

Variabilidade na aferição de medidas antropométricas: comparação de dois métodos estatísticos para avaliar a calibração de entrevistadores

Variability in the measurement of anthropometric measures: comparison between two statistical methods to assess interviewers' calibration

Vanilde de Castro¹

Suzana Alves de Moraes¹

Isabel Cristina Martins de Freitas¹

Lenise Mondini²

¹Escola de Enfermagem de Ribeirão Preto da Universidade de São Paulo - Núcleo de Epidemiologia, Ribeirão Preto, SP.

²Instituto de Saúde da Secretaria de Estado da Saúde de São Paulo

O presente trabalho está vinculado ao Projeto OBEDIARP, Projeto de Pesquisa financiado pelo CNPq-Edital CT-Saúde. Processo Nº 505622/2004-1.

Bolsa de Mestrado-Agência FAPESP. Processo Nº 06/51240-8. Concedida a Vanilde de Castro.

Contribuições dos Autores: **V Castro** colaborou na análise dos dados, na redação do manuscrito e na revisão bibliográfica do estudo. **SA Moraes** elaborou a proposta do estudo, colaborou na interpretação dos resultados, na revisão bibliográfica e na elaboração e redação da versão final do manuscrito. **ICM Freitas** colaborou no gerenciamento e análise dos dados, na revisão bibliográfica e na elaboração e redação da versão final do manuscrito. **L Mondini** colaborou no gerenciamento dos dados e na revisão da versão final do manuscrito.

Agradecimentos: Ao CNPq pelo financiamento na modalidade Auxílio à Pesquisa (Processo nº 505622/2004). À FAPESP pela concessão de 01 Bolsa de Mestrado para **V Castro** (Processo nº 06/51240-8).

Correspondência: Suzana Alves de Moraes. Av. Santa Luzia- no. 440-Apto. 81. CEP 14025-090 Ribeirão Preto, SP. E-mail: samoraes@usp.br

Resumo

Estudos que incluem medidas antropométricas exigem, além da padronização de técnicas de aferição, o emprego de métodos estatísticos para a avaliação de erros de mensuração. **Objetivo:** Comparar duas técnicas estatísticas para avaliar a calibração de entrevistadores, em treinamento para a aferição de medidas antropométricas. **Metodologia:** Treze entrevistadores participaram da fase de treinamento, que foi programado de modo que, em cada sessão, sub-grupos de entrevistadores realizassem duas aferições de medidas de peso e duas de altura em voluntários (em média, 10 voluntários/sessão). Ao todo, foram realizadas seis sessões para a aferição das medidas de peso e onze para a aferição das medidas de altura. Para as medidas de peso foram utilizadas balanças eletrônicas e, para as de altura, estadiômetros de parede. Para avaliar a calibração, dois métodos estatísticos foram comparados: a) o coeficiente de correlação intraclasse (CCI) e b) a precisão e a exatidão das aferições, segundo Habicht. **Resultados:** Os entrevistadores foram submetidos, em média, a duas sessões para a calibração das medidas de peso e a três sessões para a calibração das medidas de altura, sendo a precisão atingida antes da exatidão. Os valores dos respectivos CCIs indicaram medições fortemente concordantes desde a primeira sessão. **Conclusão:** O método de Habicht apresentou melhor desempenho que o CCI, pois, além do cálculo da precisão, indicou a magnitude da divergência das aferições realizadas pelos entrevistadores, em relação às do supervisor (exatidão).

Palavras-chave: Reprodutibilidade. Coeficiente de Correlação Intraclasse. Calibração.

Abstract

Standardized technical procedures as well as statistical methods can be used to calculate the magnitude of measurement errors in studies related to anthropometric measures. **Objective:** To compare two statistical approaches to evaluate interviewers' calibration during the training phase before data collection in an epidemiological study.

Methods: Thirteen interviewers, divided into sub-groups were trained to take two measures of weight and two of height from 10 volunteers in each training section. Training was completed in six sections for weight and eleven for height measurements. Digital scales were used to measure weight and wall stadiometers to measure height. In order to evaluate interviewers' calibration, two statistical methods were compared: a) the Intraclass Correlation Coefficient (ICC); and b) the precision and accuracy as proposed by Habicht. **Results:** On average, interviewers were submitted to 2 sections for weight measure calibration, and to 3 sections for height measure calibration, and precision was reached before accuracy. ICC values showed that measures were strongly correlated since the first section. **Conclusion:** The Habicht method seems to be better than ICC, given it allowed for corrections not only of interviewer discrepancies related to precision but it also calculated the magnitude of deviations between measurements of interviewers and supervisor (accuracy).

