginner anthropometrist	Skilful anthropometrist
Intra-evaluator	Skin folds	7.5%	5.0%
	Other measures	1.5%	1.0%
Inter-evaluator	Skin folds	10%	7.5%
	Other measures	2.0%	1.5%

Gore et al. *in* Kevin Norton e Tim Olds (2000)

TABLE 4
Physical characteristics of the sample (n = 35)

	Average	Standard deviation	Minimum value	Maximum value
Age (years)	25.45	± 9.96	15	57
Body mass (kg)	67.40	± 16.25	48	103
Stature (m)	1.70	± 0.09	1.53	1.85
Sum of the 9 skinfolds (mm)	165.38	± 78.23	61	332.5

CHART 1
Intra-evaluator relative TEM classification and results

ABDOMEN

Anthropometrist	TEM%	Classification
1	5.0	Acceptable
2	5.1	Acceptable
3	3.9	Acceptable

SUBSCAPULAR

Anthropometrist	TEM%	Classification
1	4.9	Acceptable
2	3.6	Acceptable
3	8.5	Non-acceptable

SUPRA-ILLIAC

Anthropometrist	TEM%	Classification
1	7.2	Acceptable
2	5.0	Acceptable
3	6.9	Acceptable

THIGH

Anthropometrist	TEM%	Classification
1	6.0	Acceptable
2	7.8	Non-acceptable
3	6.1	Acceptable

MEDIAL-AXILLARY

Anthropometrist	TEM%	Classification
1	3.2	Acceptable
2	7.1	Acceptable
3	4.4	Acceptable

LEG

Anthropometrist	TEM%	Classification
1	3.0	Acceptable
2	4.7	Acceptable
3	5.1	Acceptable

TRICEPS

Anthropometrist	TEM%	Classification
1	3.2	Acceptable
2	3.5	Acceptable
3	5.7	Acceptable

BICEPS

Anthropometrist	TEM%	Classification
1	3.3	Acceptable
2	3.4	Acceptable
3	4.0	Acceptable

THORAX

Anthropometrist	TEM%	Classification
1	3.7	Acceptable
2	4.2	Acceptable
3	4.6	Acceptable

CHART 2
Inter-evaluator relative TEM classification and results

ABDOMEN		
Anthropometrist	TEM%	Classification
1 and 2	8.2	Acceptable
1 and 3	8.5	Acceptable
2 and 3	6.9	Acceptable

SUBSCAPULAR		
Anthropometrist	TEM%	Classification
1 and 2	5.9	Acceptable
1 e 3	5.3	Acceptable
2 e 3	4.7	Acceptable

SUPRA-ILLIAC		
Anthropometrist	TEM%	Classification
1 and 2	6.6	Acceptable
1 and 3	7.1	Acceptable
2 and 3	9.0	Acceptable

THIGH		
Anthropometrist	TEM%	Classification
1 and 2	6.3	Acceptable
1 and 3	8.1	Acceptable
2 and 3	8.1	Acceptable

MEDIAL-AXILLARY		
Anthropometrist	TEM%	Classification
1 and 2	4.7	Acceptable
1 and 3	4.9	Acceptable
2 and 3	4.1	Acceptable

LEG		
Anthropometrist	TEM%	Classification
1 and 2	4.1	Acceptable
1 and 3	4.0	Acceptable
2 and 3	4.9	Acceptable

TRICEPS		
Anthropometrist	TEM%	Classification
1 and 2	1.7	Acceptable
1 and 3	4.7	Acceptable
2 and 3	5.8	Acceptable

BICEPS		
Anthropometrist	TEM%	Classification
1 and 2	5.8	Acceptable
1 and 3	5.0	Acceptable
2 and 3	5.9	Acceptable

THORAX		
Anthropometrist	TEM%	Classification
1 and 2	5.8	Acceptable
1 and 3	6.5	Acceptable
2 and 3	5.9	Acceptable

Charts 1 and 2 show the results of the calculations performed for each one of the nine skin folds considered in this study. They present relative TEMs of each anthropometrist for each skin fold of the intra-evaluator (chart 1) and inter-evaluator (chart 2) variability analysis.

