Anogenital distance and penile width measurements in The Infant Development and the Environment Study (TIDES): methods and predictors

Sheela Sathyanarayana[a,b,c,*], **Richard Grady**[c], **J B Redmon**[d], **Kristy Ivicek**[c], **Emily Barrett**[e], **Sarah Janssen**[f], **Ruby Nguyen**[d], **Shanna H Swan**[g], and **and the TIDES Study Team**

[a]Department of Pediatrics, University of Washington, Seattle, WA, USA

[b]Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, WA, USA

[c]Seattle Children's Research Institute, Seattle, WA, USA

[d]Department of Medicine, University of Minnesota, Minneapolis, MN, USA

[e]Department of Obstetrics and Gynecology, University of Rochester, Rochester, NY, USA

[f]Department of Occupational Medicine, Kaiser Permanente, San Francisco, CA, USA

[g]Department of Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Summary

Background—Anogenital distance (AGD), the distance from anus to genitals, is a sexually dimorphic measurement that has long been used to determine the sex of newborn rodent pups. Numerous animal studies have identified AGD as an androgen-sensitive endpoint, with a shortened distance in male offspring reflecting decreased androgen exposure *in utero*, and a longer distance in females reflecting increased androgen exposure *in utero*. Because AGD is a continuous measurement that can be assessed in all infants and adults, it can be included in epidemiologic studies as a predictive marker of reproductive health. We describe development of standardized measurement methods and predictors of AGD outcomes.

Objectives—One objective of this study was to develop a standardized protocol to measure anogenital distance (AGD) in newborns that would result in reliable and reproducible results. The second was to determine population statistics, demographics, and maternal predictors for AGD. We hypothesized that with proper training and oversight, our AGD measurement protocol would provide reliable and reproducible results.

[*]Corresponding author. University of Washington/Seattle Children's Research Institute, 2001 W. 8th Ave, Seattle, WA 98121,USA, E-mail address: sheela.sathyanarayana@seattlechildrens.org (S. Sathyanarayana).

Conflict of interest: None.

Methods—We examined infants born to 758 participants in The Infant Development and the Environment Study (TIDES) enrolled in four clinical centers from 2011-2013. We developed and implemented a detailed training protocol incorporating multiple quality control (QC) measures. In males, we measured anoscrotal distance (AGD_{AS}), anopenile distance (AGD_{AP}), and penile width (PW) and in females, anofourchette distance (AGD_{AF}) and anoclitoral distance (AGD_{AC}). A single examiner obtained three repetitions of all measurements, and a second examiner obtained independent measurements for 10% of infants. We assessed within-examiner variability and between-examiner variability.

Results—In the full TIDES cohort, including 758 mothers and children, significant predictors of AGD and PW included: age at exam, gestational age at birth, weight-for-length Z-score, maternal age and study center. In 371 males, the mean (SD) AGD_{AS}, AGD_{AP} and PW were 24.7 (4.5), 49.6 (5.9), and 10.8 (1.3) mm, respectively. In 387 females, the mean (SD) AGD_{AF} and AGD_{AC} were 16.0 (3.2) mm and 36.7 (3.8) mm, respectively. The intra-examiner and inter-examiner intraclass coefficients averaged over all subjects and examiners were between 0.89-0.92 and 0.69-0.84 respectively.

Discussion—We have sent training videos or conducted in-person training and shared protocols with six research groups. As more research groups perform these measurements, we hope this sharing of methods will allow for the comparison of measurements across studies.

Conclusions—Our study demonstrates that with appropriate training and quality control measures, AGD and PW measurements can be performed reliably and in an easily replicable manner. These measurements should be adjusted for appropriate covariates in epidemiologic analysis.

Keywords

Anogenital distance; Urology; Pediatric; Penile width

Introduction

Anogenital distance (AGD), the distance from anus to genitals, is a sexually dimorphic measurement that has long been used to determine the sex of newborn rodent pups [1]. Numerous animal studies have identified AGD as an androgen-sensitive endpoint, with a shortened distance in male offspring reflecting decreased androgen exposure *in utero*, and a longer distance in females reflecting increased androgen exposure *in utero* [2,3]. The US Environmental Protection Agency and National Toxicology Program include this measurement when evaluating the reproductive toxicity of chemicals. Several studies report shorter AGD in male infants with cryptorchidism and/or hypospadias compared with controls [4-7]. In females, a longer AGD has been associated with increased androgenization syndromes such as congenital adrenal hyperplasia [8]. In adult males, shorter AGD has been associated with infertility, reduced semen quality, and altered reproductive hormone concentrations [9-11], while in adult females, longer AGD has been associated with multifollicular ovaries [12].

