ANNALS OF HUMAN BIOLOGY, 1975, VOL. 2, NO. 2, 153–171

The 1972 Cuban national child growth study as an example of population health monitoring: design and methods

J. JORDAN, M. RUBEN, J. HERNANDEZ
Department of Growth and Development, Children's Institute and Ministry of Public Health, Havana

A. BEBELAGUA
Bureau of Census, National Planning Board, Havana

J. M. TANNER
Department of Growth and Development, Institute of Child Health, University of London

and H. GOLDSTEIN
National Children's Bureau, London

[Received 4 June 1974]

Summary. A stratified 3-stage random sample of 50 360 children ages 0–19 drawn from the whole population of Cuba was measured in a large-scale growth study during 1972–73. Fifteen anthropometric measurements were taken and puberty stages and menarche status were assessed. Hand-wrist radiographs were done on 10% of the sample. Information regarding the social and educational status of the parents was obtained and parental heights were measured. One year later, 30% of the sample were re-measured and 10% re-X-rayed. The overall response rate was 96% at pre-school and primary school age, and a little less later. Quality control sessions were held at which the nine measuring teams compared results. No significant differences were found between teams working different parts of the island. Differences between duplicate measurements of stature by individual measurers had standard deviations approximating 0·20 cm. Individual measurers' means differed from the grand means of all measurers by up to ±0·2 cm for stature, ±0·4 cm for sitting height and ±0·5 mm (7% of mean) for triceps skinfold.

One or more measurements lay outside the 3rd–97th centile limits in 21% of individuals. Scrutiny of these individuals' records resulted in elimination of measurements in amounts ranging from 0·1% (triceps skinfold) to 1·1% (crown–rump length).

Problems of planning and execution of growth surveys designed to set national standards are described, and solutions given or suggested.

1. Introduction

The physical growth and development of children is now widely recognized as a sensitive index of the health and nutrition of a population, and as developing

countries with poor nutrition and high infant mortality rates improve their health standards, growth studies become increasingly important in the evaluation of health care (Tanner, 1966; Jordan, 1973).

There are many difficulties, theoretical as well as practical, in designing and setting up comprehensive surveys to carry out this monitoring function. Standards representative of the whole country (or geographic area) have to be obtained, although in developing countries one area, usually the capital city, may be better off than others, and in all countries some social or occupational groups are better off than the remainder. It is therefore highly desirable to have samples from each of such groups; urban and rural, well-educated and little-educated, office workers and manual labourers. If the monitoring system is designed to measure the difference between such groups so that steps may be taken to diminish or eliminate them, then this implies that surveys should be repeated at suitable intervals.

The practical problems in making adequate surveys are formidable. To date, only Holland (van Wieringen, Wafelbakker, Verbrugge and de Haas, 1965) has standards for height and weight attained at each age from birth onwards based on approximately a national sample. The United States Health Examination Survey produced similar statistics for ages 6–17 (Hamill, Johnston and Grams, 1970; Hamill, Johnston and Lemeshow, 1973), and the New Zealand Department of Health (1971) for ages 5–15. No country-wide standards for velocity of growth have been published. It is the object of this paper to give details of the design and methodology of a national survey carried out during 1972 and 1973 in Cuba.

2. Subjects and methods

In 1969 the Cuban Ministry of Public Health began an intensive programme to reduce the infant mortality rate, which was 44 per 1000 live births in 1970, to 22 per 1000 live births in 1980. (By 1972 the rate was 28; the rate in Engand and Wales in 1972 was 17.) Simultaneously a national growth study was set up, on the grounds that success in achieving the hoped-for reduction in the infant mortality rate would lessen the future importance of this rate as an index of public health, and would make assessment of the growth of children an appropriate additional measure.

A study of this magnitude could not have been carried out before 1959, but since the revolution the increasing development of public health organization and the fully planned nation-wide network of health services have provided reliable statistical information covering the whole country. At the same time, the development of effective operative groups or mass organizations like the Federation of Cuban Women made possible on a voluntary basis the formidable task of visiting all houses selected in a nation-wide random sample.

In 1970 the Instituto de la Infancia of Cuba brought together representatives of the Ministry of Public Health, the Census Bureau, the Federation of Cuban Women and the University of Havana (Department of Paediatrics and Anthropology), and, through the World Health Organization, the Department of Growth and Development at the Institute of Child Health of the University of London. A team was jointly set up which carried through all stages of the planning and execution of the survey, and is collaborating in the analysis. The Ministry of Public Health provided part of the personnel, resources and logistic facilities. The Federation of Cuban Women and the National Association of Small Farmers helped with the

construction of the sample frame, the Census Bureau assisted with the selection of areas, and the Central Planning Authority provided computer facilities.

Since standards for growth velocity were required as well as for growth attained, the survey was planned in two stages; 30% of the children measured in the first stage were measured again one year later. This permits calculations of centiles for annual velocities as in the Tanner, Whitehouse and Takaishi (1966) growth velocity standards.

In 1970 the population of Cuba was 8 575 000, of whom 3 900 000 were between 0 and 19 years of age. A sample of 56 000 children was selected for measurement as described below. Between 13 and 15 anthropometric measurements were taken on each child. Ratings of pubertal development of breast, genitalia, and pubic hair were made, and at appropriate ages girls were asked whether menstruation had begun (see below). In 10% of the sample, radiographs of the left hand and wrist were made and skeletal maturity assigned following the TW2 system (Tanner, Whitehouse, Marshall, Healy and Goldstein, 1975).

The statures of both father and mother were measured, and no hearsay heights were used. No other measurements of the parents were taken, but we hope to obtain some later. Information was obtained on income and occupation of parents, educational levels of parents and child, size of family, birth order, and characteristics of previous pregnancies. Whether or not the child had attended a day-care centre (circulo infantiles) was recorded.

A pilot study was carried out in May 1971 (see below); the first stage of the survey began on 10 May 1972 and finished on 12 April 1973; the second stage began on 10 May 1973 and finished on 23 January 1974.

Sample structure

The total size of the sample was dictated by the desire to establish within fairly narrow limits the 3rd and 97th centiles for the national measurement standards. The sampling errors of these centiles are greater than those of the 50th centile (Tanner, 1952; Healy, 1974), yet it is these outlying centiles that are the critical ones for much clinical work. The limits of sampling error are smallest for measurements such as stature, in which one can assume a Gaussian distribution, and somewhat larger for those such as weight, arm and calf circumference and skinfolds, where the form of the distribution is not known with any precision. Even for stature, however, the standard error of the third centile at age 8, for example, is 0·7 cm for a sample of 200 and 0·3 cm for a sample of 1000 (Healy, 1974) and this is roughly double the standard error of the 50th centile for the same sample size.