Keywords: Reliability. Intraclass Correlation Coefficient. Calibration.

Introdução

Em estudos epidemiológicos de base populacional, que envolvem medidas antropométricas em seu instrumento de coleta de dados, torna-se imperativo o treinamento intensivo da equipe de entrevistadores, antes do início do trabalho de campo¹.

Para a investigação da prevalência de excesso de peso, diferentes medidas antropométricas têm sido propostas, sendo o Índice de Massa Corporal (IMC) o melhor indicador de obesidade global, segundo a Organização Mundial de Saúde (OMS)².

As medidas de peso e altura, necessárias para o cálculo do IMC, estão sujeitas a variações que podem ser decorrentes de processos de mensuração inadequados, falhas mecânicas ou falta de calibração dos respectivos equipamentos, ou, ainda, decorrentes de variações aleatórias (devidas ao acaso)³. Para a minimização de erros sistemáticos, destaca-se, como de fundamental importância, a padronização dos procedimentos técnicos a serem adotados durante as aferições, bem como a calibração dos entrevistadores¹.

Dentro desta perspectiva, o presente estudo teve por objetivo comparar dois métodos utilizados para a calibração de uma equipe de entrevistadores, em fase de treinamento para aferição de medidas antropométricas, antes do desenvolvimento de um estudo epidemiológico de base populacional. A aplicação destes métodos esteve fundamentada no fato de que, embora amplamente utilizados por antropometristas, considerou-se pertinente avaliar as vantagens e limitações de cada um deles.

Metodologia

Treze entrevistadores participaram da fase de treinamento, que ocorreu nos meses de outubro e novembro de 2005. Ao todo, foram realizadas seis (06) sessões para a aferição de peso e onze (11) sessões para a aferição de altura. O treinamento foi programado de modo que, em cada sessão, sub-grupos de entrevistadores realizassem

aferições de medidas de peso e altura em voluntários (em média, 10 voluntários/sessão). As 06 sessões de peso e as 11 de altura totalizaram, respectivamente, 8 e 20 horas. Durante cada sessão de treinamento, os entrevistadores tomaram duas medidas de peso (em quilogramas) e de altura (em centímetros) de cada um dos voluntários. Para evitar que guardassem na memória a primeira aferição, os entrevistadores realizaram a segunda medida somente após a finalização da primeira aferição em todos os voluntários que participaram da respectiva sessão. A segunda medida foi anotada em um novo bloco de anotações, onde não constavam os registros da primeira aferição. Além das medidas dos entrevistadores, o peso e a altura, em todos os voluntários e, em cada uma das sessões, foram duplamente aferidos por um antropometrista (supervisor), com experiência comprovada para estas aferições, sendo, portanto, considerado o padrão-ouro.

Para a tomada das medidas de peso foram utilizadas balanças eletrônicas portáteis, da marca TANITA, modelo BF 680, com precisão de 100 gramas. Para a aferição das medidas de altura foram utilizados estadiômetros da marca SECA, com escala em décimos de centímetros. Os entrevistadores foram instruídos quanto à técnica a ser utilizada para a aferição das medidas de peso, segundo a qual os voluntários deveriam estar sem sapatos, trajando roupas leves, posicionados com os dois pés sobre a balança, distribuindo seu peso igualmente sobre as duas pernas e com o olhar no horizonte. Antes de cada aferição do peso, os entrevistadores taravam a balança, segundo recomendações do fabricante. Para a aferição da altura, os entrevistadores foram instruídos para que os voluntários estivessem também sem sapatos, posicionados com os pés unidos e contra a parede, para apoiar 05 pontos do corpo: calcanhar, panturrilha, glúteos, espáduas e a cabeça, que deveria ser posicionada respeitando-se o plano de Frankfurt. O estadiômetro foi fixado a 2,20 metros verticais, em relação ao piso, e calibrado com esquadro técnico de 60 graus. O esquadro foi também utilizado para encontrar o ângulo

reto entre o piso e a parede onde o estadiômetro deveria ser fixado^{2,4}.

