

# Como medir a confiabilidade de dobras cutâneas

## *Measuring the reliability of skinfolds*

### **Rosely Sichieri**

Instituto de Medicina Social /UERJ

Rua São Francisco Xavier, 524, 7º andar

Maracanã

20551-030 Rio de Janeiro, RJ

E-mail: SICHIERI@UERJ.BR

### **Vânia de Mattos Fonseca**

Instituto de Medicina Social e Instituto Fernandes Figueira

### **Claudia de Souza Lopes**

Instituto de Medicina Social, UERJ

### **Resumo**

Um dos métodos mais usados para a obtenção de informação relacionada à quantidade de gordura e sua distribuição no organismo é a medida de dobras cutâneas. O treinamento dos antropometristas para realização de medidas de dobras cutâneas representa um grande esforço na pesquisa nutricional e poucos estudos desenvolveram métodos práticos para a padronização destas medidas. O presente estudo compara os resultados de um estudo de confiabilidade utilizando os métodos previamente referidos e o coeficiente de correlação intra-classe (ICC), baseado no modelo ANOVA. Os resultados indicam que o ICC é o melhor método para a padronização dos antropometristas.

**Palavras-chave:** Confiabilidade da tecnologia. Dobras cutâneas. Coeficiente de correlação intra-classe.

### **Abstract**

One of the most widely used and inexpensive methods for obtaining information concerning the amount and distribution of human subcutaneous fat is the measurement of skinfolds. The skinfold measurement technique requires long-term supervised practice and repetition to obtain reliable results and there are few practical methods for standardizing measures taken by anthropometrists. Two studies, developed methodology to standardize skinfold measurements. We compared data of a reliability study, using the two methods previously discussed, with intraclass correlation coefficient calculations based on an ANOVA mixed model in order to assess two technicians measuring three skinfolds. The technicians tested were considered reliable only when using intraclass correlation coefficients. The intraclass correlation coefficient appears to be a better choice for the standardization of skinfold measurements, since it is not influenced by measurements with greater discrepancies such as those for obese individuals.

**Keywords:** Technology reliability. Skinfold. Intraclass correlation coefficient.

## Introdução

Medir pregas cutâneas é um importante método de avaliação da quantidade e da distribuição da gordura corpórea em estudos epidemiológicos, por ser um método barato e não invasivo. Sua realização requer, contudo, treinamento prolongado e supervisão para que se obtenham resultados confiáveis<sup>1</sup>, fazendo com que uma importante limitação na utilização das medidas de prega cutânea seja a dificuldade na padronização dos antropometristas. A realização de medidas confiáveis depende de um bom treinamento dos antropometristas, uma vez que o instrumento de coleta parece desempenhar um papel muito pequeno na determinação do erro. Leger e col.<sup>2</sup> (1982) compararam medidores de plástico a outros mais sofisticados e, mais comumente utilizados, e não observaram importante discrepância nas medidas.

Embora os erros de medição de pregas cutâneas não possam ser completamente eliminados, existem métodos que permitem avaliá-los, sendo possível incorporar o conhecimento dos erros na análise dos dados. Uma discussão pormenorizada dos erros em medidas antropométricas é encontrada em Lohman e col.<sup>3</sup> (1988).

Duas são as questões a serem enfrentadas quando se deseja aferir medidas antropométricas capazes de retratar o estado nutricional de um indivíduo: a da validade e a da confiabilidade do método utilizado. A validade de um instrumento de medida pode ser definida como a extensão com que as diferenças de resultados obtidos com tal instrumento refletem diferenças reais entre indivíduos<sup>4</sup>. A validade das medidas de dobras cutâneas é aceita com base nas altas correlações encontradas com a medida de gordura corporal feita através de pesagem hidrostática, considerada padrão-ouro<sup>5</sup>.

A confiabilidade por sua vez é definida como a extensão na qual as medidas (ou atribuições de valores ou categorias), obtidas com determinado teste ou instrumento ou ainda por diferentes avaliadores, são

reproduzíveis ou repetíveis<sup>6</sup>. Assim, ainda que um método seja válido, testes de confiabilidade do seu uso são necessários<sup>7</sup>.

