Reliability and determinants of anogenital distance and penis dimensions in male newborns from Chiapas, México

S. Patricia Romano-Riquer[a], **Mauricio Hernández-Ávila**[a], **Beth C. Gladen**[b], **Lea A. Cupul-Uicab**[a], and **Matthew P. Longnecker**[b]

[a]Instituto Nacional de Salud Pública, 62508 Cuernavaca, Morelos, México

[b]National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina, USA

Summary

Development of the perineum as well as the external genitalia is determined by dihydrotestosterone, resulting in a greater anogenital distance (AGD) in males than females. In animal experiments with hormonally active agents, anogenital distance is used as a bioassay of fetal androgen action. Use of anogenital distance in human studies has been rare. Because anogenital distance has been an easy-to-measure, sensitive outcome in animal studies, we developed an anthropometric protocol for measurement of anogenital distance in human males. In this paper we describe the method for measurement of three anogenital distances, their reliability, and an assessment of predictors for each in the context of an epidemiological study. We compare the reliabilities and predictors to those for stretched penis length and penis width.

A cross-sectional study of 781 newly-delivered male infants was conducted in 2002–2003 in Chiapas, México. Replicate measures were obtained on nearly all subjects. The reliability of the measures of anogenital distance (0.82–0.91) were higher than for stretched penis length (0.78) and width (0.75). Birthweight and gestational length were more strongly related to anogenital distance than to penis length. Anogenital distance was not related to penis length (r = 0.03).

Our large study clearly shows that AGD can be measured well in newborn males, and that the measurements were more reliable than those of penis length. Whether AGD measures in humans relate to clinically important outcomes, however, remains to be determined, as does its utility as a measure of androgen action in epidemiological studies.

Introduction

Development of the perineum as well as the external genitalia is determined by dihydrotestosterone, resulting in a greater anogenital distance (AGD) in males than females.[1–3] In animal experiments with hormonally active agents, anogenital distance is used as a bioassay of fetal androgen action (reviewed in reference 3). Use of anogenital distance in human studies has been rare, especially among males.[3, 4] In human males, penile dimensions have traditionally been used as an indicator of androgenicity.[5] The reliability of anthropometric measures of penile dimension, however, may be compromised by the pliability of the structure.

In rodent studies, anogenital distance usually tracks through life, varies by dose of antiandrogen, and can be predictive of other androgen-responsive outcomes.[6] Because

Correspondence: Dr Matthew P. Longnecker, National Institute of Environmental Health Sciences, MD A3-05, PO Box 12233, Research Triangle Park, North Carolina 27709, USA; longnec1@niehs.nih.gov; tel 919 541-5118; fax 919 541-2511.

anogenital distance has been an easy-to-measure, sensitive outcome in animal studies, we developed an anthropometric protocol for measurement of anogenital distance in human males. Experience from our initial effort led to improvements in the technique, which we implemented in a large epidemiologic study.

In this paper we describe the method for measurement of three anogenital distances, their reliability, and an assessment of predictors for each. We compare the reliabilities to those for stretched penis length and penis width.

Methods

Subjects and Measurements

A cross-sectional study of newly-delivered male infants and their mothers was conducted in 2002–2003 in Tapachula, a city in the state of Chiapas, México. The subjects were recruited during the postpartum period, while hospitalized at either of the two participating hospitals: Hospital General de Zona # 1 del Instituto Mexicano del Seguro Social (General Hospital, Zone #1, Mexican Institute of Social Security) and Hospital General de Tapachula, de la Secretaría de Salud de Chiapas (Tapachula General Hospital, Chiapas Ministry of Health). If the eligibility criteria were met, the mothers were invited to participate and sign an informed consent form. The study protocol was approved by Institutional Review Boards at the Instituto Nacional de Salud Pública (National Institute of Public Health) in México and the National Institute of Environmental Health Sciences in the United States.

Exclusion criteria for the mother were age greater than 35 years; pre-eclampsia or pregnancy-related diabetes or hypertension; seizure disorders requiring daily medication; history of psychiatric, kidney, or cardiac disease; and non-speaker of Spanish. Infants were excluded if gestational age at delivery as estimated by the Capurro scale,[7] or the medical record (based on last menstrual period), was < 36 weeks, birthweight was < 2,500 g, pregnancy was not singleton, Apgar score at 5 minutes was 6, or child was admitted to the neonatal intensive care unit. Of the subjects who were invited to participate, 95% did so.

A questionnaire was administered to the mothers about socio-demographic characteristics, reproductive history, maternal health status, and various exposures. Measurements of weight and height were performed on the mothers and newborns. An electronic scale (Tanita©), readable to increments of 0.1 g, was used to measure the weight of the newborns. The measurements of newborn length were made with a portable infantometer, in increments of 1 mm. For the mothers, height was measured in cm using a SECA© portable stadiometer (Model 208) with a 1-mm reading precision. Weight was measured in kg using a Tanita© portable solar scale for adults (Model 1631) with a 0.2-kg reading precision for persons weighing up to 100 kg and a 0.4-kg reading precision for those weighing between 100 and 200 kg. In addition, we measured infant anogenital distance and penis size.