Análise Estatística

Para avaliar a calibração dos entrevistadores, duas técnicas estatísticas foram utilizadas: 1) a precisão e a exatidão, propostas por Habicht¹ e 2) o coeficiente de correlação intraclasse, obtido por análise de variância, com 01 critério de classificação⁵. A precisão e a exatidão propostas por Habicht, na avaliação dos entrevistadores, comparam as medidas aferidas pelo entrevistador com as do supervisor. A precisão é alcançada quando:

$$\sum d^2 \leq 2 \sum d_s^2$$

onde:

d = diferença entre as aferições realizadas pelo mesmo entrevistador.

d_s = diferença entre as aferições realizadas pelo supervisor.

A exatidão é alcançada quando:

$$\sum D^2 \leq 3 \sum d_s^2$$

onde:

D = diferença entre as somas das aferições do entrevistador e do supervisor.

d_s = diferença entre as aferições realizadas pelo supervisor.

Desta forma, depreende-se que os entrevistadores seriam aprovados quanto à precisão, quando sua respectiva somatória de diferenças não ultrapassasse o dobro da somatória de diferenças do supervisor, e, quanto à exatidão, quando a somatória das diferenças entre as medidas do entrevistador, comparadas com as respectivas medidas do supervisor, não ultrapassasse o triplo da somatória de valores alcançados pelo supervisor.

Para a obtenção das medidas de precisão e exatidão, utilizou-se uma planilha de cálculos formatada no software EXCEL⁶ e desenvolvida pelo Laboratório de Antropometria de Populações (LANPOP) do Departamento de Nutrição da Faculdade de Saúde Pública da Universidade de São Paulo.

O coeficiente de correlação intraclasse (CCI), obtido por análise de variância com

um critério de classificação, é uma função de variâncias, sendo a variância total definida pela soma dos quadrados dos desvios de todas as medidas, em relação à sua média global. Esta variância pode ser decomposta em duas frações: a variância devida às diferenças “intra indivíduos” aferidos por diferentes entrevistadores, também conhecida como a variância do resíduo (SQR) e a variância “entre indivíduos aferidos” (SQE)⁵. Desta forma, o CCI é definido por:

$$CCI = \frac{QME - QMR}{QME + QMR}$$

onde:

$$QME = SQE / (k-1)$$

$$QMR = SQR / (n-k)$$

SQE = soma dos quadrados dos desvios entre a média das medidas dos entrevistadores para um mesmo voluntário e a média geral ;

SQR = soma dos quadrados dos desvios entre as medidas de cada entrevistador e a média das medidas para o respectivo voluntário;

(k-1) e (n-k) representam os respectivos graus de liberdade de SQE e SQR.

Portanto, depreende-se que, quando SQR tende a zero, o CCI tende a 1 (100%), denotando máximo acordo “intra-medidas” tomadas de um mesmo voluntário por diferentes entrevistadores.

Para a obtenção dos coeficientes de correlação intraclasse utilizou-se o software SPSS para Windows, versão 11.0⁷.

O estudo epidemiológico que deu origem aos dados do presente estudo recebeu financiamento do CNPq, e a coleta de dados foi finalizada em julho de 2007. O Protocolo original onde constam todas as etapas para o desenvolvimento do estudo foi aprovado pelo Comitê de Ética da Escola de Enfermagem de Ribeirão Preto-USP.

Os autores, através de documentação assinada e enviada a RBE, declaram não existir nenhum tipo de conflito de interesses.

Resultados

No presente estudo, os entrevistadores foram submetidos, em média, a 2 sessões para a calibração das medidas de peso e a

3 sessões para a calibração das medidas de altura, destacando-se que a precisão das medidas foi atingida antes da exatidão.

Na Tabela 1 encontram-se as médias, desvios padrão e coeficientes de variação para as medidas de peso e altura dos voluntários, nas respectivas sessões de treinamento. Os resultados indicam que o coeficiente de variação (CV) para as medidas de peso oscilou entre 16,55% e 22,04%, enquanto que para as medidas de altura o CV oscilou entre 2,36% e 7,73%.

Na Tabela 2 observam-se os resultados correspondentes à ordem das sessões às quais compareceram diferentes grupos de entrevistadores para a aferição do peso e da altura, bem como o respectivo número de entrevistadores por sessão. Na mesma tabela encontra-se também o total de medidas aferidas e os respectivos coeficientes de correlação intraclasse para as medidas dos entrevistadores, em cada sessão, excluindo-se as medidas do supervisor.

