

EVALUACION DE LA CONFIABILIDAD DE LAS MEDICIONES ANTROPOMETRICAS

*Gilda Pareja,¹ Juan Rivera,¹ Jean-Pierre Habicht²
y Hilda Castro¹*

**Instituto de Nutrición de Centro América y Panamá
(INCAP)
Guatemala, Guatemala, C.A.**

RESUMEN

Como parte del Proyecto Desnutrición Temprana y sus Efectos en la Juventud, se condujo un estudio de confiabilidad de medidas antropométricas, en el que se remidieron 226 adolescentes y jóvenes del medio rural de Guatemala.

Para todas las variables antropométricas, los coeficientes de confiabilidad intra-medidor, mayores de 0.96, y los coeficientes de confiabilidad inter-medidor fueron mayores de 0.91.

Se encontró un efecto significativo de medidor, lo que sugiere la existencia de diferencias sistemáticas entre medidores. Esta información permitirá efectuar correcciones por efecto de medidor en el análisis de los datos.

Se evaluó el efecto de diferentes grados de "limpieza" o depuración de los datos en la confiabilidad de las medidas, encontrándose que para las condiciones en que se llevó a cabo este estudio, en el cual los medidores estaban entrenados y supervisados, la eliminación de valores fuera de rangos no resultó en una disminución muy apreciable del error técnico de medición. Como era de esperar, la "limpieza" realizada a través de la comparación de las dos mediciones efectuadas para cada sujeto, eliminando aquéllas cuyas diferencias fueran clasificadas como datos aberrantes según

Manuscrito original recibido: 2-3-90.

1 Científicos, División de Nutrición y Salud, Instituto de Nutrición de Centro América y Panamá (INCAP), Apartado Postal 1188, Guatemala, Guatemala, C.A.

2 Profesor de la Universidad de Cornell, Ithaca, Nueva York, EUA.

criterios estadísticos, disminuyó el error técnico de medición en la mayoría de las variables medidas. Sin embargo, y a excepción de una de las variables, esta disminución no fue muy acentuada, lo que permite inferir que el error técnico de medición logrado con la totalidad de los individuos incluidos en el estudio, para la mayoría de los cuales no se efectuaron remediciones, estuvo dentro de rangos aceptables.

Los resultados del estudio demuestran la importancia de incorporar en el diseño de estudios antropométricos la remediación de un porcentaje de los individuos. Esto permite evaluar la confiabilidad de las mediciones, identificar sesgos debido a medidores, y hacer correcciones de dichos sesgos.

INTRODUCCION

La evaluación de la confiabilidad de las mediciones antropométricas en un estudio es importante para controlar la calidad de la información. Una baja confiabilidad puede ocasionar una atenuación de las correlaciones entre las variables, una necesidad de mayores tamaños de muestra (o menor poder en las pruebas de hipótesis), sesgo en la selección de la muestra y coeficientes de regresión sesgados (1). Si a través de un análisis de confiabilidad se detecta que ésta no está alcanzando niveles aceptables, pueden introducirse medidas correctivas, como un programa de entrenamiento y estandarización de los medidores y/o mayor supervisión en el trabajo de campo.

Varios autores han evaluado la confiabilidad de medidas antropométricas usando el error técnico de medición y componentes de varianza (3), el coeficiente de variación (3), la diferencia entre mediciones como porcentaje del promedio (4), o el coeficiente de correlación intraclase (3). El uso del coeficiente de correlación de Pearson para este propósito ha sido justificadamente criticado (5).

Este trabajo concierne a un análisis de confiabilidad con base en el coeficiente de confiabilidad presentado por Winer (6) como estimador de un promedio de mediciones, y usado por Habicht, Yarbrough y Matorell (7) y Marks, Habicht y Mueller (8). Por otro lado, examina el efecto de medidor a través de métodos de análisis de varianza, utilizando datos antropométricos de adolescentes del medio rural de Guatemala.

Asimismo, se evalúa la necesidad de depurar o "limpiar" los datos antropométricos, es decir, corregir o eliminar datos que se presume son errores de medición. En todo estudio, por más cuidado que se haya tenido al hacer las mediciones, existen errores de medición. Los criterios para establecer cuáles son los valores que están en error y que deben eliminarse o reemplazarse por remediciones, no están bien establecidos. Un valor que es raro por ser demasiado bajo o alto, por ejemplo, puede ser que no sea un error sino que pertenezca a un individuo que es extremo respecto a esa variable. En el trabajo objeto de este artículo, se discuten diferentes procedimientos de depuración o limpieza de datos para tratar de establecer criterios convenientes.