Technique for anogenital and penile measurement

Three measures of anogenital distance were taken: anterior base of penis to anus (AGD1), posterior base of penis to anus (AGD2), and posterior of scrotum to anus (ASD). In addition, we measured penile width (PW) and stretched penile length (PL). The anogenital and PW measurements were performed using Swiss Precision Cali Max Vernier calipers, from Bel-Art Products, Pequannock, NJ, USA. The calipers were specially modified to remove sharp edges. The calipers were read in increments of 1 mm. The calipers were sterilized before use on each subject.

The newborn was placed in the dorsal decubitus position on a table with one observer at each end of the infant. From the head end, the assistant held the infant's thighs in a flexed

position with her hands while placing her forearms on the baby's arms, thus minimizing movement of the newborn. The anthropometrist positioned herself in front of the baby. For the anogenital distances, she held the caliper in a sagital plane with her right hand, with the handle of the caliper pointing up, and tilted slightly towards the child's head so that the measuring surfaces were immediately adjacent to the landmarks being measured; the fixed measuring edge of the caliper was placed at the center of the anus. The penile measurements were done when the newborn's penis was flaccid. Each measurement was taken at least two times, 5 minutes apart; the first set of readings was recorded in the questionnaire, and after these were completed, the second set was taken and noted on a sheet that was attached to each subject's file. Thus, the results of the first set of measurements were not in sight when the second was done. On each of these two occasions, PL was usually measured in duplicate, yielding up to four recorded values.

Anterior base of penis to center of anus (anogenital distance, AGD1)—The caliper was opened until the moveable edge was at the anterior base of the penis (above the penis) (Figure 1A).

Posterior base of penis to center of anus (anogenital distance, AGD2)—The caliper was opened until the moveable edge was at the posterior base of the penis, where the penis meets the scrotum (Figure 1B).

Posterior of scrotum to center of anus (anoscrotal distance, ASD)—The first scrotal fold or change in texture (rugated skin), compared with the perineum, was identified. Once it was identified, the caliper was opened until the moveable edge was at the fold (Figure 1C). In some cases it was necessary to raise the scrotum (without stretching) to be able to identify the landmark. Note that this was the measure that Salazar et al.[3] called AGD. We label it as ASD here to make the nomenclature consistent with Swan et al.[4]

Penile width (PW)—The handle of the calipers was extended laterally in front of the anthropometrist, with the fixed edge of the caliper at the right lateral base of the penis. The measurement was performed by closing the moveable edge of the caliper until it came in contact with the left lateral base of the penis (Figure 1D).

Stretched penile length (PL)—To measure the length of the penis, a wooden tongue depressor was used. The ends of the tongue depressor were marked in different colors to identify the ends used for the first and the second measurement within each set of duplicates. One end of the tongue depressor was placed on the pubic symphisis above and immediately adjacent to the penis, and firm pressure was applied against the bone. The penis was held and gently stretched, and at the same time the foreskin was lowered until the urethral meatus was visible. With an extra-fine black indelible felt tip pen held perpendicular to the depressor, a mark was made on the tongue depressor corresponding to the tip of the penis (Figure 1E).

Anthropometric training and quality control—The anthropometrists received special training before measuring anogenital and penile dimensions on subjects included in the study. In addition, they received periodic retraining with quality control feedback. Ten training sessions were held during the data collection period, from once a month to once every 3 months. The average number of training sessions per anthropometrist was 2.6. In all, 22 anthropometrists participated. Of these, 12 worked for less than 2 months; their average number of training sessions was 1.3, and they measured 38 children among them. For the remaining 10 anthropometrists, the average number of training sessions was 4.1, and each

measured from 35 to 165 children. For analysis purposes, the 12 short-term anthropometrists were combined as "other".

Study Population

Of the 872 eligible subjects who participated, 781 had anogenital distances measured with the final protocol; the remainder of the infants were measured with an earlier version of the protocol that employed tongue depressors rather than calipers, and are not considered here.

All 781 infants had measurements of each of the three anogenital distances; in all but two (AGD1 and AGD2) or three (ASD) cases, duplicate measurements were taken. All infants also had PW available (in duplicate in all but three cases). PL was measured up to 4 times, twice with the first measurement of the other distances and twice with the second. For 541 infants, all four measurements were available. Of the remainder, 43 infants had two measurements on the first and one on the second occasion, 188 had two measurements on the first and none on the second, 3 had one on each occasion, 5 had only one on the first occasion, and 1 had no measurements.

Over 80% of children were examined before they were 6 hours old. With one exception, all examinations were conducted before 38 hours of age; the remaining examination was conducted 7 days after birth.