Na Tabela 3 observa-se a magnitude dos valores alcançados para a precisão e a exatidão das medidas dos entrevistadores, em relação às do supervisor para a variável “peso”.

As avaliações indicaram que apenas 2 entrevistadores foram aprovados em uma única sessão, sendo necessária para os demais, uma sessão adicional para sua aprovação. Verifica-se ainda, na mesma tabela, que a magnitude dos valores obtidos para o CCI é compatível com os valores de máxima concordância (CCI=1).

Com relação à variável altura (Tabela 4), os resultados indicam que grande parte dos entrevistadores atingiu a precisão, a partir da segunda sessão, sendo que houve maior dificuldade para atingirem a exatidão. Ao contrário dos resultados obtidos através das técnicas propostas por Habicht, os valores correspondentes ao coeficiente de correlação intraclasse indicaram que os entrevistadores realizaram medições altamente concordantes, desde a primeira sessão.

Discussão

Os resultados obtidos confirmam a ne-

Tabela 1 - Número de ordem das sessões para medidas de peso e altura, médias*, desvios-padrão** e coeficientes de variação*** para as medidas dos voluntários nas respectivas sessões.

Table 1 - *Weight and height section numbers, means*, standard deviations**, and coefficients of variation *** for measurements of volunteers in each section.*

Nº de ordem da sessão	\bar{x}	dp	CV (%)
Peso 1	76,285	16,577	21,730
Peso 2	70,281	15,489	22,039
Peso 3	76,625	14,815	19,334
Peso 4	72,493	11,999	16,552
Peso 5	77,747	16,745	21,538
Peso 6	73,313	14,650	19,983
Altura 1	166,674	11,753	7,052
Altura 2	160,198	3,777	2,358
Altura 3	164,277	9,107	5,544
Altura 4	164,465	12,562	7,638
Altura 5	167,325	12,430	7,429
Altura 6	162,270	6,461	3,982
Altura 7	162,191	4,718	2,909
Altura 8	170,081	9,380	5,515
Altura 9	168,638	10,253	6,080
Altura 10	162,870	6,839	4,199
Altura 11	167,517	12,944	7,727

* \bar{x} = média das medidas aferidas nos voluntários. **DP = desvios-padrão.***CV = coeficiente de variação / * \bar{x} = weight and height means. **DP - standard deviation ***CV- coefficients of variation.

Tabela 2 – Número de ordem das sessões de peso e altura, número de entrevistadores, número de medidas e respectivos coeficientes de correlação intraclasse (CCI)*, com intervalos de confiança (IC: 95%)

Table 2 – *Weight and height section numbers, number of interviewers and measurements, intraclass correlation coefficients (ICC)* and confidence intervals (CI: 95%)*

Nº de ordem das Sessões	Entrevistadores (n)	Medidas (n)	CCI	(IC: 95%)
Peso 1	2	40	1,000	(1,000 - 1,000)
Peso 2	4	80	1,000	(1,000 - 1,000)
Peso 3	4	80	1,000	(1,000 - 1,000)
Peso 4	2	40	1,000	(1,000 - 1,000)
Peso 5	2	36	1,000	(1,000 - 1,000)
Peso 6	11	308	1,000	(1,000 - 1,000)
Altura 1	4	64	1,000	(1,000 - 1,000)
Altura 2	3	48	0,999	(0,998 - 1,000)
Altura 3	3	48	1,000	(1,000 - 1,000)
Altura 4	2	20	0,999	(0,997 - 1,000)
Altura 5	2	20	1,000	(1,000 - 1,000)
Altura 6	2	20	0,999	(0,998 - 1,000)
Altura 7	4	32	1,000	(1,000 - 1,000)
Altura 8	4	32	1,000	(1,000 - 1,000)
Altura 9	2	16	1,000	(1,000 - 1,000)
Altura 10	5	80	1,000	(1,000 - 1,000)
Altura 11	8	96	1,000	(1,000 - 1,000)

*CCI = Coeficiente de Correlação Intraclasse para as medidas obtidas pelos entrevistadores, excluindo-se as aferições realizadas pelo supervisor.