The standard adopted for the evaluation of the TEM found was the beginners standard (table 3), once the anthropometrists of this study are trainee graduation students. The results indicate acceptable variability in the accuracy of measurements of most skin folds for all anthropometrists. Unacceptable values were only observed in the subscapular skin fold for anthropometrist # 3 and thigh skin fold for anthropometrist # 2. The inter-observer variability presented acceptable results in all points analyzed.

It is worth emphasizing that, despite results being acceptable, a higher variation on the relative TEM may be observed in anatomical points where the thickness of the fat tissue layer is higher[6].

We have observed, along many years of practice in anthropometry that in these sites (supra-illiac, abdomen), the difficulty to measure skin fold thickness is higher in function of the higher fat accumulation, what may lead to more frequent measurement errors. In these cases, besides the precise localization of the anatomical point to be measured, it is recommended to ask for the aid of an assistant to grasp the fold with both hands in order for the anthropometrist to perform the measurement[7].

DISCUSSION

The skin folds measurements have been employed to estimate the body fat level, being considered as an adequate method for the evaluation of a large number of people due to its easy execution, low cost and relative accuracy.

The highest accuracy in this measurement is associated to the employment of calibrated devices, the presence of skilful anthropometrists and to the periodic accuracy control of both[8].

The accuracy of the anthropometrical measurements is associated to the minimization of the systematic error probably in function of either the equipment or the anthropometrist. The anthropometrical measurements are even subjected to non-systematic errors with regard to occurrences beyond control. In the specific case of skin folds measurements, the non-systematic error is related to presuppositions in which the method is based on and what its limitations are[9,10].

The employment of skin folds measurements is conditioned to the presupposition that the skin folds composed of skin and fat, even when compressed, may represent fat from non-compressed subcutaneous layers. It is observed, however, that the skin thickness and its intra and inter-individuals variations are not considered in this presupposed as well as the compressibility variations of the fat layer that changes according to the place measured, age, gender, hydration level, cells size and health state[11].

These presupposed compose the specific error of the anthropometrical method for the body fat estimation. Once these limitations are known, it becomes necessary to minimize the occurrence of errors both in the execution of the skin folds thickness measurements technique and in the control of the equipment calibration, which compose the errors passible of intervention.

The calibration of the equipment (plicometer) used in the skin folds measurements must be periodically performed. The shaft (central factor) around which the plicometer's movable nipper moves must annually calibrated, and the accuracy in the separation between nippers must be verified each six months[12].

With regard to the improvement of the anthropometric measurement technique, one knows that it is directly related with the number of evaluations performed by the anthropometrist[13-15]. Vegelin *et al.*[13] (2003) analyzed anthropometrists of different skill levels and verified that the best scores for stature and triceps skin folds measurements were reached by anthropometrists with over than six years of experience. Thus, training and the performance of periodic controls of the measurement technique quality will allow anthropometrists to reach better accuracy and hence acceptable reliability in the measurement determinations.

In the case of the stature, a better accuracy is very important, above all in recent studies of human growth modeling[16] as part of the daily variations follow-up. Only with a full control of the technical error one can distinguish change on stature due to genuine growth from change resulting from the measurement error.

Once the TEM is periodically calculated, it is possible to quantify intra and inter-evaluator variations. In the present study, the attainment of the intra-evaluator TEM was used to indicate the necessity of the technique improvement of three trainee anthropometrists.

The TEMs obtained in the comparison between these measurements were classified as acceptable in the intra-evaluator analysis for most skin folds (chart 1). In other words, the variation verified between measurements performed by these anthropometrists in two different moments suffers no influence (or a little) of the systematic error. When these anthropometrists performed measurements before and after a training period for skin folds with acceptable TEM, the variations observed will be a result of adaptations to the training applied if all other factors that affect the measures were controlled (nutrition, diseases).