O treinamento dos antropometristas para realização de medidas de dobras cutâneas representa um grande esforço na pesquisa nutricional, sendo que poucos estudos desenvolveram métodos práticos para a padronização destas medidas. Dois estudos claramente relacionados a este ponto são os de Habicht<sup>8</sup> (1974) e o de Sauerborn e col.<sup>9</sup> (1991), descritos na metodologia deste trabalho. Contudo estes dois métodos, na realidade, não indicam o grau de concordância das medidas, mas simplesmente testam hipóteses sobre o erro de tipo I, quando o real problema nestas avaliações diz respeito a quão próximas são as medidas.

Este estudo tem como objetivo comparar os métodos propostos por Habicht (1974)<sup>8</sup> e Sauerborn et al. (1991)<sup>9</sup> com uma proposta de padronização de antropometristas, usando o coeficiente de correlação intra-classe (ICC).

## Material e Método

Foram avaliados 22 adolescentes (11 meninos e 11 meninas) de uma escola privada do RJ, com idade entre 11 e 13 anos. Este é o tamanho de amostra necessário para medir um valor de ICC ou Kappa igual a 0,70, com um nível de significância de 0,05 e um poder de 0,95<sup>10</sup>. Foram realizadas, em cada estudante, três medidas de cada uma das dobras cutâneas (tríceps, subescapular e supra-ílica). Dois antropometristas tomaram as medidas. Um deles foi uma das autoras deste trabalho (VMF) e o outro, o treinador desta, com grande experiência neste tipo de atividade, foi considerado como padrão ouro. A antropometrista (VMF) foi treinada por um período de 16 horas.

As dobras cutâneas foram obtidas utilizando-se um plicômetro da marca Cescorf, com pressão constante, tomadas segundo padronização<sup>11</sup>. A dobra tricipital foi medida na parte posterior do braço direito, so-

bre o músculo tricipital, no ponto médio entre o acrômio e o olécrano, pinçando-se a pele e o tecido subcutâneo entre o polegar e o indicador, onde se aplicou o plicômetro 1 cm abaixo dos dedos que pinçavam a prega, sendo a leitura feita após 2 a 3 segundos no milímetro mais próximo. A dobra subescapular foi medida logo abaixo da extremidade inferior da escápula direita. A pele e o tecido subcutâneo foram pinçados neste local e a dobra angulada em 45° a partir do plano horizontal, dirigindo-se superiormente para dentro, onde se colocou o plicômetro 1 cm abaixo dos dedos que pinçavam a dobra. Para a medida da dobra supra-iliaca, afastava-se levemente o membro superior direito do corpo e pinçava-se a dobra logo acima da crista ilíaca no sentido vertical, segundo a linha média axilar. Estas medidas foram tomadas na ordem: tríceps, subescapular e supra-iliaca, por três vezes. Em cada medição era reiniciado o procedimento de colocação do plicômetro. As medidas foram feitas com o mesmo instrumento; um antropometrista desconhecia as medidas obtidas pelo outro e um anotador as registrava. A calibração do aparelho foi realizada no início das mensurações de cada antropometrista.

### **Análise de variância**

Classicamente, uma análise de variância avalia a significância estatística das diferenças entre diferentes médias. Entretanto, ela também provê uma expressão para a variância intra conjunto de medidas tão bem quanto para a variância entre conjunto de medidas. Estes são os componentes da variância usados para compor as medidas do ICC.

Utilizamos um modelo misto de análise de variância conforme Landis e Koch<sup>12</sup> (1975) e Bartko<sup>13</sup> (1994), o qual contempla a repetição das medidas para um mesmo sujeito com o antropometrista considerado efeito fixo, e a repetição das medidas um efeito randômico, e cujo modelo é descrito como:

$$y_{ijk} = m + p_i + r_j + (pr)_{ij} + e_{ijk}$$

onde:

$y$  = variável dependente, medida de interesse,

$m$  = efeito geral comum a todas as observações,

$p_i$  = são os valores aleatórios de cada um dos sujeitos analisados,

$r_j$  = é o efeito fixo associado ao antropometrista,

$(pr)_{ij}$  = é o termo de interação,

$e_{ijk}$  = erro residual para sujeito "i", observador "j" e a medida "k".