Statistical Analysis

We estimated the intra-anthropometrist and inter-anthropometrist reliabilities of the anogenital and penile measurements from variance components. For the anogenital distances and PW, we used PROC MIXED in SAS (SAS Institute Inc., Cary, NC, USA) to estimate the random variation from one child to another, the random variation from one anthropometrist to another, and the residual variation (which represents the variation from one measurement occasion to another). For PL, the replicate measurements on each occasion allowed estimation of child, anthropometrist, occasion, and residual (which represents the variation from one replicate to another) variation. The reliability (also called inter-rater reliability) is the ratio of the child variance component to the sum of all the variance components. For the intra-rater reliability, the numerator of the ratio is instead the sum of the child and anthropometrist variance components. The relationship between variance components and reliabilities has been described previously.[8]

All 781 subjects were included when examining reliability. When calculating descriptive statistics and examining potential predictors, subjects were excluded from a particular measurement if replicates differed by 30% or more; there were two such cases for ASD, one for PW, and three for PL. Means of replicates were used for descriptive statistics and examination of predictors.

Potential predictors of anogenital and penile dimensions were examined in linear models, again using SAS. Predictors included child size (weight and length, each linear), gestational age (<38, 38, 39, 40, 41+ weeks), maternal size (height and pre-pregnancy body mass index (BMI), both linear), maternal age (linear), parity (1, 2–3, 4+), hospital, rural/urban residence, education (0, 1–6, 7–9, 10–12, 13+ years), marital status (married, partner, separated/divorced/widowed, single), and income per capita (poorest, other poor, not poor). Anthropometrist, nested within hospital, was adjusted for as a random effect. In some cases, pre-pregnancy weight was not available, and weight at the first followup visit (from a second study not described here) was substituted. Per capita income categories (2002 versions) were defined separately for urban and rural areas.[9] In urban areas, those with monthly per capita income of 672.25 pesos or less were in the poorest category, comprising those who would have difficulty buying adequate food; of the remainder, those with income

of 1367.35 or less lacked adequate income for other human needs, so were still considered poor. In rural areas, the cutpoints were 494.77 and 946.49, respectively.

Results

All of the subjects were Latino; further ethnic characterization was not possible. The mothers were relatively young and most had had children previously (Table 1). They were about evenly split between the two hospitals; the Social Security hospital serves those who qualify for insurance through employment. More than half lived in urban areas, primarily Tapachula, with the remainder in surrounding villages. Only 32% had gone beyond ninth grade. Most were either married or living with their partner. Chiapas is one of the poorest states in México; most of the women were not only poor, but in the poorest category. The median height (152.4 cm) was above the regional median (149.9 cm) and close to the national median (153.0 cm) for women of reproductive age.[10] Median BMI (23.1 kg/m^2) was below the regional (24.3 kg/m^2) and national (25.0 kg/m^2) medians.[10] Given the eligibility criteria, the normal median birthweight among term babies was as expected. Smoking was uncommon, with only 12% reporting ever smoking, and the typical frequency being one or two cigarettes a month; only 8 women reported any smoking at all during pregnancy. Virtually all women reported at most occasional alcohol consumption, with even less during pregnancy.

The reliability of the measures were in general relatively high (Table 2). The (inter-rater) reliability, the proportion of the total variation in an anthropometric measurement due to true differences among subjects, can also be interpreted as the correlation between measurements obtained by two anthropometrists on the same subject. The reliability ranged from 0.76 for PL to 0.91 for AGD1, and was smaller for the penile than for the anogenital measurements. Compared to the variance among children, the anthropometrist-to-anthropometrist variability was greatest for PL (27% of the child variance for PL, as compared to 3% for AGD1), causing lower reliability. Again compared to the variance among children, the occasion/replicate variability was greatest for PW (20% of the child variance for PW, as compared to 7% for AGD1), also causing lower reliability. The intra-rater reliability, interpretable as the correlation of repeated measures done by the same anthropometrist, ranged from 0.83 for PW to 0.96 for PL.

Distributions of anogenital and penile measurements are shown in Table 3. AGD1 and AGD2 were highly correlated with each other (r = 0.81); correlations with ASD were lower (r = 0.37 and 0.26, respectively). The correlation of AGD1 with PW was 0.23; with PL was 0.03; and with penile volume was 0.20; correlations of the other two anogenital distances with penile measures were similar. PW and PL were not correlated with each other (r = −0.04).

The principal known predictor of anogenital distance is body size.[11] For example, the relationship of AGD1 with birthweight is shown in Figure 2, along with a smoothed curve. The relationship was essentially linear; the correlation was 0.37. Relationships with AGD2 and ASD were similar but weaker (correlations 0.27 and 0.24, respectively).

The results of considering multiple predictors of anogenital distances are shown in Table 4; four of the potential predictors influenced one or more of the three measures. AGD1 increased with both birthweight and gestational age. Urban and rural residents also differed in AGD1. The other two anogenital distances showed similar patterns with birthweight and gestational age, and also showed some differences between hospitals.

Predictors of penile measures were different (Table 5). PW increased with birthweight and decreased with maternal age. PL increased with maternal height and decreased with maternal age.