For this reason a figure of approximately 1000 boys and 1000 girls was chosen as the basic number for each year of age in the age range 3–9, giving 95% confidence limits of $\pm 0{\cdot}6$ cm for the 3rd height centile. At earlier ages and in adolescence the fact that growth is faster implies that a larger sample (effectively a more frequent age sampling) is needed. Population standards should be estimated with the same accuracy at each age. Thus if the total age range is divided into intervals in each of which the amount of growth is approximately the same, equal numbers of children should be sampled within each interval. This leads to sampling numbers of children proportionate to the growth rate at each age. A one-year interval was chosen for the age range 3–9 years, since in a 1 year age range it is reasonable to approximate the average growth by a straight line at these ages (Goldstein, 1972).

One complication which arises is that growth curves are not the same for all measurements. The final age structure chosen was an average for height, weight and triceps skinfold measurements. Boys and girls were treated separately because their adolescent spurts take place at different ages. Table 1 shows the sample structure planned, as a percentage of total estimated child population of Cuba at each age. Since growth continues for two years longer in boys than in girls, a somewhat larger sample of boys is required to obtain the same precision at each age in both sexes. As reliable estimates of adult measurement were desired, and the age at which growth ceases in Cuba was not accurately known, the sample was augmented at older ages (i.e. 18 and over in boys, and 16 and over in girls).

The age groups from which the statistics of growth attained were to be calculated were chosen so that approximately the same number of subjects was present in each (see table 2). These differ from the ones conventionally used in the first three years. At puberty half-year intervals were used, at different ages for boys and girls.

Age group	Boys	Girls	Age group	Boys	Girls
0+	13	13	10+	4	8
1+	8	8	11+	4	8
2+	5	5	12+	8	8
3+	5	5	13+	8	3
4+	4	4	14+	8	3
5+	4	4	15+	3	3
6+	4	4	16+	3	3
7+	4	4	17+	3	3
8+	4	4	18+	3	3
9+	4	4	19+	3	3
			Total	102	100

Table 1. Approximate percentages of total sample for each sex at each year of age (drawn proportional to rate of growth of average of 3 measurements).

Sample selection

The sampling frame was based on the results of the 1970 national census. Cuba is divided into six provinces (figure 1), but the two most highly populated (Havana and Oriente) were sub-divided, giving eight in all. As a first step the number of children of each year of age measured in each province was made proportional to the percentage of total Cuban children of that age living in that province, separately for rural and urban areas of the province as defined by the census. Thus the children aged 5 years registered in the urban areas of Havana province in 1970 constituted 17·8% of all urban Cuban children of this age while those of the rural areas of Oriente province constituted 23·7% of all children in rural areas. Since the total sample called for approximately 1100 girls and 1100 boys of this age (to give the numbers in table 2) we needed approximately 200 urban Havana children of each sex, and 260 of each sex from Oriente.

Interval (yr)	Number of subjects		Age centre (yr)
0·00–0·33	1200		0·167
0·34–0·66	1200		0·500
0·67–0·99	1200		0·833
1·00–1·49	1000		1·250
1·50–1·99	1000		1·750
2·00–2·74	1000		2·375
2·75–3·49	1000		3·125
3·50–4·49	1300		4·000
4·50–5·49	1100		5·000
5·50–6·49	1100		6·000
6·50–7·49	1100		7·000
7·50–8·49	1100		8·000
8·50–8·49	1100		9·000
	Boys	Girls	
9·50–10·49	1100	1700	10·000
10·50–10·99	600	1100	10·750
11·00–11·49	600	1100	11·250
11·50–11·99	600	1100	11·750
12·00–12·49	1100	1100	12·250
12·50–12·99	1100	1100	12·750
13·00–13·49	1100	400	13·250
13·50–13·99	1100	400	13·750
14·00–14·49	1100	400	14·250
14·50–15·49	1500	800	15·000
15·50–16·49	800	800	16·000
16·50–17·49	800	800	17·000
17·50–18·49	800	800	18·000
18·50–19·99	1300	1300	19·250
	28 100	28 100	

Table 2. Age intervals and approximate target numbers in each

Figure 1. Map of the Republic of Cuba to show provinces, with numbers of children selected in each set of urban and rural areas.

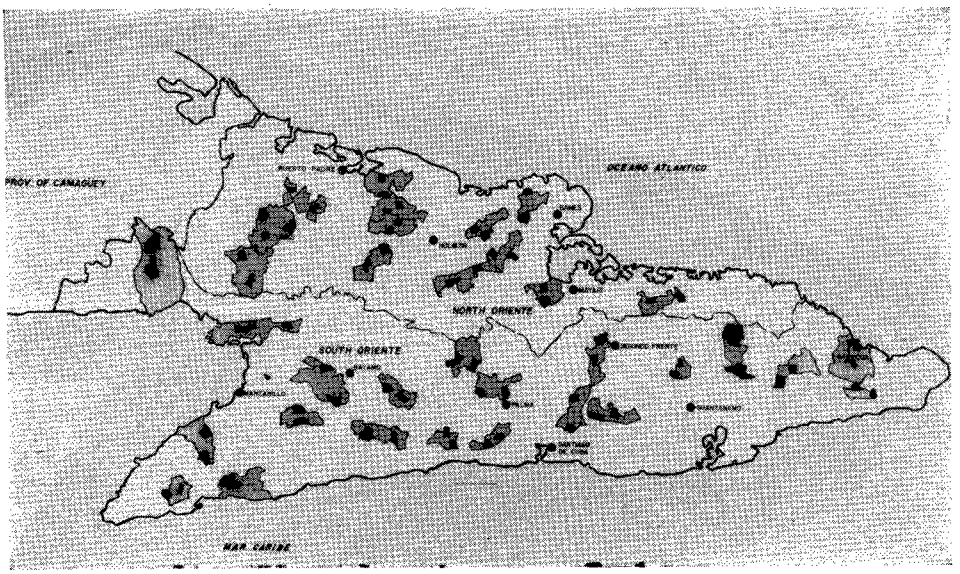

Figure 2. Illustration of three-stage sampling in Oriente province, rural areas. 34 selected areas (out of 254) shown in grey. Within each area, two sectors were randomly selected, shown in black. Average number of all sectors per area was 12.