MATERIAL Y METODOS

En el Proyecto Desnutrición Temprana y sus Efectos en la Juventud, se obtuvo información antropométrica, socioeconómica, psicológica, clínica y fisiológica en una muestra de 2,484 individuos con edades comprendidas entre 11 y 26 años, procedentes de comunidades rurales del Oriente de Guatemala que, en su niñez, habían participado en un estudio longitudinal de suplementación alimentaria (véase por ejemplo, Habicht y Yarbrough) (9). Como parte de este Proyecto, y con el objetivo de introducir medidas de control de calidad y posibilitar la detección de problemas y su corrección durante el curso del mismo, se condujo un estudio de confiabilidad.

Alrededor del 10% de los individuos de la muestra del Proyecto fue remedido para estimar la confiabilidad de sus medidas antropométricas. El intervalo entre mediciones varió entre tres y 56 días, no pudiéndose fijar en menos días debido a que los individuos estaban dispersos en siete comunidades y era costoso movilizar los equipos de medidores. En los 19 individuos con mayor intervalo (entre 35 y 56 días) se examinó la diferencia entre sus mediciones por sexo y edad, y se constató que no era resultado de crecimiento, especialmente en las edades en que ocurre el pico máximo de crecimiento durante la pubertad.

Cuatro medidores estandarizados participaron en este estudio, obteniendo mediciones para las 21 variables antropométricas siguientes: peso (kg); talla y talla sentado (cm); diámetros biacromial, biiliocrestal, bicondilar, biestiloideo, biepicondilar (cm); circunferencias de pantorrilla, cefálica, de brazo, de cintura, de cadera y de muslo (cm); y pliegues cutáneos de pantorrilla, muslo anterior, tríceps, bíceps, subescapular, axilar medio y supra-ilíaco (mm).

En la Tabla 1 se exponen los números de sujetos remedidos, en una clasificación cruzada por identificación del primer y segundo medidor.

Los sujetos remedidos se dividieron en dos grupos: el grupo remedido por el mismo medidor que realizó la primera medición, formado por 55 sujetos y denominado grupo intra-medidor, y el grupo integrado por los 171 sujetos que fueron remedidos por un medidor diferente, denominado grupo inter-medidor.

Los pliegues se transformaron a logaritmos naturales (Ln), después de examinar su distribución y encontrar que el desvío estándar era proporcional al promedio (10).

En cada grupo se estimó la confiabilidad de un promedio de mediciones con el siguiente estimador presentado por Winer (6) y utilizado por Marks, Habicht y Mueller (8):

$$R = 1 - Sr^2/s^2$$

donde Sr^2 y s^2 son el cuadrado medio dentro y entre sujetos, respectivamente, en un análisis de varianza a una vía. En el caso del grupo inter-medidor, $1-R$ contiene cualquier diferencia que exista entre medidores. Asimismo, cualquier crecimiento que exista en los

individuos estará contenido en 1-R y confundido con el error de medición.

TABLA 1

**FRECUENCIA DE REMEDICIONES CLASIFICADAS SEGUN
PRIMER Y SEGUNDO MEDIDOR**
Segundo Medidor

		A	B	C	D	
Primer medidor	A	15	15	12	14	56
	B	19	15	16	15	65
	C	15	13	13	13	54
	D	13	14	12	12	51
		62	57	53	54	226

El error técnico de medición, estimado como la raíz cuadrada de Sr^2 , fue calculado para los datos con tres diferentes grados de limpieza: el primer grado de limpieza designa el estado de los datos antes de realizar limpieza, es decir, cuando todos los valores obtenidos fueron incluidos, aunque estuvieran fuera de rango. En el segundo grado de limpieza los valores de las medidas antropométricas que estuvieron fuera de los rangos aceptables definidos para cada grupo de edad y sexo y que fueron considerados como error después de que los supervisores revisaron el conjunto de medidas antropométricas, se omitieron. En el tercer grado de limpieza, se calcularon las diferencias entre mediciones para cada sujeto, y aplicando criterios desarrollados por Tukey (11), se detectaron diferencias fuera de rango, omitiendo luego las mediciones correspondientes. Una diferencia era considerada fuera de rango si 1) excedía el tercer cuartil de su distribución en más de tres veces el rango intercuartil (diferencia entre tercer y primer cuartil), ó 2) estaba por debajo del primer cuartil de su distribución en más de tres veces el rango intercuartil.