O termo de interação  $(pr)_{ij}$  mede a extensão em que o antropometrista se afasta do seu padrão usual quando mede o sujeito "i"; a variância desta interação reflete quanto os observadores variam na sua classificação geral de um mesmo sujeito. Assume-se a independência dos efeitos.

O ICC nesta situação é estimado pela razão:

$$\frac{MSB - MSE}{MSB + MSE}$$

$$MSB + MSE$$

onde:

MSB é o estimador da variância devido a diferenças nas medidas dos sujeitos para os dois antropometristas e indica o efeito entre-sujeitos.

MSE é o estimador do componente da variância relativo ao erro intrínseco da medida.

Valores do ICC acima de 0.70 são considerados como apresentando substancial confiabilidade<sup>12</sup>. Os valores de variância foram obtidos através do programa Statistical Analysis System (SAS), pela inclusão de um termo de interação no modelo (antropometrista \* repetição das medidas).

### **Padronização proposta por Habicht**

A metodologia permite uma identificação rápida, em condições de campo, da confiabilidade entre e intra antropometrista. Este autor propõe que cada técnico repita a medida duas vezes para dez observações diferentes, sendo que a soma dos quadrados das diferenças para o mesmo

antropometrista define a confiabilidade intra individual (o autor chamou esta aferição de precisão), enquanto que a soma dos quadrados das diferenças entre dois antropometristas para a mesma observação, define a confiabilidade entre indivíduos (o autor chamou esta aferição de exatidão). Quando um antropometrista atinge uma confiabilidade intra individual menor do que duas vezes a confiabilidade intra individual do supervisor ele é considerado preciso; quando ele apresenta uma confiabilidade entre indivíduos menor do que três vezes a confiabilidade intra individual do supervisor, o treinando é considerado padronizado.

Neste estudo, foram utilizadas as duas últimas medidas efetuadas pelos antropometristas, sendo elas avaliadas conforme segue:

- 1) cada avaliador faz 2 medidas de cada adolescente, sendo calculadas as diferenças (d) de cada par destas medidas. Estas diferenças são elevadas ao quadrado e é feito seu somatório (d<sup>2</sup>);
- 2) cada par de medidas é somado, sendo calculadas as diferenças das somas entre o avaliador em treinamento e o supervisor (padrão-ouro). As diferenças são elevadas ao quadrado e seu somatório calculado (D<sup>2</sup>).

Os resultados são analisados dentro dos seguintes critérios:

- 1) o somatório de d<sup>2</sup> do supervisor será o menor, supondo-se que ele seja o mais competentes dos antropometristas;
- 2) o somatório do avaliador em treinamento tem relação inversa com a precisão, e assim não deve exceder o dobro do somatório de d<sup>2</sup> do supervisor;
- 3) o somatório de D<sup>2</sup> tem relação inversa com a exatidão, não devendo exceder o triplo de somatório de d<sup>2</sup> do supervisor.

### **Padronização proposta por Sauerborn**

Este método para identificação de concordância baseia-se nas médias de duas medidas por antropometrista para cada observação.

São calculadas a média da diferença dos pares de medidas (d) e seu desvio padrão (DP). O intervalo de concordância inclui  $d \pm 2(DP)$ . Assim, calcularam-se a média de cada par de medidas e a diferença entre o treinando e o padrão-ouro, obtendo-se as médias das diferenças (d) e seus desvios padrões. Quanto mais próxima de zero estiver a média das diferenças, menor será a possibilidade de erro.

## **Resultados**

A Tabela 1 mostra os valores médios das pregas cutâneas e o desvio padrão obtidos pelos dois antropometristas. A normalidade da distribuição foi testada em SAS e os testes não levaram à rejeição da hipótese de normalidade. Os valores do ICC, para as três medidas efetuadas, foram todos acima de 0,90, considerados como indicadores de muito boa concordância entre as medidas (Tabela 2). A Tabela 3 apresenta os resultados referentes ao proposto por Habicht<sup>8</sup> (1974), mostrando que, para as três medidas, se excederia o limite aceitável para considerar o antropometrista como padronizado, ou seja, D do antropometrista 2 é maior do que três vezes o valor de d. Contudo, há reprodutibilidade para as pregas supra-ílica e tricípital.