The sampling procedure within each province is illustrated in figure 2. A stratified 3-stage random sample was selected. For the purposes of the study each province was divided into areas comprising 1200–2000 households, each exclusively urban or rural. A number of such areas was selected randomly from each "urban-province" or "rural-province", the number being proportionate to the child population there. This led to a total of 220 areas being selected. In the rural part of Oriente, for example, the shaded areas of figure 2 were selected. Areas were further divided in the census into sectors, each comprising about 120 households. Within each of the selected areas two sectors were selected at random (indicated in black in figure 2), leading to a total of 440 sectors. A selection of the children in these 440 sectors were measured. A list was made of all children (persons aged 0 to 19) living in a given sector, together with their ages. This list was checked by the local division of the Women's Federation who updated the census data by visiting each dwelling in the sector. In order to obtain a proper sample of the 0–0·99-year age-group all women in the sixth to ninth month of pregnancy were listed, and children who were born at the time of the visit of the measuring team were measured.

These lists were sent to the central office, and a computer randomly selected individuals within each year of age, to give the numbers required plus 8%. This 8% was added to allow for the expected number of non-respondents. The number of areas and sectors per area had been so chosen that about twice as many individuals appeared on the list as were needed for the samples. The sampling was from a list of children, therefore, not a list of householders, as is sometimes done. No notice was taken of whether siblings (by chance) entered the list, though the fact of their being sibs was indicated on the forms. Since there is variation in family size between provinces this led to selection of different average numbers of children per household. For example, in Havana it was 1·5 and in Oriente 3·1.

In listing in each province no notice was taken of disease or congenital malformations. At the measuring stage obvious cases of such things as chondrodystrophy and Down's syndrome were excluded; further exclusions were made by the medical officers at the centre when the forms (which had a space for clinical assessment) were received. However, only cases of specific disorders such as congenital heart disease, sickle cell anaemia and congenital malformations were removed. Children of all birthweights, and also twins, were included.

Such a complete sampling frame requires primarily a large cadre of field workers to check the lists of individuals. This was provided in Cuba by personnel from the Federation of Cuban Women. Listing was carried out 50 days before the date fixed for the beginning of measurement in the area in question. Lists of individuals selected were returned to the sectors 10 days before measurement began. This schedule was successfully followed in every instance.

Measuring teams and training

One measuring team was assigned to each province, with one in reserve nationally. There were nine teams, each of seven persons (provincial co-ordinator, clerk, two measurer-recorders who each measured or recorded on alternate days, X-ray technician, and two drivers). The measurers were specially recruited for the survey. They were secondary school graduates, but had no higher education. All were females, which facilitated measuring girls, but created some problems in respect of adolescent boys (see below). All X-ray technicians were male, and did puberty ratings of the male adolescents and measurement of their thigh circumferences. Each team had a converted truck, in which all measuring equipment was set up, as measuring centre, which in most areas remained in one place for one week. Each team also had a Land Rover for transport of the team and for fetching children from outlying homes to the truck. In urban areas trucks were less used, the measuring stations usually being in public buildings.

Each team had a Harpenden stadiometer, anthropometer and skinfold caliper, portable infant-measuring table, a tape measure, weighing machine, portable X-ray equipment and portable generator. The measurers were trained together by a single expert, Mr. R. H. Whitehouse, for a period of two weeks, ending with a test of each individual against the instructor. After the survey started, continuous monitoring of the inter-reliability of the individuals and teams was carried out, as described below.

Pilot study

The teams were trained in March 1971, and in May one year before the main survey started, a pilot study was undertaken. Children were assembled and measured in four of the sectors, two being in Havana rural and two in Havana urban areas. The attendance was 85% in the urban and 96% in the rural areas, and no follow-up effort was made on the non-respondents. Various problems emerged, which would have caused great difficulty if a main survey had started without a full pilot study having been done. In rural areas, more frequent moves of the teams were necessary than had originally been envisaged; in a few areas the measuring station was set up for only a single day (though never less). Thus trucks were acquired, which carried all the measuring and X-ray equipment, as well as tents in which the measuring teams were from time to time constrained to live.

Secondly, the number of persons who could be measured in the working day was reduced to 35–40, instead of 45–50 as originally estimated. This was largely due to the time taken in filling in forms, undressing and dressing, etc., rather than to the measuring itself.

Thirdly, the use of cassetted film for X-rays had to be abandoned in favour of the much more convenient non-cassette film, despite the greater exposure required.

X-ray exposure

Tests were carried out during the pilot study to determine the amount of radiation to which the children were subjected. Parents were aware of the issues involved and reassured as to safety. Exposure meters were placed by the gonads, chest and the exposed hand. All measurements were well within international safety limits, as assessed by the Cuban National Program for Protection Against X-rays and Ionising Radiations, and were approved by the National Board of Radiology of the Ministry of Public Health. The skin exposure (non-cassette film) was about 50 millirads; the exposure of the gonads negligible.

Publicity and liaison

A considerable effort of publicity was made, so that as many people as possible became aware of the existence and aims of the study before they were actually approached for co-operation. This aspect of the planning was very important. About three months before the listing began there was a press conference followed by front page publicity in the national press. The national TV network for a period of several months featured a brief news announcement daily, either at peak viewing hours or at the hour of the youth programme: the study had a symbol of a child growing and this was impressed on the viewers. Local radio stations also carried news announcements, specially adapted to the rural audiences, stressing the aims and importance of the study; in each region these were timed to precede the visits of the members of the Women's Federation for listing. As the study has proceeded special follow-up weeks of publicity have been made at intervals of three to four months, and it is intended to continue this, including results as they become available.

A doctor, statistician, and a Women's Federation representative from the central co-ordinating group visited each province for a one-day conference with local Federation workers, etc., two weeks before the listing started. During the pilot study a representative of the Women's Federation from each region came to Havana and also one representative of the regional office of the Ministry of Public Health. These persons stayed for a week; they had two days of theoretical instruction and then they themselves carried out the listing in the four sectors used as the pilot. They were thus able to instruct the sector personnel for the main study from a basis of genuine experience.

Quality control: Monitoring the reliability and validity of measurement

It was initially planned to have each team remeasure routinely some 5% of its children, repeating on the last day of each month its precise procedure of the day before. In addition it was hoped that teams could be paired so that Team A would on occasion measure children previously measured by Team B, in a site on the junction of the two teams' areas. This plan had to be abandoned. Instead, particular quality control sessions were substituted, at which either all teams or

else the four eastern and four western ones met for three days and measured a selection of persons. During the first stage five such sessions took place, at roughly two-month intervals. In the second stage two meetings took place, but not all teams attended the second, since some had by that time finished working. Each team took part in seven quality control sessions.