*ICC = Intraclass Correlation Coefficient for interviewers measurements, except for the supervisor's measurements.

Tabela 3 -Precisão (P), exatidão (E), resultados da avaliação segundo Habicht*, coeficientes de correlação intraclasse (CCI) e respectivos intervalos de confiança (IC:95%) para as medidas de peso.

Table 3 – Precision (P), accuracy (E), interviewers evaluation as proposed by Habicht*, intraclass correlation coefficients (ICC), and confidence intervals (CI: 95%) for weight measurements.

Entrevistadores	Nº de Sessões	Desempenho 1			Desempenho 2		
		P	E	CCI (IC: 95%)	P	E	CCI (IC: 95%)
B	1	1,20 (A)	2,20 (A)	1,000 (1,000-1,000)	-----	-----	-----
C	1	1,50 (A)	2,50 (A)	1,000 (1,000-1,000)	-----	-----	-----
A	2	1,75 (A)	4,75 (R)	1,000 (1,000-1,000)	1,0 (A)	2,37 (A)	1,000 (1,000-1,000)
D	2	3,50 (R)	6,50 (R)	1,000 (1,000-1,000)	1,12 (A)	1,25 (A)	1,000 (1,000-1,000)
E	2	2,75 (R)	7,75 (R)	1,000 (1,000-1,000)	1,13 (A)	1,13 (A)	1,000 (1,000-1,000)
F	2	2,50 (R)	9,50 (R)	1,000 (1,000-1,000)	1,00 (A)	1,88 (A)	1,000 (1,000-1,000)
G	2	1,50 (A)	4,00 (R)	1,000 (1,000-1,000)	1,00 (A)	1,75 (A)	1,000 (1,000-1,000)
H	2	0,67 (A)	5,00 (R)	1,000 (1,000-1,000)	1,63 (A)	1,00 (A)	1,000 (1,000-1,000)
I	2	2,50 (R)	3,00 (A)	1,000 (1,000-1,000)	1,13 (A)	1,25 (A)	1,000 (1,000-1,000)
J	2	1,00 (A)	6,67 (R)	1,000 (1,000-1,000)	1,50 (A)	1,00 (A)	1,000 (1,000-1,000)
K	2	4,00 (R)	8,00 (R)	1,000 (1,000-1,000)	1,38 (A)	1,50 (A)	1,000 (1,000-1,000)
L	2	2,00 (A)	7,00 (R)	1,000 (1,000-1,000)	1,00 (A)	2,71 (A)	1,000 (1,000-1,000)
M	2	3,25 (R)	6,75 (R)	1,000 (1,000-1,000)	0,86 (A)	3,00 (A)	1,000 (1,000-1,000)

*A = Aprovado; R – Reprovado/*A = Approved; R - Failed

cessidade de treinamento para a aferição adequada de medidas antropométricas, antes do início do trabalho de campo.

A variabilidade de medidas antropométricas pode ser decorrente de variação biológica ou decorrente de variações técnicas. Tendo em vista que a maior incidência de erros é decorrente de variações técnicas e que as mesmas são passíveis de controle, a calibração dos entrevistadores é fundamental para a garantia de acurácia das medidas, promovendo-se, desta forma, a minimização de erros sistemáticos⁹⁻¹¹.

No presente estudo, os autores compararam duas técnicas estatísticas para a avaliação da acurácia de medidas antropométricas aferidas pelos entrevistadores, em fase de treinamento para o trabalho de campo de um estudo epidemiológico.

Avaliando-se os resultados obtidos por meio da aplicação do coeficiente de correlação intraclasse (CCI) com aqueles obtidos para as estimativas de precisão e exatidão, propostas por Habicht¹, observa-se que, enquanto o CCI indicou elevado nível de concordância para as aferições, já

na primeira sessão, os valores obtidos para a precisão e exatidão indicaram a necessidade de sessões adicionais de treinamento para que os entrevistadores atingissem a calibração desejada. A discrepância encontrada pode ser decorrente do fato de o CCI ser altamente dependente da variabilidade das respectivas medidas na população do estudo, sendo, portanto, afetado por sua heterogeneidade¹¹⁻¹⁵. Isto significa que elevados valores do CCI podem ser resultantes da grande variabilidade “entre” os sujeitos aferidos, mesmo em situações onde exista baixa concordância “intra” medidas^{12,13,15-18}. Esta situação decorre do fato de que, diante da heterogeneidade da amostra, a proporção da variância total determinada pela fração da variância devida às diferenças “entre indivíduos aferidos” aumenta, acarretando, conseqüentemente, a diminuição da proporção de variância total que é determinada pela fração da variância “intra-indivíduos”, o que poderia conduzir a uma elevação da magnitude do CCI, mesmo diante de baixa reprodutibilidade das medidas¹³. Por outro lado, cabe destacar que, quando a popula-