Para investigar si existían diferencias sistemáticas entre medidores, se utilizó análisis de varianza, estimando las diferencias entre medidores ajustadas por el efecto de sujetos. Como los datos no son balanceados, se utilizaron técnicas de modelos lineales, ajustando un modelo que incluía un término para el efecto de medidor y otro para el efecto de sujeto. Para cada variable, el efecto de cada medidor se estimó con respecto al medidor que obtuvo las mediciones con el menor promedio, mediante la técnica de contrastes.

RESULTADOS

En la Tabla 2 se presentan estadísticas descriptivas de las variables antropométricas.

TABLA 2
ESTADÍSTICAS DESCRIPTIVAS DE LAS VARIABLES
ANTROPOMETRICAS

Variable (x)	Número de mediciones	Promedio		Desvío estándar	
		\bar{x}	Ln (\bar{x})	\bar{x}	Ln (\bar{x})
Peso	452	43.42	-	11.58	-
Talla	452	148.69	-	11.17	-
Talla sentado	451	78.52	-	6.03	-
Diám. biacromial	450	32.04	-	3.39	-
Diám. biiliocrestal	452	24.28	-	2.64	-
Diám. bicondilar	452	8.24	-	0.66	-
Diám. biestiloideo	452	4.89	-	0.44	-
Diám. biepicondilar	450	5.85	-	0.55	-
Circ. de pantorrilla	452	30.53	-	3.61	-
Circ. cefálica	452	51.28	-	1.73	-
Circ. de brazo	452	22.48	-	3.30	-
Circ. de cintura	450	69.39	-	8.19	-
Circ. de cadera	452	76.42	-	9.12	-
Circ. de muslo	452	42.27	-	5.81	-
P1. de pantorrilla	452	9.42	2.0959	5.63	0.5285
P1. de muslo anterior	452	13.53	2.4578	7.64	0.5384
P1. de tríceps	451	9.75	2.1868	4.43	0.4167
P1. de bíceps	452	5.32	1.5751	2.66	0.4220
P1. subescapular	451	9.31	2.1157	5.16	0.4557
P1. axilar medio	451	7.74	1.9077	4.92	0.4907
P1. supra-ilíaco	452	11.49	2.2773	7.17	0.5574

Los errores técnicos de medición obtenidos con los tres grados de limpieza, para el grupo intra-medidor, se detallan en la Tabla 3. Se observa que en este grupo hubo una reducción en el error técnico de medición en tres de las variables (diámetro biiliocrestal, circunferencia de muslo y pliegue supra-ilíaco), como resultado del tercer grado de limpieza. Sin embargo, no hubo ningún cambio como resultado del segundo grado de limpieza. En la última columna de la misma Tabla, se exponen los valores del coeficiente de confiabilidad para el grupo intra-medidor.

En la Tabla 4 se consignan, los errores técnicos de medición correspondientes al grupo inter-medidor, y se observa que solamente hubo una pequeña reducción en el error técnico de medición en cuatro de las variables después de efectuar el segundo grado de limpieza. En cambio, con el tercer grado de limpieza, el error técnico de medición en la mayoría de las variables se redujo. En algunas, esta reducción fue considerable, como en circunferencia de pantorrilla, donde el error se redujo a la mitad, y en circunferencia de cintura, donde se redujo un 27%. En la última columna de la Tabla 4 se

presentan los valores del coeficiente de confiabilidad para el grupo intra-medidor.

En todos los casos, fue el coeficiente de confiabilidad (Tablas 3 y 4) mayor de 0.91. El coeficiente de confiabilidad intra-medidor fue siempre mayor de 0.96; para nueve variables fue mayor de 0.99, y para 17, fue mayor de 0.98. El coeficiente de confiabilidad inter-medidor acusó el valor más bajo para pliegues de biceps (0.9151), fue mayor de 0.99 para cuatro variables, y mayor de 0.98 para 11 variables. Como era de esperar la confiabilidad intra-medidor en general fue mayor que la confiabilidad inter-medidor.