The quality control procedures had two main purposes. First, they were to monitor and control the extent of variation shown by each measurer, and second, they were to give an estimate of differences between teams working in different provinces, since between-province differences were to be examined in the results.

At each meeting each measurer measured stature, sitting height, bi-iliac diameter, head and arm circumference twice and triceps skinfold once on a sample of adolescents aged 11 to 16 and on a sample of adults. The individual measurer's duplicates were made after a minimum of six other persons had been measured in between. The effects of change in measurement during the day were minimized by each subject having each particular measurement (e.g. stature) completed by all measurers within the space of one hour. Three dimensions only were taken on each individual but each dimension was measured in 10 different persons by each measurer. Supine length and crown–rump length were measured once each by each measurer in a group of 15 children aged 1–3 years. The children and adolescents measured were different at each meeting, but the adults were the same at six of the 11 meetings.

3. Results

We shall describe here the degree of completeness of the sample, the results of the quality controls, and the results of scrutinizing the data for errors. Results of the survey itself will be presented in another paper.

Age	Boys %		Girls %	
0+	93	(3212)	93	(3152)
1+	93	(1935)	94	(2210)
2+	96	(1448)	95	(1608)
3+	96	(1765)	96	(1457)
4+	94	(1114)	96	(1091)
5+	94	(1105)	96	(1083)
6+	96	(1111)	96	(1107)
7+	96	(1115)	95	(1124)
8+	96	(1108)	96	(1122)
9+	95	(1123)	95	(1168)
10+	95	(1145)	95	(2259)
11+	93	(1127)	95	(2281)
12+	93	(2159)	94	(2261)
13+	93	(2041)	90	(883)
14+	88	(1998)	88	(896)
15+	80	(870)	89	(879)
16+	69	(895)	87	(891)
17+	51	(915)	86	(857)
18+	55	(941)	84	(1022)
19+	45	(670)	80	(632)
Total	89	27 833	93	27 983

Table 3. Percentage of children measured at each age out of total selected for measurement. Total numbers selected shown in brackets.

Age	Sickness		Vacation		Boarding school		Military service	Refusal		Other and unknown causes		Total	
	B	G	B	G	B	G	B	B	G	B	G	B	G
0+	28	25	3	5	—	—	—	4	1	89	181	124	212
1+	11	5	2	4	—	—	—	—	2	70	111	83	122
2+	—	2	—	—	—	—	—	3	—	23	74	26	76
3+	3	1	3	3	—	—	—	—	—	35	58	41	62
4+	2	2	1	1	—	—	—	1	—	34	39	38	42
5+	6	2	2	1	—	—	—	1	1	25	43	34	47
6+	0	1	—	1	—	—	—	—	2	25	43	25	47
7+	1	1	1	2	1	3	—	—	1	20	48	23	55
8+	2	1	5	5	4	—	—	—	—	18	36	29	42
9+	5	2	1	5	5	4	—	2	—	25	46	38	57
10+	1	2	3	4	4	10	—	3	—	20	104	31	120
11+	2	3	3	1	14	8	—	—	1	32	100	51	113
12+	2	4	8	4	21	23	—	1	5	63	99	95	135
13+	3	3	4	3	55	27	—	8	—	63	59	133	92
14+	3	6	6	4	65	26	—	7	2	101	77	182	115
15+	4	3	1	1	45	21	—	2	2	77	74	129	101
16+	2	6	1	1	63	20	62	5	—	116	94	249	121
17+	3	5	1	—	51	17	169	1	4	156	97	381	123
18+	5	4	6	1	35	19	181	1	4	131	130	359	158
19+	3	3	2	1	18	16	157	4	6	117	101	301	127
Total	86	81	53	47	381	194	569	43	31	1240	1614	2372	1967

Table 4. Reasons for non-response. B, boys; G, girls.

Completeness of sample

In table 3 the numbers of children in each group selected are given, together with the percent response. The reasons for non-response are given in table 4. Arrangements were initially made to fetch boarding school pupils and military service personnel for measurement but they were not adequately followed, since it was thought that the numbers involved would be smaller than turned out to be the case. The percentage of children at boarding school was relatively small, though in future surveys it will be necessary to make workable arrangements, probably by sampling boarding schools specially, since the number of these schools is rising rapidly. Arrangements with the armed forces are clearly highly important for the age groups 17, 18 and 19 in boys.

In other age groups and in girls, however, the chief reason for non-attendance was unspecified. We must suppose that this was chiefly a form of passive refusal ("refusal" in the table represents a spoken statement). Puberty and simply getting older and more independent seem the most likely causes. Girls in this category rose from 5% at 10 and 11 to 7% at 13, 9% at 14 and 15, 11% at 16 and 15% thereafter. Boys began to rise only at age 15 (10%) and thereafter increased to 20% at 16 and 25–35% from 16 to 19. Part of the problem was the preponderance of female measurer-recorders. Another problem was the increasing independence of older boys, who were often unwilling to attend measurement sessions especially with younger children present. Urban attendance was at most ages lower than rural.

In the second stage a special effort was made to secure as high a number as possible of the non-respondents, to see if they differed in their measurements from the remainder, and thus to determine whether a bias had been introduced. In Havana province a particularly detailed follow-up study was made. The results are not yet to hand.

After the end of the second stage a sample of boys between 16 and 19 years of age in military service was secured through the co-operation of the armed services. Their measurements are to be compared to those in the first stage to see whether they can be used to augment the first stage data.

Team	Control No. 1	2	3	4	5	6	7	Average
Pinar del Rio	0·212	0·205	0·363	0·240	(a)	0·234	(b)	0·251
Havana								
Metropolitan	(a)	(a)	0·264	0·254	0·183	0·149	0·184	0·207
Interior	0·234	0·240	0·166	0·186	0·242	0·190	(b)	0·210
Matanzas	0·240	0·122	0·084	0·180	0·166	0·151	(b)	0·157
Las Villas	0·290	0·172	0·198	0·230	0·209	0·279	0·236	0·231
Camaguey	0·252	0·230	0·213	0·225	0·198	0·267	(b)	0·231
N. Oriente	0·264	0·161	0·196	0·200	0·154	0·303	0·206	0·212
S. Oriente	0·262	0·148	0·206	0·185	0·156	0·176	0·186	0·188
All Teams	0·251	0·183	0·211	0·213	0·187	0·219	0·203	0·211

(a) The team did not participate due to illness of one measurer.