Os resultados obtidos com a proposta de Sauerborn e col.<sup>9</sup> (1991) indicam que, tanto para um mesmo antropometrista, quanto para as medidas entre antropometristas (Tabela 4), os valores obtidos apresentam ampla variação, tornando-se difícil aceitar a confiabilidade das medidas. Os valores médios das diferenças apresentadas na Tabela 4 referem-se à diferença entre a média de duas medidas por antropometrista.

## **Discussão**

Os resultados deste estudo confirmam os achados de outros autores sobre a dificuldade de realização das medidas de dobras cutâneas<sup>14</sup>. Dos resultados apresentados e dada a facilidade, nos dias atuais, de utilização de programas que permitam a realização de análise de variância (ANOVA),

**Tabela 1** - Valores médios e desvio-padrão (DP) das 3 pregas cutâneas, segundo antropometrista.  
**Table 1** - Means and standard deviation (DP) for three skinfolds, by technicians.

	Antropometristas			
	Em padronização		Padrão	
	média	DP	média	DP
Supra-ílica	17,9	10,3	20,3	9,2
Tríceps	18,2	7,8	16,1	6,2
Subescapular	14,7	7,5	15,0	6,5

**Tabela 2** - Quadrados médios, coeficiente de correlação intraclasse (ICC) e intervalo de confiança de 95% (IC) para cada dobra cutânea.

**Table 2** - Variance, intraclass correlation coefficient (ICC), and 95% confidence interval (IC) for each skinfold.

Quadrados médios	Dobra Supra-ílica	Dobra Tríceps	Dobra Subescapular
$\sigma_d^2$ entre antropometristas	18.697	14.281	191
$\sigma_s^2$ entre sujeitos	55.774	26.716	28.386
$\sigma_{sd}^2$ interação	4431	3086	204
$\sigma_e^2$ erro	482	694	494
ICC	0.97	0.94	0.96
IC	0.95 – 0.99	0.85 – 0.97	0.89 – 0.98

$\sigma_s^2$  = é o estimador da variância devido a diferenças nas medidas dos sujeitos para os 2 antropometristas. Indica o efeito entre-sujeitos calculado pela diferença entre o quadrado médio entre os sujeitos menos o quadrado médio do resíduo, dividido pelo número de antropometristas.

$\sigma_d^2$  = é o estimador da variância devido a diferenças das medidas médias dos antropometristas; reflete o efeito entre os observadores.

$\sigma_{sd}^2$  = é o estimador da variância dos observadores nos julgamentos de um mesmo sujeito; indica a variância da interação antropometrista-sujeito.

$\sigma_e^2$  = é o estimador do componente da variância relativo ao erro intrínseco da medida.

95% IC calculado segundo Bartko<sup>13</sup>, 1994.

$(F_L - 1) / (F_L + 1) < ICC < Fu - 1 / Fu + 1$ .

Onde:  $F_L = F / F_{(1-\alpha/2); N-1, N-1}$ ;  $Fu = F_{(1-\alpha/2); N-1, N-1}$ .

$\sigma_s^2$  = Between subjects variance. Calculated by the difference between subjects mean square and residual mean square divided by the number of technicians.

$\sigma_d^2$  = Between raters variance.

$\sigma_{sd}^2$  = Raters variance measuring the same subject. Indicates the interaction raters-subjects.

$\sigma_e^2$  = error associated with observations.

95% CI calculated according to Bartko, 1994.

ICC- Intraclass correlation coefficient

$(F_L - 1) / (F_L + 1) < ICC < Fu - 1 / Fu + 1$ .

Where:  $F_L = F / F_{(1-\alpha/2); N-1, N-1}$ ;  $Fu = F_{(1-\alpha/2); N-1, N-1}$ .