La Tabla 5 consigna los efectos de medidores ajustados por el efecto de sujetos, para ocho de las variables medidas. Los efectos de medidores se calcularon para el grupo inter-medidor y en relación al medidor que obtuvo el menor promedio. Las pruebas "t" mostraron efecto de los medidores ($P < 0.05$) para todas las variables presentadas, a excepción de peso y circunferencia cefálica. Lo mismo se obtiene usando el criterio de Bonferroni para inferencias simultáneas, con el cual, para que una diferencia sea significativa al 5%, P debe ser menor de 0.016. Se observa que el medidor B parece medir mayor talla que el medidor C; el medidor D menor circunferencia de brazo que el A y el B; el A menores pliegues de triceps, biceps y subescapular que los medidores C y D, y el medidor D, mayor pliegue supra-iliaco que el medidor B.

TABLA 3

ERROR TECNICO DE MEDICION CON DIFERENTES GRADOS DE LIMPIEZA, Y COEFICIENTE DE CONFIABILIDAD (R) CON EL SEGUNDO GRADO DE LIMPIEZA, PARA EL GRUPO INTRA-MEDIDOR

Variable	Error técnico de medición para diversos grados de limpieza			
	Primero	Segundo	Tercero	R
Peso	0.6581	0.6581	0.6581	0.9983
Talla	0.4527	0.4527	0.4527	0.9993
Talla sentado	0.6174	0.6174	0.6174	0.9951
Diám. biacromial	0.6722	0.6722	0.6722	0.9836
Diám. biiliocrestal	0.4306	0.4306	0.3161	0.9861
Diám. bicondilar	0.1176	0.1176	0.1176	0.9857
Diám. biestiloideo	0.1023	0.1023	0.1023	0.9774
Diám. biepicondilar	0.0792	0.0792	0.0792	0.9908
Circ. de pantorrilla	0.3317	0.3317	0.3317	0.9945
Cir. cefálica	0.2185	0.2185	0.2185	0.9934
Circ. de brazo	0.3249	0.3249	0.3249	0.9944
Circ. de cintura	1.2513	1.2513	1.2513	0.9866
Circ. de cadera	0.6557	0.6557	0.6557	0.9970
Circ. de muslo	0.9826	0.9826	0.6131	0.9825
Ln P1. de pantorrilla	0.1013	0.1013	0.1013	0.9822
Ln P1. de muslo anterior	0.0946	0.0946	0.0946	0.9875
Ln P1. de triceps	0.0588	0.0588	0.0588	0.9907
Ln P1. de biceps	0.1146	0.1146	0.1146	0.9628
Ln P1. subescapular	0.0904	0.0904	0.0904	0.9834
Ln P1. axilar medio	0.1051	0.1051	0.1051	0.9769
Ln P1. suprailíaco	0.1226	0.1226	0.0902	0.9763

TABLA 4

**ERROR TECNICO DE MEDICION CON DIFERENTES
GRADOS DE LIMPIEZA, Y
COEFICIENTE DE CONFIABILIDAD (R) CON EL
SEGUNDO GRADO DE
LIMPIEZA, PARA EL GRUPO INTER-MEDIDOR**

Variable	Error técnico de medición para diversos grados de limpieza			R
	Primero	Segundo	Tercero	
Peso	0.7584	0.7584	0.5694	0.9979
Talla	0.6381	0.6381	0.5573	0.9983
Talla sentado	0.9058	0.7368	0.6757	0.9924
Diám. biacromial	0.8713	0.8343	0.8271	0.9664
Diám. biliocrestal	0.5064	0.5064	0.4879	0.9815
Diám. bicondilar	0.2002	0.2002	0.2002	0.9496
Diám. biestiloideo	0.1444	0.1444	0.1444	0.9383
Diám. biepicondilar	0.1610	0.1381	0.1394	0.9655
Circ. de pantorrilla	0.6403	0.6403	0.3263	0.9852
Circ. cefálica	0.3784	0.3784	0.2540	0.9737
Circ. de brazo	0.4348	0.4348	0.4348	0.9916
Circ. de cintura	2.4152	2.1110	1.5537	0.9670
Circ. de cadera	1.4208	1.4208	1.1997	0.9883
Circ. de muslo	0.9646	0.9646	0.6949	0.9669
Ln P1. de pantorrilla	0.1015	0.1015	0.1015	0.9810
Ln P1. de muslo anterior	0.1029	0.1029	0.1029	0.9797
Ln P1. de tríceps	0.0709	0.0709	0.0709	0.9848
Ln P1. de bíceps	0.1660	0.1660	0.1660	0.9151
Ln P1. subescapular	0.1146	0.1146	0.1100	0.9652
Ln P1. axilar medio	0.0939	0.0939	0.0939	0.9812
Ln P1. suprailíaco	0.1388	0.1388	0.1388	0.9675

DISCUSION

El mejoramiento en el error técnico de medición después de realizar el segundo grado de limpieza de los datos fue casi inapreciable. No obstante, habría que señalar que la limpieza se hizo con criterios muy conservadores: los valores que estaban fuera de los rangos definidos por los supervisores eran examinados junto con las otras mediciones antropométricas de los sujetos, y muy pocos valores fueron eliminados. En la totalidad de las variables medidas, no hubo ningún cambio en el grupo intra-medidor, y sólo en cuatro variables se observó una disminución en el error técnico de medición en el grupo inter-medidor.