(b) The study in these provinces had ended.

Table 5. Standard deviation (cm) of differences between duplicate measurements, averaged for the two measurers of each team; stature in adolescents. (Each based on 10 duplicates.)

Quality control

Table 5 compares the reliability of the eight teams in their measurements of stature. The values are the standard deviations of the differences between duplicate measurements, with the figures for the two measurers of each team averaged. (No significant differences occurred between individual measurers within teams.) Typically, the SD of the differences was about 0·2 cm. There was no significant difference between teams, nor between different quality control occasions (though the first occasion appears to be a little worse than the others).

Table 6 shows the variance between measurers within each provincial team averaged over all the quality control occasions, giving values for sitting height, head circumference, bi-iliac diameter and arm circumference as well as stature. The teams did not differ significantly in their reliability in any of the measurements.

	Pinar del Rio	Havana Interior	Havana Metropolitan	Matanzas	Las Villas	Camaguey	N. Oriente	S. Oriente	Average variance	Bartlett test of variance χ^2 (7 d.f.)
Stature	0·066	0·045	0·037	0·027	0·054	0·054	0·047	0·037	0·045	0·95
Sitting height	0·142	0·123	0·051	0·068	0·195	0·146	0·103	0·117	0·130	13·44
Head circum.	0·020	0·012	0·016	0·007	0·013	0·017	0·021	0·013	0·014	4·58
Bi-iliac diam.	0·010	0·028	0·018	0·050	0·014	0·051	0·036	0·008	0·021	3·44
Arm circum.	0·021	0·016	0·022	0·070	0·023	0·020	0·020	0·017	0·020	2·66

$\chi^2_{0·05}(7 \text{ d.f.}) = 14·1$

Table 6. Variance (cm^2) between measurers with each provincial team (adolescents).

Figure 3 illustrates the changes of reliability with time in seven individual measurers, one from each team being randomly selected. The difference of reliability between measurers was slight and did not vary significantly during the course of the survey. For the most part the SD of duplicates lay between 0·1 and 0·3 cm. Nobody significantly improved or deteriorated.

Turning to the consideration of the bias in individuals and teams, table 7 shows the differences for stature between the means of each team and the overall means on each occasion. The Las Villas team seemed usually to measure a trifle taller than most of the other West area teams, but the difference was small and did not change significantly with time. Table 8 summarizes the results for the other measurements. Mean differences were slight.

Figure 4 shows the deviations of four measurers on six control sessions, from their collective mean on each occasion (stature, adolescents). Though some measurers consistently measured higher than others, the differences were less than $\pm 0·2$ cm. In sitting height (figure 5) the differences were somewhat larger. Figure 6 shows the situation for triceps skinfold, figure 7 for supine length and figure 8 for crown–rump length in infants.

The nine adults measured, unlike the adolescents, were the same in five of the sessions and hence absolute measurements are plotted for stature and sitting height (figures 9 and 10). Measurers varied from occasion to occasion over a range of about $\pm 0·2$ cm for stature and $\pm 0·4$ cm for sitting height.

The results of the quality control did not indicate that it was necessary to make any adjustment to sample means or variances because of measurer errors or consistent differences between measuring teams.

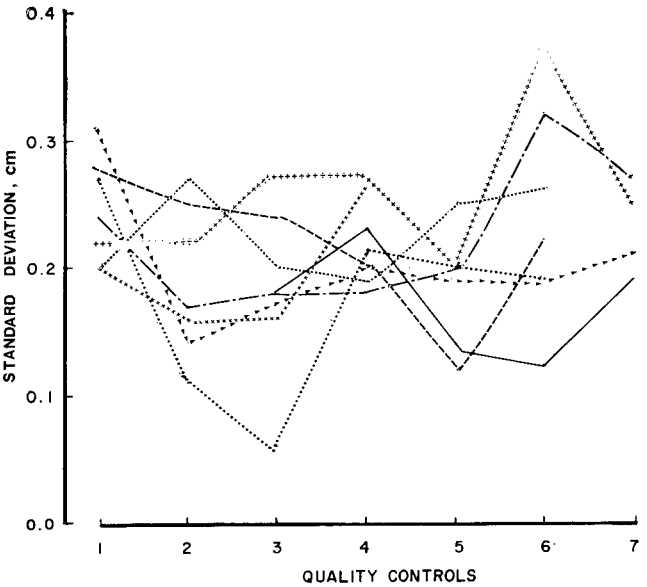

Figure 3. Reliability of eight measurers as assessed at successive quality control sessions. Within measurer standard deviations for stature in adolescents.

Location	Team	Quality Control 1	2	3	4	5	6	7
	Pinar del Rio	−0·171	0·0	—	—	(a)	−0·030	(b)
	Havana Metrop.	(a)	(a)	−0·005	−0·080	−0·026	0·136	−0·185
West	Havana Interior	−0·218	−0·135	−0·149	−0·037	−0·234	−0·198	(b)
Meetings	Matanzas	0·137	−0·035	—	—	0·154	0·197	(b)
	Las Villas	0·232	0·170	0·211	0·085	0·106	0·167	0·185
	Camaguey	—	—	−0·057	0·033	—	—	—
	Oriente North	0·0	0·022	−0·123	−0·044	0·053	0·034	0·190
	Oriente South	0·150	0·046	−0·091	−0·029	0·045	0·064	−0·190
East	Camaguey	−0·150	−0·068	—	—	−0·097	−0·099	(b)
Meetings	Pinar del Rio	—	—	0·050	−0·006	—	—	—
	Matanzas	—	—	0·162	0·081	—	—	—

(a) Team not complete due to illness of one measurer.

(b) Survey had already ended by that date.

Table 7. Differences between means of each team, and overall means of all teams at central meeting (cm). Stature, adolescents.

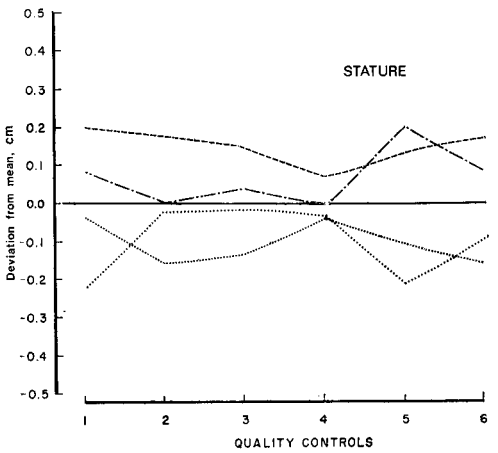

Figure 4. Difference of four individual measurers on six occasions from their mean on each occasion. Stature in adolescents.