Tabela 4 – Precisão (P), exatidão (E), resultados da avaliação segundo Habicht*, coeficientes de correlação intraclasse (CCI), e respectivos intervalos de confiança (IC: 95%) para as medidas de altura.

Table 4 - Precision (P), accuracy (E), interviewers evaluation as proposed by Habicht*, intraclass correlation coefficients (ICC), and confidence intervals (CI:95%) for height measurements

Entrevistadores	Nº de Sessões	Desempenho 1			Desempenho 2			Desempenho 3			Desempenho 4		
		P	E	CCI (IC: 95%)	P	E	CCI (IC: 95%)	P	E	CCI (IC: 95%)	P	E	CCI (IC: 95%)
I	2	24,14 (R)	133,71 (R)	0,998 (0,992-1,000)	1,00 (A)	1,62 (A)	1,000 (1,000-1,000)	-----	-----	-----	-----	-----	-----
M	2	0,73 (A)	6,40 (R)	1,000 (1,000-1,000)	0,69 (A)	3,00 (A)	1,000 (1,000-1,000)	-----	-----	-----	-----	-----	-----
A	3	4,23 (R)	22,77 (R)	0,999 (0,997-1,000)	3,86 (R)	94,86 (R)	0,999 (0,995-1,000)	0,87 (A)	2,37 (A)	1,000 (1,000-1,000)	-----	-----	-----
B	3	24,00 (R)	35,00 (R)	1,000 (1,000-1,000)	10,33 (R)	25,33 (R)	1,000 (1,000-1,000)	1,92 (A)	2,33 (A)	1,000 (1,000-1,000)	-----	-----	-----
C	3	4,33 (R)	8,00 (R)	1,000 (1,000-1,000)	0,33 (A)	14,67 (R)	1,000 (1,000-1,000)	0,58 (A)	3,08 (A)	1,000 (1,000-1,000)	-----	-----	-----
D	3	16,00 (R)	154,60 (R)	1,000 (0,999-1,000)	0,20 (A)	102,80 (R)	1,000 (0,999-1,000)	0,58 (A)	2,83 (A)	1,000 (1,000-1,000)	-----	-----	-----
E	3	5,50 (R)	73,50 (R)	0,999 (0,997-1,000)	13,50 (R)	65,50 (R)	0,999 (0,997-1,000)	0,75 (A)	2,75 (A)	1,000 (1,000-1,000)	-----	-----	-----
F	3	2,33 (R)	19,33 (R)	1,000 (1,000-1,000)	1,67 (A)	28,00 (R)	1,000 (1,000-1,000)	0,96 (A)	2,79 (A)	1,000 (1,000-1,000)	-----	-----	-----
G	3	2,69 (R)	20,15 (R)	0,999 (0,997-1,000)	5,00 (R)	19,71 (R)	1,000 (1,000-1,000)	0,25 (A)	2,25 (A)	1,000 (1,000-1,000)	-----	-----	-----
H	3	32,40 (R)	76,60 (R)	1,000 (0,999-1,000)	26,00 (R)	172,20 (R)	0,999 (0,998-1,000)	0,62 (A)	3,12 (A)	1,000 (1,000-1,000)	-----	-----	-----
L	3	14,33 (R)	14,67 (R)	1,000 (0,999-1,000)	30,57 (R)	324,14 (R)	0,999 (0,995-1,000)	1,12 (A)	3,04 (A)	1,000 (1,000-1,000)	-----	-----	-----
J	4	26,20 (R)	121,20 (R)	1,000 (0,999-1,000)	5,33 (R)	25,00 (R)	1,000 (1,000-1,000)	3,33 (R)	29,00 (R)	1,000 (0,999-1,000)	0,65 (A)	3,04 (A)	1,000 (1,000-1,000)
K	4	5,62 (R)	17,38 (R)	0,999 (0,997-1,000)	13,50 (R)	64,50 (R)	0,999 (0,997-1,000)	1,00 (A)	45,00 (R)	1,000 (0,998-1,000)	1,23 (A)	2,31 (A)	1,000 (1,000-1,000)