Discussion

Whether assessment of AGD in human males has utility over that of traditional measures of androgenization, such as stretched penis length, remains to be determined. Nonetheless, as the utility of AGD in humans is evaluated, the present data demonstrate that in the context of an epidemiological study the determinations can be done well by specially trained observers and that the reliability may be greater than for traditional measures such as stretched penis length. The reliability of AGD was similar to that of well accepted anthropometric measures of skinfold thickness (Table 6).[8, 12–16]

Body weight was an important determinant of AGD in our data, as would be expected from observations in animals.[11] Gestational age also was independently related to AGD1 and AGD2 (Table 4). The longer AGD among those with urban residence may reflect a subtle difference in the racial-ethnic background between the urban and rural subjects. Racial-ethnic differences in stretched penis length has been reported for other groups.[17, 18] Residence can also be a surrogate for a host of other lifestyle factors and exposures, making interpretation of our finding difficult.

To our knowledge, values of AGD1 and AGD2 in newborns have not previously been reported in the literature. ASD, however, was reported in another study of newborn Mexican males. The median in that study was 22 mm,[3] as compared with 19 mm in the present study. Birthweight and gestational age were similar in the two populations. As noted above, differences in the racial-ethnic compositions of the two Mexican populations could account for the difference, or it could be due to the improved method of measurement used in the present study. The posterior border of the scrotum is not always clearly identifiable.[19] This may contribute to the lower reliability of ASD compared with the other AGD measures, thus rendering it less suitable for comparisons between populations. Values of AGD1 and ASD in boys aged 2 to 30 months have recently been reported;[4] as expected, values in those older boys were larger than in newborns.

The mean stretched penis length of the infants in the present study (27.4 mm), was lower than reported in other populations. For example, among Indonesian newborns a mean of 28.6 was reported,[20] though values near 35 mm are more typical.[17, 18,21–23] The relatively low inter-rater reliability of the penis length measure means caution is needed when making comparisons among studies. However, in two previous comparisons of penis length across ethnic-racial groups,[17, 18] one observer examined all subjects, and differences were found. Therefore, true variation in mean penis length may exist across groups, and could be due to genetics or other effects not related to endocrine abnormalities.

In these data penis length was lower among those with older mothers, and was increased among those with taller mothers. In other studies of newborns, penis length has been examined in relation to only birthweight and length, and gestational age.[18, 20, 21, 23] In most of these studies direct associations were found, especially with gestational age.[18, 20, 21] Our subjects, however, had normal birthweight and were full term, which limited our ability to detect associations. The associations with maternal age and height are intriguing but these have not been examined in other studies; in the absence of confirming evidence we believe these findings merit little current attention.

Penis width has been studied less than penis length; in our study it was directly related to birthweight. In other studies of penis width, however, the measurement was done when the

penis was stretched,[22] or the exact technique was not clear.[18, 21] Although flaccid and stretched penis width are highly correlated,[24] our results may not be directly comparable to those in other studies.

Swan et al.[4] recently reported a correlation of 0.49 between AGD1 and penile volume (length times the square of width) in boys aged 2 to 30 months. In our data, the correlation of AGD1 with penile length was 0.03 and with penile volume was 0.20. The low correlation of penis length and AGD1 in our data suggests that they might be differentially responsive to androgenic stimuli. The larger correlation of AGD1 and penile volume in the study by Swan et al.[4] may have been due to the greater range of values being correlated. For example, the interquartile range for AGD1 in the Swan et al. study was 12.7 mm, as compared with 5.5 mm in our study. We also note that the results of the study by Swan et al.[4] suggested that anogenital distance in human males may be affected by early exposure to antiandrogens. Thus, detailed documentation of the technique, and characterization of the reliability of anogenital distance measurements may be of general interest.

In conclusion, our large study clearly shows that AGD can be measured well in newborn males, and that the measurements were more reliable than those of penis length. Whether AGD measures in humans relate to clinically important outcomes, however, remains to be determined, as does its utility as a measure of androgen action.

Acknowledgments

References

1. Bowman CJ, Barlow NJ, Turner KJ, Wallace DG, Foster PM. Effects of in utero exposure to finasteride on androgen-dependent reproductive development in the male rat. Toxicologic Sciences. 2003; 74:393–406.

2. Marty MS, Chapin RE, Parks LG, Thorsrud BA. Development and maturation of the male reproductive system. Birth Defects Research Part B, Developmental and Reproductive Toxicology. 2003; 68:125–136.

3. Salazar-Martinez E, Romano-Riquer P, Yañez-Marquez E, Longnecker MP, Hernandez-Avila M. Anogenital distance in human male and female newborns: a descriptive, cross-sectional study. Environmental Health: A Global Access Science Resource. 2004; 3:8.