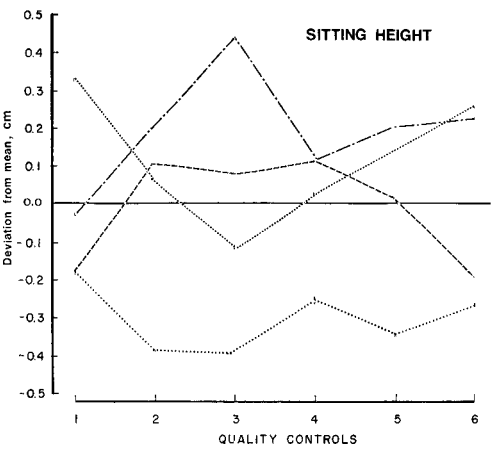

Figure 5. Differences of four individual measurers on six occasions from their mean on each occasion. Sitting height in adolescents.

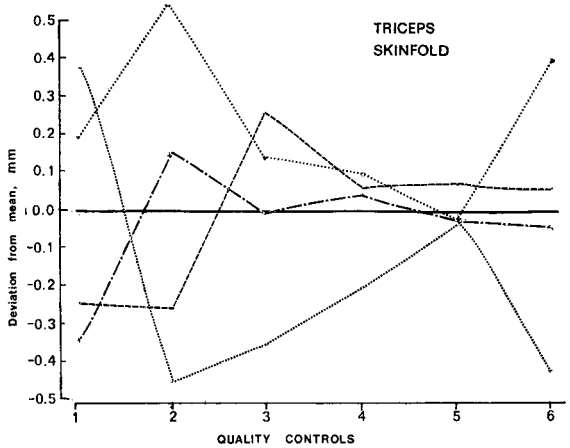

Figure 6. Differences of four individual measurers on six occasions from their mean on each occasion. Triceps skinfolds in adolescents.

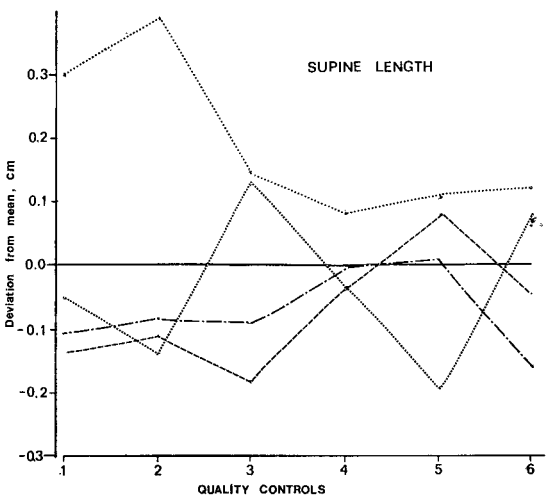

Figure 7. Differences of four individual measurers on six occasions from their mean on each occasion. Supine length.

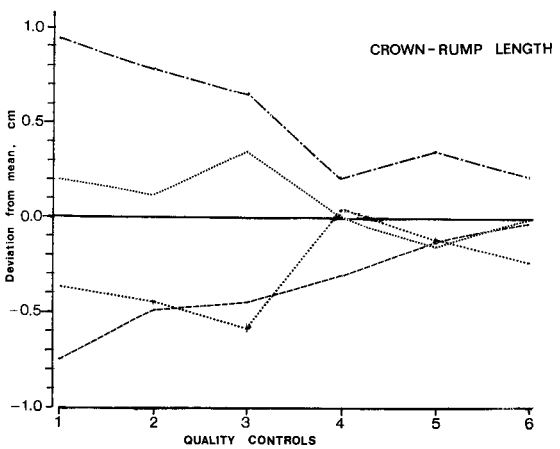

Figure 8. Differences of four individual measurers on six occasions from their mean on each occasion. Crown–rump length.

Measurements	WEST			EAST	
	Havana Metrop.	Havana Interior	Las Villas	Oriente North	Oriente South
Stature (cm)	161·79	161·66	162·00	158·10	158·12
Sitting height (cm)	84·25	84·21	84·42	82·31	82·48
Head circumference (cm)	54·17	54·22	54·22	53·99	54·01
Arm circumference (cm)	22·57	22·69	22·72	21·90	21·92
Triceps skinfold (mm)	7·3	7·2	7·0	7·3	7·2
Bi-iliac diameter (cm)	24·24	24·40	24·42	24·69	24·64

Table 8. Average values (cm) of measurements for each of five teams in five quality controls (adolescents).

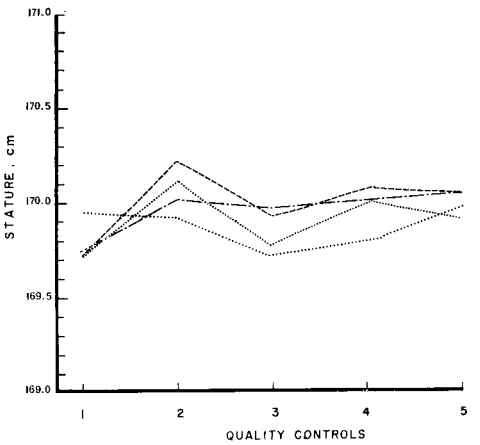

Figure 9. Means of measurements of stature of 9 adults obtained by each of 4 measurers on five occasions.

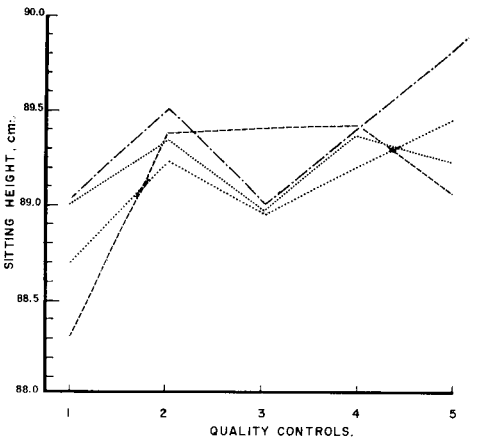

Figure 10. Means of measurements of sitting height of 9 adults obtained by each of 4 measurers on five occasions.