A= Aprovado; R= Reprovado - A= Approved; R = Failed

ção é homogênea, ou seja, quando existe pequena variabilidade entre os sujeitos, o poder do CCI para detectar a real concordância tem valor limitado^{12-14,17}. Neste caso, a maior proporção da variância total observada seria determinada pela variância “intra-indivíduos”, o que conduziria a uma diminuição da magnitude do CCI, mesmo diante da boa reprodutibilidade^{13,19}. Diante do exposto, torna-se plausível a recomendação de que, para uma correta interpretação dos valores obtidos para o coeficiente de correlação intraclasse, deve-se levar em consideração as características da população do estudo, no que diz respeito à sua homogeneidade/heterogeneidade, em relação aos valores das medidas de interesse^{14,15,17,19}.

No presente estudo, calculou-se o coeficiente de variação (CV)^{9,20} para verificar a heterogeneidade das medidas de peso e altura dos voluntários, nas respectivas sessões de treinamento. A elevada magnitude do CCI encontrada no estudo poderia ser explicada, ao menos em parte, pela variabilidade das medidas dos voluntários, principalmente no que diz respeito às medidas de peso.

Diante das limitações apresentadas, mesmo sendo utilizado como um indicador de concordância para variáveis quantitativas¹³, o emprego do coeficiente de correlação intraclasse como único indicador de calibração pode ser desaconselhável, tornando-se, por vezes, necessária, a combinação de diferentes técnicas estatísticas com o propósito de avaliar-se, adequadamente, o desempenho dos entrevistadores em fase de treinamento^{18,21}.

Os valores obtidos, segundo a técnica proposta por Habicht¹, indicaram a necessidade de maior número de sessões de treinamento para a aprovação dos entrevistadores. Este método parece ser mais fidedigno, pois, além de permitir o cálculo da precisão, indica também a magnitude da divergência das aferições dos entrevistadores, quando comparadas com as aferições de um supervisor (exatidão), sendo portanto de grande valor para a correção de inconsistências técnicas, por parte da equipe submetida ao treinamento. Há que se considerar, por

outro lado, que se as diferenças de um único par de valores obtidos pelos entrevistadores se distancia das respectivas diferenças do supervisor, as medidas de precisão e exatidão podem ficar comprometidas.

Em relação às medidas antropométricas, torna-se oportuno ressaltar que, embora o peso apresente menor potencial de variabilidade do que as medidas de circunferências ou de pregas cutâneas, há que se considerar que, mesmo com a utilização de balanças eletrônicas de alta precisão, erros de mensuração foram introduzidos durante o treinamento, quando o entrevistador não atendeu, principalmente, a uma ou mais das seguintes recomendações, relacionadas ao posicionamento adequado do voluntário na balança: a) instrução ao voluntário para que ao se posicionar na balança procurasse distribuir o peso do corpo de forma simétrica, evitando apoiar-se com mais firmeza em uma perna que em outra e b) instrução para que o voluntário, ao se posicionar na balança, mantivesse o olhar no horizonte. Com relação à altura, erros de mensuração foram introduzidos quando o entrevistador não atendeu, principalmente, à recomendação do posicionamento da cabeça do voluntário sob o estadiômetro, segundo o plano de Frankfurt^{2,4}.

A existência de diferentes técnicas de aferição para o peso e a altura possibilita a escolha da mais adequada, de acordo com as condições do estudo. Entretanto, apesar da diversidade de técnicas de aferição, são registrados poucos métodos para análise dos erros de medição, o que dificulta a avaliação da calibração dos entrevistadores²¹⁻²³.

O presente estudo foi conduzido com o propósito de comparar duas técnicas estatísticas, freqüentemente utilizadas para a calibração de entrevistadores, em treinamento para a aferição de medidas antropométricas. Com este propósito, os resultados permitiram concluir que a técnica proposta por Habicht¹ se mostrou mais adequada que os coeficientes de correlação intraclasse, visto que permite avaliar além da precisão, a exatidão das medidas quando

comparadas àquelas aferidas por um supervisor (exatidão).