Because AGD is a continuous measurement that can be assessed in all infants and adults, it can be included in epidemiologic studies as a predictive marker of reproductive health. To

be a useful measurement in human studies, methods should be standardized, validated, and easily reproducible across research groups. While several groups have measured AGD in infants, methods differ, making comparison across studies difficult. We describe the development of a standardized clinical protocol to measure AGD and penile width (PW) that produced results that were replicable and repeatable over time.

The Infant Development and the Environment Study (TIDES) was designed to examine prenatal exposure to phthalates and other environmental chemicals in relation to infant reproductive development. Using the protocol described here, we measured AGD and PW at birth and identified significant predictors of these genital measurements.

Methods

Study population

Recruitment of the TIDES cohort and its characteristics are discussed in Barrett *et al*. 2014 [13]. Briefly, TIDES recruited pregnant mothers and their children from four study centers from 2010 to 2012. All women who were less than < 13 weeks pregnant with a singleton pregnancy, English or Spanish speaking, aged 18 or over, with no serious threat to the pregnancy, and who planned to deliver at a study hospital were eligible for inclusion. After signing an informed consent, study participants were asked to give a urine sample in each trimester of pregnancy, a serum sample in the first trimester, and to complete a questionnaire in each trimester. Here, we report on 758 women who completed questionnaires, provided urine and serum samples, and delivered a healthy infant. Infants born prematurely were measured once the medical team deemed they were healthy and able to undergo the examination. All study centers received human subjects approval for conducting these study procedures.

Clinical methods development

A multi-disciplinary clinical team led by a pediatrician (Clinical Director), an endocrinologist, a pediatric urologist, a pediatric nurse, and the TIDES PI developed the standardized protocol. All AGD measurements are distances from the center of the anus to a genital landmark. In males the genital landmarks are 1) the anterior base of the penis where the penile tissue meets the pubic bone (AGD_{AP}), and 2) the base of the scrotum where the skin changes from rugated to smooth (AGD_{AS}) (Figure 1a). In females, the genital landmarks are 1) the anterior tip of the clitoral hood (AGD_{AC}), and 2) the base of the posterior fourchette where skin folds fuse (AGD_{AF}) (Figure 1b). For PW, we measured the diameter of the base of the penis while the penis was flaccid.

The clinical protocol was developed and refined during a study-wide training session and includes three components: inspection, positioning, and measurement. The infant is placed on a flat surface (e.g. a flat hospital bed or neonatal cart) in a supine position with the lower half of the body exposed. The examiner visually inspects the infant for any genital abnormalities. The baby's buttocks are placed at the edge of the flat surface to facilitate measurement. The infant is then positioned with the legs held back in a frog leg posture at a 60-90° angle from t he torso at the hip. The examiner then uses a disposable marker to mark

the mid-anus position used for all AGD measurements. The most important factors are ensuring that the infant is as still and calm as possible throughout the exam and that he is positioned properly. We made our measurements with dial calipers (Dial Vernier Caliper 6"mm) held with the numbers facing away from the examiner. Three independent measurements were made for each AGD measurement with the calipers closed and zeroed out between each measurement. PW, the easiest measurement to obtain in an infant, was measured last.

Quality control measures

Our team addressed challenges to obtaining accurate and reproducible AGD measurements during a 2-day standardized training. The core components of the standardized training included: 1) development of a training manual with clinical protocols and references, 2) didactic teaching on infant genital anatomical landmarks with demonstrations of how to perform measurements, 3) repeat measurements on male and female infants by all study examiners in two sessions, the first followed by feedback/discussion and the second with different infants on the second day, 4) calculation of the measures of between- and within-examiner variability in real time with discussion of results, and 5) debriefing of examiners regarding their measurements and discussion of any needed protocol changes. In addition, we created a video for examiners who might join the study and to train those in future studies. This video is available for any study teams that would like to use it for research purposes. Study teams may contact the first author to obtain a copy of the video protocol.