In the case of TEM results classified as **non-acceptable**, these anthropometrists should be encouraged to participate in technical improvements based on detailed anthropometrical protocols of measurement standardization and later to perform new TEM calculations.

In the inter-evaluator analysis, one observes that, despite anthropometrists have reached TEMs classified as acceptable for all skin folds (chart 2), the relative TEM value are closer to the cut point to be acceptable (10%) in regions of higher fat accumulation (supra-illiac, thigh and abdomen), especially when the comparison involves anthropometrists 2 and 3. The linear growth of the measurement error with the increase on the skin fold has already been described by Marks *et al.*[6] (1989).

For improvement purpose, it is recommended that further efforts should be done in order to obtain lower TEMs.

CONCLUSIONS

The method presented was of easy execution and allowed analyzing the performance of trainee anthropometrists.

The performance of TEM periodic evaluations with the objective of controlling and minimizing the anthropometrists' intra and inter-evaluator TEM is recommended.

The detailed presentation of the TEM calculation method will allow other groups to apply it, thus assuring a better quality control of the measurement performed.

ACKNOWLEDGMENTS

FAPERJ, FUJB and UFRJ.

All the authors declared there is not any potential conflict of interests regarding this article.

REFERENCES

1. Pozo J, Argente J. Técnicas auxológicas. An Esp Pediatr 2000;52:192-8.
2. Ulijaszek SJ, Kerr DA. Anthropometrics measurement error and the assessment of nutritional status. Br J Nutr 1999;82:165-77.
3. Hawes MR, Martin AD. Human body composition. In: Eston R, Reilly T, editors. Kinanthropometry and exercise physiology laboratory manual: tests, procedures and data. New York: Routledge Taylor & Francis Group, 2004;5-43.
4. Pederson D, Gore C. Error en la medición antropométrica. In: Norton K, Olds T, editors. Antropométrica. Argentina: Biosystem Servicio Educativo, 2000;71-86.
5. Norton K, Olds T, editors. Antropometrica. Argentina: Biosystem, 2000.
6. Marks GC, Habicht JP, Mueller WH. Reliability, dependability, and precision of anthropometric measurements – The Second National Health and Nutrition Examination Survey 1976-1980. Am J Epidemiol 1989;130:578-87.
7. Oliveira FP, Guimarães JNF. Antropometria. Rio de Janeiro: NCE-UFRJ, 2003.
8. Ward R, Anderson GS. Resilience of antropometric data assembly strategies to imposed error. J Sports Sci 1998;16:755-9.
9. Harrison GG, Buskirk ER, Carter LJE, Johston FE, Lohman TG, Pollock ML, et al. Skin folds thickness and measurement technique. In: Lohman TG, Roche AF, Martorell R, editors. Anthropometric standardization reference manual. Illinois: Human Kinetics, 1988;148-55.
10. Whitehead JR. A study of measurement variation among different skinfold calipers. Br J Phys Educ 1990;7:10-4.
11. Martin AD, Ross WD, Drinkwater DT, Clarys JP. Prediction of body fat by skinfold caliper: assumptions and cadaver evidence. Int J Obes 1985;9:31-9.
12. Clarlyon R, Gore C, Woolford S, Bryant R. Calibración de los calibres de pliegues cutáneos Harpenden. In: Norton K, Olds T, editors. Antropometrica. Argentina: Biosystem, 2000;87-106.
13. Vegelin AL, Brukx LJ, Waelkens JJ, Van den Broeck J. Influence of knowledge, training and experience of observers on the reliability of anthropometrics measurements in children. Ann Hum Biol 2003;30:65-79.
14. Kouchi M, Mochimaru M, Tsuzuki K, Yokoi T. Interobserver errors in anthropometry. J Hum Ergol (Tokyo) 1999;28:15-24.
15. Caíno S, Adamo P, Kelmansky D, Lejarraga H. Impacto del entrenamiento sobre el error de mediciones antropométricas. Arch Argent Pediatr 2002;100:110-3.
16. Lampl M, Veldhuis JD, Johnson ML. Saltation and stasis: a model of human growth. Science 1992;258:801-3.