**Tabela 3** - Cálculos segundo Habicht<sup>8</sup>, valores de d (precisão) e D (exatidão)

**Table 3** – Values of precision (d) and reliability (D) according to Habicht<sup>8</sup>

		d ( $\Sigma d^2$ )	D ( $\Sigma D^2$ )
Supra-ílica	Antropometrista 1 (padrão)	5.991	
	Antropometrista 2	10.234	203.365
Tricipital	Antropometrista 1 (padrão)	5.724	
	Antropometrista 2	5.944	231.950
Subescapular	Antropometrista 1 (padrão)	3.383	
	Antropometrista 2	19.230	89.083

**Tabela 4** - Cálculos segundo Sauerborn e col.<sup>9</sup> para média das diferenças intra e entre antropometristas (d) e desvio padrão (DP).

**Table 4** - Differences within and between (d) skinfold measures of the two technicians and standard deviation (DP), according to Sauerborn et al.<sup>9</sup>

		Diferenças intra antropometristas	
		Diferenças	DP
Supra-iliaca	Antropometrista 1 (padrão)	-4,27	15,8
	Antropometrista 2	5,68	21,6
Tricipital	Antropometrista 1 (padrão)	0,45	16,8
	Antropometrista 2	-6,72	15,0
Subescapular	Antropometrista 1 (padrão)	-2,92	12,3
	Antropometrista 2	4,00	29,8

  

Diferenças entre antropometristas		
	Diferenças (mm) d	DP(d) (mm)
Supra-iliaca	-23,8	39
Tricipital	20,8	48
Subescapular	-2,4	31

parece mais adequado avaliar a concordância de antropometristas através do ICC.

Comparando os valores obtidos através da análise de variância com aqueles obtidos pelos dois outros métodos, conclui-se que, enquanto a análise de variância mostrou resultados para o ICC maiores do que 70%, valor considerado de bom nível de concordância<sup>6</sup>, os valores de concordância/exatidão, segundo Habicht, excederiam o limite aceitável que seria de 3 vezes o valor de d (precisão) a partir do antropometrista padrão. Conclusão semelhante se atingiria com a proposta de Sauerborn e col.<sup>9</sup>, cujos intervalos de confiança excederiam em muito os valores do percentil 95 da distribuição das pregas, em adolescente americanos (WHO 1995).

Observe-se que tanto na proposta de Habicht<sup>8</sup>, quanto na de Sauerborn e col.<sup>9</sup>

uma ou poucas medidas muito discrepantes têm uma influência muito grande nas medidas de concordância. Este é um problema particularmente importante na aferição de pregas cutâneas, pois as maiores discrepâncias ocorrem para as pessoas mais gordas, que vão influenciar sobremaneira os valores de “d” avaliados. Este problema não tem tanta importância para o cálculo do ICC, pois sendo uma razão, os obesos vão contribuir tanto na variância entre as medidas, quanto na variância entre os indivíduos medidos. Daí a menor influência de valores aberrantes sobre o ICC.

Os resultados deste estudo de confiabilidade nos permitem concluir que, embora o cálculo do ICC seja mais complexo do que as medidas alternativas avaliadas, ele é a medida mais desejável para o cálculo de confiabilidade.

---

## Summary

One of the most widely used and inexpensive methods for obtaining information concerning the amount and distribution of human subcutaneous fat is the measurement of skinfolds. The skinfold measurement technique requires long-term extended supervised practice and repetition to obtain reliable results and there are few practical methods for standardizing anthropometrists. Two studies, Habicht, 1974 and Sauerborn et al, 1991, developed a method to standardize skinfold measurements. Neither method, however, indicated agreement between technicians. The authors tested a hypothesis of a type I error, when the real issue was determining how close measures were. We compared data of a reliability study, using two methods previously discussed with intraclass correlation

coefficient (ICC) calculations based on an ANOVA mixed model in order to compare two technicians measuring three skinfolds. For the three skinfolds tested, ICC values were greater than 0.90, considered good agreement, and indicating that the technician tested was reliable. The technicians tested were not considered reliable when both the Habicht and the Sauerborn methods were used. The high influence of a few large discrepancies between measures using the Habicht and Sauerborn methods may explain the difference found. This kind of discrepancy is particularly important when measuring skinfolds of obese individuals. The intraclass correlation coefficient appears to be a better choice for the standardization of skinfold measurements, since it is not influenced by measurements.