We used several quality control measures during the study to ensure accurate and reliable AGD measurements. We obtained repeat measures by a second examiner on at least every 10th infant of each gender to monitor between-examiner reliability. All genital measurements were repeated three times. We examined within-examiner variability in two ways. We calculated the intra-examiner inter-class coefficient (ICC). In addition, a second examiner repeated all measurements on 108 (14%) of infants to examine between-examiner variability, which was measured by the inter-examiner ICC.

Statistical analysis

We examined descriptive characteristics within male and female infants separately for the four measures of AGD and PW. We then examined a number of potential predictors of these measures. AGD is known to vary with both age and body size [14,15] but because these predictors are highly correlated, we chose not to include them both in our models but rather to find a measure of body size that is both independent of age and predictive of AGD. That measure is weight-for length which, among alternative measures (such as weight-for-age, and BMI-for-age) explains the greatest proportion of the variance across all AGD measurements. We used World Health Organization (WHO) size-for-age norms because Center for Disease Control (CDC) norms are not recommended for very young infants [16].

After examining univariate predictors of AGD and PW, we used general linear models to examine covariates in relation to study outcomes. Based on prior analyses, we initially considered age at exam, gestational age at birth, weight-for-length Z score, maternal age, race, smoking, and study center as potential covariates. For center effects, we used UMN as

a reference group. All were retained in final models except smoking and race/ethnicity. Excluding these variables resulted in only small changes in beta-coefficients (<10%) for any retained covariates.

Results

Among the 758 TIDES mothers included in these analyses, the mean age was 31.1 years and the majority (66%) were white (Table 1). Over 70% of women had some graduate education. Most infants were born at term (mean, 39 gestational weeks) with normal birth weight. Fifty-one percent of births were to female infants.

Among TIDES infants mean (SD) male AGD_{AS}, AGD_{AP}, and PW were 24.7mm (4.5), 49.6mm (5.9), and 10.8 mm (1.3), respectively. The mean female AGD_{AF} and AGD_{AC} were 16.0mm (3.2), and 36.7mm (3.9), respectively (Table 2). In univariate analyses, weight and length at exam were significantly associated with all measures of AGD and PW (results not shown). Gestational age at birth was significantly associated with increased AGD_{AS} and AGD_{AP} in males and increased AGD_{AC} in females, while birthweight was associated with all measurements except PW.

Among the body size measures we considered, weight for length Z-score explained the greatest proportion of the variance for our multivariable models predicting AGD and PW. Age at exam was positively associated with all AGD and PW measurements (Table 3). Maternal age was directly related to both measures of AGD in boys, but not to PW. In girls, there was a suggestion of an inverse relationship between maternal age and AGD. The greatest center effects were seen with AGD_{AF}.

Six examiners measured 70.7% of infants, with a total of 20 examiners across sites. Within-examiner variability was somewhat greater for the male AGD measurements than the female, with intra-examiner ICCs ranging from 0.92 for AGD_{AF} and AGD_{AC} to 0.95 for AGD_{AS} (Table 4). For PW, the ICC was somewhat lower (0.89). The inter-examiner ICCs exhibited a similar pattern with the highest ICC seen for AGD_{AS} (0.84) and the smallest for PW (0.69).

Discussion

We describe the successful development and implementation of a clinical protocol to measure infant AGD and PW. Based on our quality control results, we believe that AGD and PW are clinical anatomic measurements that can be performed reliably by trained examiners. Several infant and maternal characteristics were significant predictors of AGD and PW measures and should be considered in epidemiologic statistical analyses including age at exam, body size (modeled as weight-for length Z-score here), gestational age at birth, maternal age, and study center. In the future, other study investigators and research teams will be able to use our clinical protocols, training materials, and model building data to effectively integrate AGD and PW measurements into epidemiologic research studies.

Several studies have found that infant anthropometric characteristics are important predictors of AGD, but few studies describe how best to model AGD for epidemiologic

statistical analysis. In our population of infants, examined at median age 1 day, we found that age at exam, gestational age at birth, and a measure of body size independent of age (weight-for-length Z-score) as well as maternal age are important predictors of measures of AGD and PW. Although 82.6% of TIDES infants were examined within 3 days of birth, which should reduce the influence of age at exam, it still remained one of the strongest predictors of AGD. Despite our intensive quality control methods, some between center-differences remained and study center was a significant covariate in our models.