4. Swan SH, Main KM, Liu F, Stewart SL, Kruse RL, Calafat AM, et al. Decrease in anogenital distance among male infants with prenatal phthalate exposure. Environmental Health Perspectives. 2005; 113:1056–1061. [PubMed: 16079079]

5. American Academy of Pediatrics Committee on Genetics. Evaluation of the newborn with developmental anomalies of the external genitalia. Pediatrics. 2000; 106:138–142. [PubMed: 10878165]

6. Gray LE Jr, Wolf C, Lambright C, Mann P, Price M, Cooper RL, et al. Administration of potentially antiandrogenic pesticides (procymidone, linuron, iprodione, chlozolinate, p,p′-DDE and ketaconazole) and toxic substances (dibutyl- and diethylhexyl phthalate, PCB 169, and ethane dimethane sulphonate) during sexual differentiation produces diverse profiles of reproductive malformations in the male rat. Toxicology and Industrial Health. 1999; 15:94–118. [PubMed: 10188194]

7. Capurro H, Konichezky S, Fonseca D, Caldeyro-Barcia R. A simplified method for diagnosis of gestational age in the newborn infant. Journal of Pediatrics. 1978; 93:120–122. [PubMed: 650322]

8. Ferrario M, Carpenter MA, Chambless LE. Reliability of body fat distribution measurements. The ARIC study baseline cohort result. International Journal of Obesity and Related Metabolic Disorders. 1995; 19:449–457. [PubMed: 8520633]

9. Technical note on the measurement of poverty based on the results of the National Survey of Household Income and Expenditures. 2002. (in Spanish). http://www.sedesol.gob.mx/subsecretarias/prospectiva/main_medicionpobreza.htm

10. González de Cossío, T.; Rivera Dommarco, J.; Shamah Levy, T.; Ramírez Silva, I.; Barquera Cervera, S.; Morales Ruán, MC., et al. Mujeres. In: Rivera Dommarco, J.; Shamah Levy, T.; Villalpando Hernández, S.; González de Cossío, T.; Hernández Prado, B.; Cuernavaca, Sepúlva J., editors. Encuesta Nacional de Nutrición 1999. Estado nutricio de niños y mujeres en México. México: Instituto Nacional de Salud Pública; 2001. p. 103-177.

11. Gallavan RH Jr, Holson JF, Stump DG, Knapp JF, Reynolds VL. Interpreting the toxicologic significance of alterations in anogenital distance: potential for confounding effects of progeny body weights. Reproductive Toxicology. 1999; 13:383–390. [PubMed: 10560587]

12. Jackson AS, Pollock ML, Gettman LR. Intertester reliability of selected skinfold and circumference measurements and percent fat estimates. Research Quarterly. 1978; 4:546–551. [PubMed: 741092]

13. Mueller WH, Malina RM. Relative reliability of circumferences and skinfolds as measures of body fat distribution. American Journal of Physical Anthropology. 1987; 72:437–439. [PubMed: 3605318]

14. Johnston, FE.; Hamill, PVV.; Lemeshow, S. Skinfold thickness of children 6–11 years, United States. Washington, DC: US Department of Health and Human Services; 1972. (Vital and Health Statistics. Series 11, No. 120)

15. Marks GC, Habicht JP, Mueller WH. Reliability, dependability, and precision of anthropometric measurements. The Second National Health and Nutrition Examination Survey 1976–1980. American Journal of Epidemiology. 1989; 130:578–587. [PubMed: 2764002]

16. Klipstein-Grobusch K, Georg T, Boeing H. Interviewer variability in anthropometric measurements and estimates of body composition. International Journal of Epidemiology. 1997; 26:S174–S180. [PubMed: 9126545]

17. Phillip M, De Boer C, Pilpel D, Karplus M, Sofer S. Clitoral and penile sizes of full term newborns in two different ethnic groups. Journal of Pediatric Endocrinology & Metabolism: JPEM. 1996; 9:175–179.

18. Cheng PK, Chanoine JP. Should the definition of micropenis vary according to ethnicity? Hormone Research. 2001; 55:278–281. [PubMed: 11805431]

19. Ballard JL, Khoury JC, Wedig K, Wang L, Eilers-Walsman BL, Lipp R. New Ballard Score, expanded to include extremely premature infants. Journal of Pediatrics. 1991; 119:417–423. [PubMed: 1880657]

20. Sutan-Assin M, Rukman J, Dahlan A. Penile dimensions of newborn infants. Paediatrica Indonesiana. 1989; 29:146–150. [PubMed: 2812828]

21. Feldman KW, Smith DW. Fetal phallic growth and penile standards for newborn male infants. Journal of Pediatrics. 1975; 86:395–398. [PubMed: 1113226]

22. Flatau E, Josefsberg Z, Reisner SH, Bialik O, Iaron Z. Letter: Penile size in the newborn infant. Journal of Pediatrics. 1975; 87:663–664. [PubMed: 1159603]

23. Lian WB, Lee WR, Ho LY. Penile length of newborns in Singapore. Journal of Pediatric Endocrinology & Metabolism: JPEM. 2000; 13:55–62.