TABLA 5

ESTIMACION DEL EFECTO DE MEDIDORES EN RELACION AL MEDIDOR EXTREMO

Variable	Efecto	Estimación del efecto	Error estándar	t	Valor-p a dos colas
Peso	B — A	0.0932	0.1367	0.68	0.4963
	C — A	0.1801	0.1443	1.25	0.2138
	D — A	0.1433	0.1443	0.99	0.3222
Talla	A — C	0.1508	0.1176	1.28	0.2016
	B — C	0.3727	0.1157	3.22	0.0015
	D — C	0.0368	0.1205	0.31	0.7605
Circunferencia cefálica	A — C	0.1386	0.0714	1.94	0.0540
	B — C	0.1233	0.0703	1.75	0.0812
	D — C	0.1113	0.0732	1.52	0.1301
Circunferencia de brazo	A — D	0.2652	0.0662	4.01	0.0001
	B — D	0.5539	0.0651	8.51	0.0001
	C — D	0.0151	0.0678	0.22	0.8241
Ln Pliegue de tríceps	B — A	0.0141	0.0115	1.23	0.2200
	C — A	0.0650	0.0121	5.37	0.0001
	D — A	0.0591	0.0121	4.88	0.0001
Ln Pliegue de bíceps	B — A	0.0060	0.0217	0.28	0.7835
	C — A	0.2482	0.0229	10.86	0.0001
	D — A	0.1191	0.0229	5.21	0.0001
Ln Pliegue subescapular	B — A	0.0117	0.0170	0.69	0.4927
	C — A	0.0831	0.0179	4.64	0.0001
	D — A	0.1434	0.0179	8.00	0.0001
Ln Pliegue supra-ilíaco	A — B	0.0046	0.0234	0.20	0.8432
	C — B	0.0002	0.0243	0.01	0.9929
	D — B	0.1047	0.0243	4.30	0.0001

En cambio, se observó una disminución apreciable del error técnico de medición en el grupo inter-medidor al realizar el tercer grado de limpieza, es decir, al omitir los valores de las mediciones cuya diferencia estaba fuera de rango con respecto a su distribución. Esta disminución se observó en 12 de las 21 variables medidas.

En cuanto a la confiabilidad de las mediciones, se observa — como era de esperar— que es más alta en el grupo intra-medidor, y más baja en el grupo inter-medidor, ya que este último contiene el efecto de medidores, que fue significativo (al 5%) en la mayoría de las variables. El coeficiente de confiabilidad de las mediciones para

todas las variables fue mayor de 0.91 y en la mayoría de los casos, mayor de 0.96. En este estudio los valores obtenidos están tan cerca de 1, que se podría afirmar que las mediciones fueron muy confiables. En general, sin embargo, es necesario establecer criterios para definir valores del coeficiente que permitan evaluar si las mediciones en un estudio son poco, medianamente o muy confiables.

Tanto la alta confiabilidad obtenida como el bajo efecto de la limpieza de datos según rangos en el error técnico de medición pueden atribuirse al período de entrenamiento y estandarización de los medidores antes del comienzo del estudio, y al cuidado que se tomó en asegurar la calidad de los datos, cuya recolección se llevaba a cabo bajo la supervisión directa de personal altamente calificado.

Como consecuencia de lo expuesto, la limpieza de los datos usando criterios univariados (limpieza de rangos), sin eliminar la mayoría sino únicamente aquéllos que evidentemente están en error, no resultó en mucha mejoría del error técnico de medición. En cambio, cuando se obtuvieron remediciones, realizando luego limpieza de rangos de las diferencias, se observó una disminución apreciable del error técnico de medición inter-medidor, pero no intra-medidor. La disminución del primero, sin embargo, a excepción de una de las variables, si bien apreciable no fue muy acentuada, lo que permite inferir que el error técnico de medición presente en la totalidad de individuos del estudio está dentro de rangos aceptables. Cuando sea posible efectuar remediciones en un estudio, sería de esperar una reducción del error inter-medidor. Cuando esta medida no sea posible, habría que aumentar el tamaño de muestra de las mediciones únicas para alcanzar el mismo poder que se lograría con las remediciones.