Enfatiza-se, por último, que, para a calibração dos entrevistadores, possíveis diferenças nas condições locais para a exe-

cução do trabalho de campo, em relação às condições do treinamento, sejam levadas em consideração, procurando-se, sempre que possível, adequar o espaço de treinamento ao espaço de coleta^{10,22}.

Referências

1. Habicht JP. Estandarizacion de Métodos Epidemiológicos Cuantitativos sobre el terreno. *Bol Oficina Sanit Panam* 1974; Mayo: 375-84.
2. World Health Organization Report. *Obesity: Preventing and managing the global epidemic*. Geneva; 1997.
3. Bohme MTS. Cineantropometria - Componentes da constituição corporal. *Rev Bras Cineantropom Desempenho Hum* 2000; 2(1): 72-9.
4. Habicht JP, Butz WP. Measurement of health and nutrition effects of large-scale nutrition intervention projects. In: Klein RE. *Evaluation of the impact of nutrition and health programs*. New York: Plenum Press; 1979. p. 133-89.
5. Snedecor GW, Cochran WC. *Statistical Methods*. 6th ed. Ed. Iowa State University Press; 1967.
6. EXCEL. Microsoft Office XP. Microsoft Windows XP Professional. Microsoft Corporation, 1981-2001.
7. SPSS for Windows. SPSS – *Statistical Package in Social Science: Release 10.0.1* (27/outubro/1999), standard version Copyright – SPSS INC, 1989 – 1999.
8. Perini TA, Oliveira GL, Ornellas JS, Oliveira FP. Cálculo do erro técnico de medição em antropometria. *Rev Bras Med Esporte* 2005; 11(1): 81-5.
9. Goto R, Mascie-Taylor N. Precision of measurement as a component of human variation. *J Physiol Anthropol* 2007; 26(2): 253-6.
10. Oliveira Filho A, Oliveira AAB, Oliveira ER, Kurata DM, Pineda M. Variabilidade intra-avaliador e inter-avaliadores de medidas antropométricas. *Acta Sci. Health Sci* 2007; 29(1): 1-5.
11. Rosati P, Bartolozzi F, Guariglia L. Intra- and inter-observer repeatability of femur length measurement in early pregnancy. *Ultrasound Obstet Gynecol* 2004; 23: 599–601.
12. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005; 19(1): 231-40.
13. Bédard M, Martin NJ, Krueger P, Brazil K. Assessing reproducibility of data obtained with instruments based on continuous measurements. *Exp Aging Res* 2000; 26: 353-65.
14. Hopkins WG. Measures of Reliability in Sports Medicine and Science. *Sports Med* 2000; 30(1): 1-15.
15. Olds T. Five errors about error. *J Sci Med Sport* 2002; 5: 336–40.
16. Bartko JJ. On various Intraclass Correlation Reliability Coefficients. *Psychol Bull* 1976; 83(5): 762–5.
17. Looney MA. When is the Intraclass Correlation Coefficient misleading? *Meas Phys Educ Exerc Sci* 2000; 4(2): 73–8.
18. Bruton A, Conway JH, Holgate ST. Reliability: what is it, and how is it measured? *Physiotherapy* 2000; 86(2): 94-9.
19. Shultz SJ, Nguyen AD, Windley TC, Kulas AS, Botic TL, Beynon BD. Intratester and intertester reliability of clinical measures of lower extremity anatomic characteristics: implications for multicenter studies. *Clin J Sport Med* 2006; 16(2): 155-61.
20. Berquó ES, Souza JMB, Gotlieb SLD. *Bioestatística*. 2^a ed. São Paulo: Ed. Pedagógica e Universitária Ltda; 1981.
21. Vegelin AL, Brukx LJCE, Waelkens JJ, Van den Broeck J. Influence of knowledge, training and experience of observers on the reliability of anthropometric measurements in children. *Ann Hum Biol* 2003; 30(1): 65-79.
22. Ulijaszek SJ, Kerr DA. Anthropometric measurement error and the assessment of nutritional status. *Br J Nutr* 1999; 82: 165-77.
23. WHO Multicentre Growth Reference Study Group. Reliability of anthropometric measurements in the WHO Multicentre Growth Reference Study. *Acta Paediatrica* 2006; 450(S): 38–46.

Recebido em: 16/07/07

Versão final reapresentada em: 16/01/08

Aprovado em: 19/03/08