---

## Referências

1. Jelliffe DB. *The assessment of the nutritional status of the community*. Geneva; WHO; 1966. (WHO - Monograph Series, 53).
2. Léger AL, Lambert J, Martin P. Validity of skinfold caliper measurements. *Wayne State University Press* 1982; 54:667-75.
3. Lohman TG, Roche AF, Martorell R. *Anthropometric standartization reference manual*. Illinois: Human Kinetics Books; 1988.
4. Last JM. *A dictionary of epidemiology*. 3<sup>rd</sup>ed. New York: Oxford University Press; 1995.
5. Revicki ED, Israel RG. Relationship between body mass and measures of body adiposity. *Am J Public Health* 1986; 76:992-4.
6. Streiner DL, Norman GR. *Health measurement scales: a practical guide to their development and use*. Oxford: Oxford University Press; 1989.
7. Bartko JJ, Carpenter WT. On the methods and theory of reliability. *J Nerv Ment Dis* 1976; 163:307-17.
8. Habicht JP. Estandartización de métodos epidemiológicos cuantitativos sobre el terreno. *Bol Oficina Sanit Panam* 1974; 76: 375-84.
9. Sauerborn R, Morley DC, Bullough CHW. Un método estadístico simple para obtener confiabilidad en las mediciones antropométricas. *Salud Pública Méx* 1991; 33:106-11.
10. Dunn G. *Designs for reliability studies*. New York: Oxford University Press; 1989. Designs for reliability studies; p. 59-72.
11. Durnin JVGA, Rahaman VM. The assessment of the amount of fat in the human body from measurements of skinfold thickness. *Br J Nutr* 1967; 21: 681-8.
12. Landis JR, Koch GG. A review of statistical methods in analysis of data arising from observer reliability studies (part I). *Stat Neerland* 1975; 3: 103-23.
13. Bartko JJ. General methodology. II - Measures of agreement: a single procedure. *Stat Med* 1994; 13:737-45.
14. Flegal KM. Defining obesity in children and adolescents: epidemiologic approaches. *Crit Rev Food Nutr* 1993; 33:307-12.
15. WHO Working Group. Use and interpretation of anthropometric indicators of nutritional status. *Bull World Health Organ* 1995; 64:929-41.

## Anexo

DATA ANTRVAN;

INPUT ANTROPO 1 OBS 2-3 SUPR1 4-6 SUPR2 7-9 SUPR3 10-12 TRIC1 13-15 TRIC2 16-18

TRIC3 19-21 SUBS1 22-24 SUBS2 25-27 SUBS3 28-30 ; CARDS;

2 1140150120200210220175150160  
2 2165175180 80 85 75100105100  
1 2 78102112 79 85 75 92 94 94  
1 1100130105196189188125140125  
2 3215240245160240205135135130  
1 3155152145254257248175125124  
1 4125120108175179179105106120  
2 4145170165182195165170165165  
2 5178181165135158155110113136  
1 5143150144154145158145138118  
2 6280330280215220240210180190  
1 6229233243245238240143143173  
2 7210220192175175200165165160  
1 7135158188145172175122126129  
2 8255300290235240205175200205  
1 8186221215233205245155163161  
2 9130140140170170175150130140  
1 9 85 98 93189180186 94102 92  
210130130160150160150100120120  
110 89103100149146133 96 94 91  
211 75 75 75 60 65 65 80 75 70  
111100 70 70 70 70 70100100 90  
212410420405240250255350340370  
112440500500440440470380400470  
213261291290278278300185218195  
113220221223232240238212239224  
214 70 75 75 75 74 75105 95100  
114110100100150100100100120110  
215 70 85 85 70 80 75 70 70 70  
115100110 90120 95110 70 70 90  
216180240230130145140 85 95105  
116220210220140150170120160160  
217380370370175160155290270255  
117370380350210210220280290270  
218250235230120115115135130145  
118180210240150160200160150140  
219320300300170165180130135145  
119350320350195200200150150140  
120117120115116110116112112112  
220 90140150 65 65 65 95 95105  
221210210210200200220130160160  
121204174240209209220126220115  
222140140135165150155130135120  
122 90 79104138158145 99 93 99