24. Schonfeld WA, Beebe GW. Normal growth and variation in the male genitalia from birth to maturity. Journal of Urology. 1942; 48:759–777.
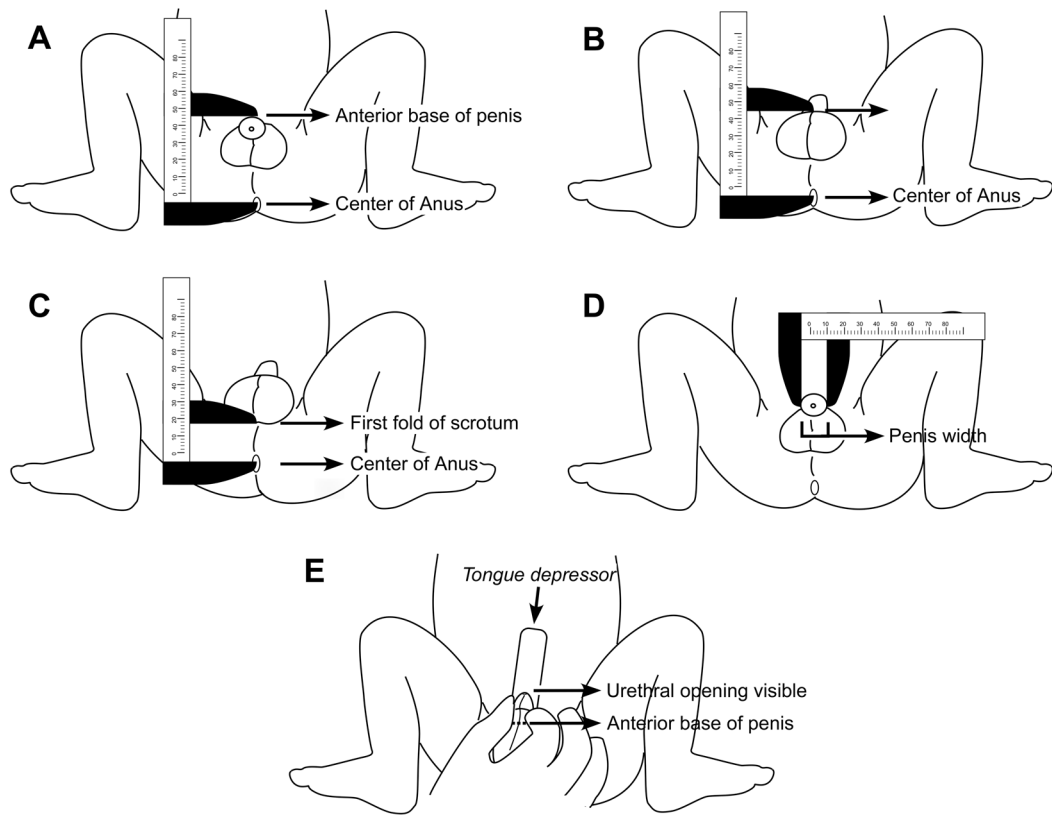
Figure 1.
A) Anterior of penis to center of anus distance (AGD1), B) Posterior of penis to center of anus distance (AGD2), C) Posterior of scrotum to center of anus distance (ASD), D) Penis width (PW), E) Stretched penis length (PL)
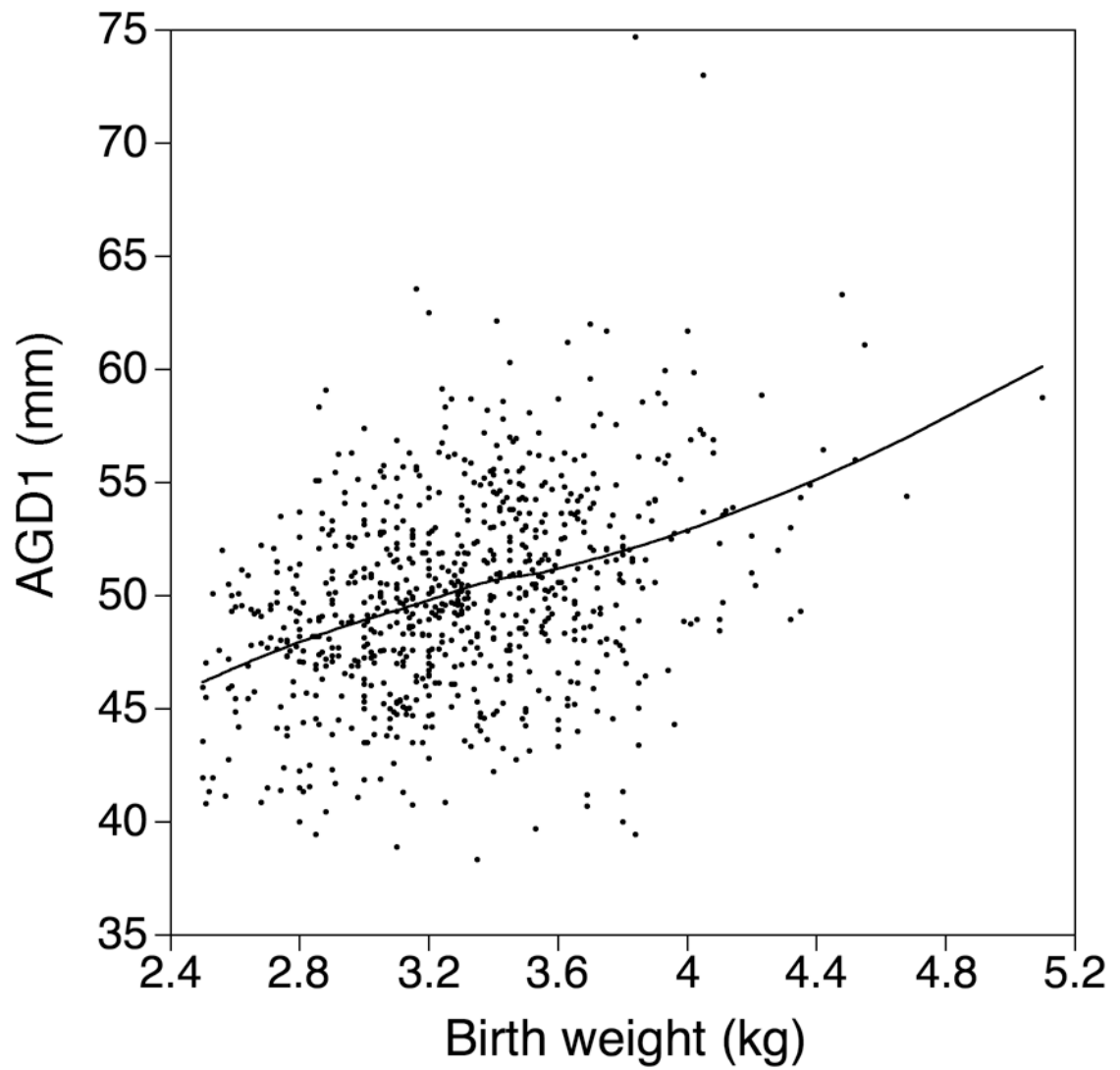
Figure 2.
Relationship of anogenital distance to birthweight. Smooth curve is a LOESS smoother.