Editing the data

All anthropometric survey data need to be scrutinized for errors and edited. Approximate 3rd and 97th centile values were fixed for each measurement at each age, based on a sample of 4000 plus the 800 of the pilot study, and the computer printed out the data of all individuals whose values lay outside these limits. It also printed data from individuals in whom similar limits for age at given puberty stage were exceeded or whose dates for ages or for examination were clearly impossible. This resulted in 21% of all individuals being listed. All these were scrutinized by hand, a job which took two people a month.

This erred on the safe but laborious side; 1st and 99th centiles limits are probably adequate for large-scale surveys.

In the approximately 10 000 cards examined as suspicious about 200 were found with one or more punching errors, an incidence of 0·3% in the total sample. These cases were repunched. A number of cards with errors could be reconstructed; for example, correct dates of birth or examination could be found and substituted for those written down wrongly. In addition one measurer (only) misunderstood the puberty rating 2 and her ratings had to be changed, after discussion, back to 1;

one team forgot for a period to add 10 cm to the sitting height measurement as required by the particular instrumentation used and this was added; in one team the words for 0 and for 6 were muddled and in many instances this led to easily recognizable absurdities which could be corrected. The reconstructed cards totalled 1·2%. These repunched and reconstructed cards were scrutinized again in a second run through the same edit programme; only a very small number were listed as erroneous.

On the remaining 19·5% of cards listed as suspicious, the "impossible" measurements were rejected, but the rest of the measurements were retained. To have eliminated the entire individual where a single measurement was wrong would have lost about 10% of subjects. However, cards in which three or more measurements were impossible were wholly eliminated, but these were few, amounting to 0·2% of the whole.

4. Discussion

Planning and execution

The survey was designed to be a national one suitable for constructing standards for clinical use in Cuba, of height and supine length, weight, sitting height and crown–rump length, skinfolds, head, arm, calf and thigh circumferences, bi-iliac and biacromial diameter, and foot length. It was also designed to sample enough geographical areas and occupational groups so that comparisons could be made between them for purposes of public health control and government health planning policy. All ages were considered to be important and steps had to be taken to secure proper samples, especially of pre-school children. This involved sampling the actual population in selected sectors of the island using a very large number of local enumerators, similar to those employed in census-taking. In Cuba the country-wide network of the Federation of Cuban Women, together with the Small Farmers' Association made this possible, and indeed very successful, since 96% of the sample of pre-school age attended for measurement. Since the survey had the co-operation of these organizations, schools were not sampled as is usual in this type of survey, but instead the same system was used at all ages. This resulted in an equally successful (96%) attendance of primary school children.

In the secondary school ages, however, the numbers of non-respondents became important; this was partly due to some boys and girls being away at boarding school, partly to boys on military service and partly to non-attendance for reasons never stated. At secondary school level it might be advisable to change to a sample frame of schools rather than children in geographical sectors. This would ensure adequate numbers from the boarding schools and it would very likely diminish the number of adolescent non-attenders, especially if the measuring station were set up in the school itself. There would, however, be a difficulty in obtaining representative age samples up to 16, which is school leaving age in Cuba. In particular, because the school system is organized by grades rather than age groups, problems would arise at the ages where a change in sampling procedure occurred. Above this age, samples of military service personnel could be used for boys. Further thought must be given to the question of finding the non-respondents. If secondary schools were selected perhaps the best design would be to have one or more all-male measuring teams and use these exclusively to measure boys aged 15 and upwards.

The high response figure depends not only on the network of enumerators, but also on continuous publicity, stressing the usefulness of the survey for purposes of health. A pilot study is highly important and led to several modifications of procedure. In addition it is very useful for training measurers and others. The design by which local head organizers were themselves the enumerators in the pilot project proved to be a useful one.

The second stage of the survey was made a year after the first, in order to obtain velocity standards; once measuring personnel have been trained and the organization run in, this is not difficult and yields important rewards. A 30% sample was chosen and only further experience will show whether this was sufficient. It is intended to keep the monitoring organization in being permanently (using many of the measurers on longitudinal and teaching projects) so that regular surveys may be made at 5-year intervals.

The parental measurements in the first survey were restricted to height and weight but ideally once parents are successfully obtained for measurement all the same measurements should be taken as were done on their children.

A departure from normal practice was made by dividing the age ranges for sampling according to the velocity of growth. Thus in the first year of life, when growth is fastest, three times as many children were measured, and in the age range 1·0 to 2·0 years twice as many as in the primary school years. Twice as many adolescents were measured at each year of age also. This structure is more logical than those usually used.

Quality controls

In any survey the number of measurers should be kept to the minimum administratively possible; there should be two trained measurers per team and all should be trained together by one acknowledged expert. Furthermore, quality control sessions should be held at fairly frequent intervals. This did indeed produce highly reliable measurements and eliminated differences between measuring teams which if they had existed would have prejudiced inter-regional comparisons (at the worst, the control sessions would have permitted bias to be measured and allowed for). The typical figure of the standard deviation of differences between repeated height measurements was 0·20 cm, which compares well with the value of 0·18 cm found for Mr. R. H. Whitehouse, the trainer of the measuring teams (Whitehouse, Tanner and Healy, 1974).

Editing the data

As a screen for detecting measurement error the 3rd and 97th centiles were used for each measurement and this resulted in 21% of the 56 000 individuals being listed for individual scrutiny. Probably a screen using the 1st and 99th centiles would be nearly as effective at finding the real errors, especially if combined with suitable bivariate listings such as stature with sitting height, one skinfold with another and arm circumference with calf circumference.

To eliminate the entire individual where any measurement was erroneous would have resulted in an unacceptable loss of nearly 10% of individuals. Eliminations of single measurements were highest for supine length and crown–rump length (1·1%) and lowest for triceps and subscapular skinfolds (0·1%). The other measurements varied between 0·2% and 0·5%.

Acknowledgments

We wish to acknowledge the great help given to the study by Mr. R. H. Whitehouse of the Department of Growth and Development, Institute of Child Health, London, who trained the measuring teams and advised on the ratings of the X-rays. We owe thanks to the following members of the steering committee of the study; Dr. D. Alonso, Dr. C. del Busto, M. Gancedo, Lic. L. Alvarez and Lic. A. Martinez. Dr. F. Rojas Ochoa and Dr. A. Tejeiro also gave invaluable help. We are indebted to the World Health Organisation for its continuing help and to the National Children's Bureau, London, for support.