El efecto de medidor requiere ciertas precauciones en la extracción de conclusiones. Es aconsejable realizar análisis ajustados por dicho efecto, utilizando técnicas como las descritas por Fleiss (1). Para evaluar la importancia de realizar un ajuste por el efecto del medidor para un análisis particular, las conclusiones obtenidas de un análisis ajustado podrían compararse con las obtenidas del análisis correspondiente sin ajustar. Si las conclusiones son las mismas, podría presentarse el análisis sin ajustar, pero si varían, sería más apropiado presentar el análisis ajustado.

SUMMARY

EVALUATION OF THE RELIABILITY OF ANTHROPOMETRIC MEASURES

As part of the Early Malnutrition and Its Effects on Youth Project, a study of the reliability of anthropometric measures was carried out through the re-measurement of 226 adolescents of the rural area of Guatemala.

In all anthropometric variables, the intra-measure coefficients of reliability were higher than 0.96 and the inter-measure coefficients of reliability higher than 0.91.

A significant effect of the measure was found, suggesting the existence

of systematic differences among measures. This information will permit corrections by measure effect in the analysis of data.

The effect of different grades of "data editing" and "validity, reliability and agreement check", of measures was evaluated. The conclusion was that even though the persons in charge of measurements were trained and supervised, the deletion of values outside ranks did not result in an appreciable decrease of the technical error of measurement. As expected, data editing and validity carried out through the comparison of two measurements taken from each subject and eliminating those whose differences were classified as error according to statistical criteria, decreased the technical error of measurement, in most of the variables measured. Nevertheless, with the exception of one of the variables, this decrease was not marked. This indicates that the technical error of measurement obtained from all the subjects, many of whom were not re-measured, was within acceptable ranks.

Results of the study demonstrate the importance of including re-measurement of a certain percentage of subjects when designing anthropometric studies. This will allow evaluation of the reliability of measurements, identification of bias due to measurements, and their correction.

BIBLIOGRAFIA

1. Fleiss, J.L. *The Design and Analysis of Clinical Experiments*. New York, N.Y., Wiley & Sons, 1986.
2. Martorell, R., J-P, Habicht, C. Yarbrough, G. Guzmán & R. E. Klein. The identification and evaluation of measurement variability in the anthropometry of preschool children. *Am. J. Phys Anthropol.*, 43: 347-352, 1975.
3. Foster, T.A., L. S. Webber, R. Sathanur *et al.*, Measurement error of risk factor variables in an epidemiologic study of children. The Bogalusa Heart Study. *J. Chron. Dis.*, 33: 661-672, 1988.
4. Bray, G.A., F.L. Greenway, M.E. Molitch, W.T. Dahns, R.L. Atkinson & K. Hamilton. Use of anthropometric measures to assess weight loss. *Amer. J. Clin. Nutr.*, 31: 769-773, 1978.
5. Bland, J.M. & D.G. Altman. Statistical methods for assessing agreement between two methods of clinical measurements. *The Lancet*, Feb. 8: 307-310, 1986.
6. Winer, B.J. *Statistical Principles in Experimental Design*. 2nd ed. New York, N.Y., McGraw Hill, 1971.
7. Habicht, J-P., C. Yarbrough & R. Martorell. Anthropometric field methods: Criteria for selection. In: *Nutrition and Growth*. D.B. Jelliffe and E.F.P. Jelliffe (Eds.). New York, N.Y., Plenum Press, 1979, p. 365-387.
8. Marks, G.C., J-P. Habicht & W.H. Mueller. Reliability, dependability and precision of anthropometric measurements-NHANES-II. *Am. J. Epidemiol.* (In press).
9. Klein, R.E., J-P. Habicht & C. Yarbrough. Some methodological problems in field studies of nutrition and intelligence. In: *Nutrition, Development and Social Behavior*. D.K. Kallen (Ed.) Washington, D.C., United States Government Printing Office, 1973, p. 61-75. (DHEW Publication No. (NIH) 73-242).
10. Snedecor, G.W. & W.G. Cochran. *Statistical Methods*. 7th ed. Ames, Iowa, The Iowa State University Press, 1980.
11. Tukey, J.W. *Exploratory Data Analysis*. Wesley Publishing Company, 1977.