Table 1

Characteristics of 781 mothers and their newborn sons, Tapachula, Chiapas, México

Characteristic	Percent or Median (25th, 75th Percentile)
Mothers	
Age (years)	23 (20, 27)
Parity (%)	
1	41.1
2–3	53.9
4	5.0
Hospital (%)	
Social Security Hospital	47.0
General Hospital	53.0
Residence (%)	
Rural	40.6
Urban	59.4
Education (%) [a]	
None	4.5
1–6 years	32.7
7–9 years	30.8
10–12 years	22.3
12+ years	9.7
Marital Status(%) [a]	
Married	44.8
Living with partner	44.7
Single	8.1
Separated, divorced, widowed	2.4
Income per capita (%) [a]	
Poorest	72.6
Poor, but not poorest	17.6
Not poor	9.8
Anthropometry	
Weight (kg, pre-pregnancy) [a]	54 (48, 63)
Height (cm)	152.4 (148.2, 156.0)
Body mass index (kg/m^2) [a]	23.1 (20.9, 26.3)
Infants	
Gestational age (weeks)	40 (39, 40)
Weight (g)	3280 (3020, 3540)
Length (cm)	49.6 (48.5, 51.0)

[a] Missing: education (n=1), marital status (n=2), weight/BMI (n=19), income per capita (n=18)

Table 2

Variance components and reliabilities of anogenital and penile measures in 781 male newborns, Tapachula, Chiapas, México[a]

Measure	Child variance component	Anthropometrist variance component	Occasion (/Replicate) variance component[b]	R inter	R intra
AGD1	18.39	0.59	1.28	0.91	0.93
AGD2	24.60	1.93	1.46	0.88	0.94
ASD	15.33	1.81	0.80	0.85	0.95
PW	0.93	0.10	0.18	0.77	0.83
PL [c]	17.06	4.64	0.11 / 0.51	0.76	0.96

[a] Abbreviations: AGD1, anogenital distance 1, anterior of penis to center of anus; AGD2, anogenital distance 2, posterior of penis to center of anus; ASD, anoscrotal distance, posterior of scrotum to center of anus; PW, penis width; PL, stretched penis length

[b] Measurements were taken on two occasions. For PL, there were usually two replicates per occasion; for the others, there was only one measurement per occasion, so that replicate variability within occasion could not be assessed separately.

[c] One subject had no measurements.

Table 3

Percentiles of anogenital and penile measurements of 781 male newborns, Tapachula, Chiapas, México[a]

Measure	Minimum	10th	25th	Median	75th	90th	Maximum
Anogenital distance (mm)							
AGD1	38.4	44.6	47.3	49.9	52.8	55.5	74.7
AGD2	19.8	39.1	42.4	45.4	48.7	51.8	70.4
ASD [b]	8.7	13.3	16.5	19.1	21.5	23.4	35.8
Penile dimensions (mm)							
PW [b]	6.3	9.3	9.8	10.5	11.2	11.8	15.7
PL [b]	15.5	21.6	24.3	27.1	30.4	33.2	42.9

[a]See list of abbreviations in first footnote of Table 2.