Our gratitude goes to the members of the measuring teams who carried out their assignments over two years, often in difficult conditions, and we should especially like to thank all the children and their parents who took part in the study.

References

Goldstein, H. (1972). The construction of standards for measurements subject to growth. *Human Biology,* **33,** 255–261.

Hamill, P. V. V., Johnston, F. E., and Grams, W. (1970). *Height and weight of children: United States.* Rockville Md.: U.S. Department of Health, Education and Welfare; National Centre for Health Statistics, Series 11, No. 104.

Hamill, P. V. V., Johnston, F. E., and Lemeshow, S. (1973). *Height and weight of youths 12–17 years: United States.* Rockville Md.: U.S. Department of Health, Education and Welfare, National Center for Health Statistics, Series 11, No. 124.

Healy, M. J. R. (1974). Notes on the statistics of growth standards. *Annals of Human Biology,* **1,** 41–46.

Jordan, J. R. (1973). Los estudios sobre el crecimiento de niño en los paises en desarollo. *Carnets de l'Enfance,* **23,** 46–62.

New Zealand Department of Health (1971). *Physical Development of New Zealand school children 1969.* Department of Health Special Report Series No. 38. Wellington: Department of Health.

Tanner, J. M. (1952). The assessment of growth and development in children. *Archives of Disease in Childhood,* **27,** 10–33.

Tanner, J. M. (1966). Growth and physique in different populations of mankind. In: *The Biology of Human Adaptability,* ed. P. T. Baker and J. S. Weiner. Oxford: Clarendon Press.

Tanner, J. M., Whitehouse, R. H., and Takaishi, M. (1966). Standards from birth to maturity for height, weight, height velocity and weight velocity. British Children 1965. *Archives of Disease in Childhood,* **41,** 454–471; 613–635.

Tanner, J. M., Whitehouse, R. H., Marshall, W. A., Healy, M. J. R., and Goldstein, H. (1975). Assessment of skeletal maturity and prediction of adult height: TW 2 method. London: Academic Press.

Whitehouse, R. H., Tanner, J. M., and Healy, M. J. R. (1974). Diurnal variation in stature and sitting height in 12–14-year-old boys. *Annals of Human Biology,* **1,** 103–106.

van Wieringen, J. C., Wafelbakker, F., Verbrugge, H. P., and de Haas, J. H. (1971). *Growth Diagrams 1965, Netherlands.* Leiden: Netherlands Institute of Preventive Medicine.

Address correspondence to: Professor J. M. Tanner, Department of Growth and Development, Institute of Child Health, 30 Guilford Street, London WC1N 1EH.

Zusammenfassung. Ein geschichtetes in drei Abschnitten aufs Geratewolh genommene Beispiel von 50 360 Kindern im Alter von 0–19 aus der Bevölkerung von Cuba, wurde in einer gross angelegten Wachstum Studie während 1972–73 gemessen. 15 anthropometrische Messungen wurden genommen, Pubertät-Phasen und Menarche Zustand wurden in Betracht gezogen. Radiograph-Aufnahmen wurden an Handgelenken gemacht an 10% des Beispiels. Information wurde erlangt hinsichtlich des sozialen und Erziehungs-Status der Eltern, und die Grösse der Eltern wurde gemessen. Ein Jahr später wurden 30% des Beispiels wieder gemessen und 3% wieder Röntgen bestrahlt.

Die allgemeine Antwort-Ratio betrug 96% bei Kindern vor dem Schulalter und erste Schuljahre und etwas weniger später. Es wurden Sitzungen gehalten zur Qualitätskontrolle in welchen die neun messenden Gruppen Resultate verglichen. Keine bedeutenden Unterschiede wurden entdeckt zwischen den Gruppen, die in verschiedenen Teilen der Insel arbeiteten.

Unterschiede zwischen Duplikat-Grössen-Messungen von einzeiten Messern zeigten normale Abweichungen von annähernd 0,20 cm. Der Durchschnitt des individuellen Messers unterschied sich vom ganzen Durchschnitt aller Messer mit ±0,2 cm für Grösse, ±0,4 cm für Sitzhöhe und ±0,5 cm (7% vom durchschnitt) Oberarm Muskel-Hautfalten. Eine oder mehr Messungen lagen ausserhalb der 3.–97. Centile-Beschränkung in 21%. Individuums Genaue Prüfung der Rekords dieser Individuums resultierte in der Ausscheidung von Messungen in Mengen schwankend von 0,1% (Oberarmmuskel Hautfalte) bis 1,1% (Scheitel–Körperlänge). Probleme zur Planung und Ausführung von Wachstum-Uebersichten, so entworfen, dass sie eine nationale Norm setzen, werden beschrieben und Lösungen gegeben oder vorgeschlagen.

Résumé. Un échantillon aléatoir stratifié à 3 degrés de 50 360 enfants âggés de 0 à 19 ans, provenant de l'ensemble de Cuba, a été mesuré lors d'une étude de croissance menée sur une large échelle en 1972–73. On a procédé à 15 mesures anthropométriques, et les stades pubertaires et la présence d'un cycle menstruel ont été évalués. Des radiographies de la main et du poignet ont été prises sur 10 pour cent de l'échantillon, et des renseignements sur le statut social et le niveau d'instruction des parents ont été relevés, et la taille de ceux-ci a été mesurée. Un an plus tard, 30 pour cent de l'échantillon a été remesuré, et des radiographies ont été relaites sur 3 pour cent.

Le taux général de reponse a été de 96 pour cent des âges préscolaires et scolaires, et peu inférieur ensuite. Des sessions de contrôle de qualité ont montré peu de différences entre les techniques de mesure selon les différentes régions de l'île. Les différences entre mesures appariées ont un écart-type d'environ 0,20 cm pour la taille. Les différences maximales des observateurs individuels par rapport à l'ensemble de leurs collègues sont de ±0,2 cm pour la taille debout, ± 0,4 cm pour la taille assise, ±0,5 cm (7 pour cent de la moyenne) pour le pli tricipital.

Une mensuration ou plus se trouve en dehors des limites des 3ème–97ème centiles chez 21 pour cent des individus. Les dossiers de ces individus ont été examinés, ce qui a résulté en l'élimination de 0,1 pour cent (pli tricipital) à 1,1 pour cent (longueur occiput–coccyx) des mesures.

Les difficultés de la planification et de l'exécution des enquêtes de croissance sont décrites, et des solutions sont proposées ou suggérées.