[b]Number missing: ASD (2 unreliable), PW (1 unreliable), PL (1 missing, 3 unreliable)

Table 4

Regression coefficients (± standard error) for anogenital distance (mm)[a], [b]

	AGD1		AGD2		ASD	
	Coefficient	p value	Coefficient	p value	Coefficient	p value
Birthweight (per kg)	3.73 ± 0.49	<0.001	3.27 ± 0.59	<0.001	1.90 ± 0.46	<0.001
Gestational age		0.015		0.018		0.48
36–37	−0.93 ± 0.83		−1.36 ± 0.99		−0.22 ± 0.77	
38	−0.49 ± 0.47		−0.38 ± 0.56		0.02 ± 0.43	
39	0.18 ± 0.36		0.04 ± 0.43		−0.43 ± 0.33	
40	Reference		Reference		Reference	
41+	2.04 ± 0.68		2.43 ± 0.82		0.68 ± 0.63	
Residence		0.065		0.12		0.16
Urban	0.58 ± 0.31		0.59 ± 0.38		0.41 ± 0.29	
Rural	Reference		Reference		Reference	
Hospital		0.40		0.066		0.064
Social Security	Reference		Reference		Reference	
General	0.61 ± 0.71		1.77 ± 0.86		−1.46 ± 0.70	

[a] See list of abbreviations given in first footnote to Table 2. The model results indicate, e.g., a) that for each additional kg of baby, AGD1 increased by 3.73 mm, after adjusting for the other covariates in the model, and b) compared with infants born after 40 completed weeks of gestation, those born after 36–37 weeks had an AGD1 that was 0.93 mm lower, after adjusting for the other covariates in the model.

[b] Models included all variables shown plus birth length, maternal height, pre-pregnancy BMI, age, parity, education, marital status, household income per capita and random effect of anthropometrist. For variables not shown, all p>.20 for all three anogenital measures. Due to missing information, the number of subjects was 749 for AGD1/AGD2 and 747 for ASD.

Table 5

Regression coefficients (± standard error) for penile measurements (mm)[a, b]

	PW		PL	
	Coefficient	**p value**	**Coefficient**	**p value**
Birthweight (per kg)	0.77 ± 0.12	<0.001	−0.41 ± 0.50	0.41
Maternal age (per 10 yr)	−0.22 ± 0.09	0.017	−1.06 ± 0.39	0.007
Maternal height (per 10 cm)	0.01 ± 0.06	0.92	0.73 ± 0.27	0.008
Parity		0.16		0.69
1	Reference		Reference	
2–3	0.03 ± 0.09		−0.01 ± 0.37	
4+	0.36 ± 0.19		0.63 ± 0.82	

[a]See list of abbreviations given in first footnote to Table 2. The model results indicate that, e.g., a) for each additional 10 years of mother's age, PL decreased by 1.06 mm, after adjusting for the other covariates in the model, and b) for each additional 10 cm of maternal height, PL increased by 0.73 mm, after adjusting for the other covariates in the model.

[b]Models included all variables shown plus birth length, gestational age, maternal pre-pregnancy BMI, hospital, rural/urban, education, marital status, household income per capita and random effect of anthropometrist. For variables not shown, all p>.20 for both penile measures. Due to missing information, the number of subjects was 748 for PW and 745 for PL.

Table 6

Inter- and intra-observer reliability from published results of selected studies of reliability of skinfold measurements[a]

Study	Descriptions of studies	Skinfold	Reliability (Inter)	Reliability (Intra)
Jackson 1978[14]	35 healthy males, relatively heavy, 3 measurements on different days by different testers	Chest	0.980	
		Axilla	0.978	
		Triceps	0.976	
		Subscapula	0.977	
		Abdomen	0.969	
		Suprailium	0.974	
		Thigh	0.971	
Mueller 1987[15] (Johnston 1974)[16]	HES cycle III; children 12–17 years old; repeat measurement taken 2–3 weeks later; in 77 cases, by same tester, in 224 cases, by different tester	Triceps	0.897	0.981
		Subscapular	0.940	0.914
		Midaxillary	0.939	0.878
		Suprailiac	0.921	0.954
		Medial calf	0.876	0.957
Marks 1989[17]	NHANES II: 95 males, 134 females; duplicate measurements on first occasion by two testers, repeated measurements at least 6 days later by one old and one new tester, gross discrepancies eliminated, analysis on log scale	Males:		
		Triceps	0.88	
		Subscapular	0.95	
		Females:		
		Triceps	0.81	
		Subscapular	0.93	
Ferrario, 1995[10]	ARIC Study, n=374; adults, 45–65 years; random sample selection one reps. Two interviewers	Triceps	0.91	0.99
		Subscapular	0.91	0.99
Klipstein, 1997[18]	10 healthy volunteers (4 men, 6 women),18 interviewers, measures taken for 3 different days	Biceps	0.86	0.98
		Triceps	0.97	0.99
		Subscapular	0.94	0.99
		Suprailiac	0.95	0.99

[a]This table was modeled after Table 2 in Klipstein *et al.*[18]