In a recent newborn cohort from Greece and Spain (n=352) researchers reported female AGD_{AC} and AGD_{AF} mean measurements of 34.9mm and 14.1mm. These were slightly smaller than our TIDES girls, but the mean weights of our female infants were also higher [17]. Male AGD_{AP}, AGD_{AS}, and PW in this cohort were very similar to those seen in our TIDES boys [17]. In a cohort of over 500 infants in England, Thankamony reports only AGD_{AS} was 19.8mm and AGD_{AF} was 9.1mm, which were both smaller than those in the TIDES study [18]. This study used a somewhat different measurement, as described by Salazar-Martinez [19] in a study of 87 infants in Mexico in which measurements were smaller than those we report here. Using the Salazar-Martinez/Thankamony method infants are placed with feet and legs touching the exam table, while our protocol requires that infants' legs are held in a frog leg or lithotomy type position. Although methods are described in detail in each study, the angle at which the legs are placed relative to the hips is not described, and we found this to be the biggest determinant of between-study differences in measurements. Based on this discrepancy, we post-hoc measured AGD_{AS} and AGD_{AP} using both measurement techniques in 17 male infants. Anecdotally, we found that infants measured using the Salazar-Martinez/Thankamony technique had, on average, a 3mm smaller AGD_{AS} and 5mm smaller AGD_{AP} compared with our frog leg/lithotomy position. Both measurement techniques may be reliable for research studies, but systematic comparison of measurement methodologies is necessary to confirm this.

Key components of TIDES that allowed for successful AGD and PW measurements included: standardized positioning of infants in a quiet, non-moving state, careful planning of the standardized training, on-going dialogue with feedback on measurements, and consistent quality control tracking of measurements. The standardized training allowed for discussion of methods to position and quiet the infant, appropriate identification of anatomic markers, real time feedback on measurements, and comparisons of measurements between examiners. Monthly phone meetings and re-training at the second annual meeting allowed for feedback from all examiners, which led to consistent measurements over time as the study progressed.

The ICC for each measurement of AGD and PW was 0.89 or above. We were less successful in controlling between-individual and between-center differences. Study centers and staff reported that the AGD_{AF} was the most difficult measurement to conduct and replicate, and this measurement showed the largest center effects. To address this issue, we focused on the female genital anatomy at the second annual meeting retraining with photographs of difficult exams to discuss appropriate landmarks. Although examiners reported little difficulty with the male AGD_{AS} measurement because the landmark for where the scrotal skin meets smooth skin was easily visualized in the majority of infants, the inter-

examiner ICC was slightly higher for AGD_{AS} compared with AGD_{AP}, which examiners reported to be a bit more difficult because of increased skin/suprapubic fat pad between the penis and the abdomen.

In the past two decades, AGD has been measured in numerous human studies and its potential utility as a clinical marker has evolved. Several studies report that AGD is shorter in kids with hypospadias, cryptorchidism and frank genitourinary abnormalities, and in adult men with abnomal sperm parameters, enlarged prostate, and reduced fertility compared with healthy controls. To further establish its clinical utility, much larger scale studies should confirm that children and men with disorders have shorter AGDs when compared with age- and body size-matched controls. Then, the measure should be trialed in the clinical setting. It may have its greatest utility in urologic clinics where the measurement may predict future fertility or genitourinary outcomes.

Conclusions

Our data suggest that AGD and PW measurements can be performed easily and reliably in clinical research settings. We have sent training videos or conducted in-person training and shared protocols with six research groups. We have been responsive to questions from researchers on how to perform these measurements accurately and reliably, and each research group has reported successful measurements with excellent ICC. As more research groups perform these measurements, we hope this sharing of methods will allow for comparison of measurements across studies. Epidemiologic analyses should examine multiple potential predictors of AGD including infants' body size, age at exam, gestational age at birth, between-center/-examiner differences as well as maternal characteristics, such as age and, potentially, smoking and race/ethnicity depending on the study population.

Acknowledgments

References

1. Griffith JQ. The breeding of the rat. The rat in laboratory investigation. 1942:1–18.

2. Wolf C Jr, Lambright C, Mann P, Price M, Cooper RL, Ostby J, Gray LE Jr. Administration of potentially antiandrogenic pesticides (procymidone, linuron, iprodione, chlozolinate, p,p′-DDE, and ketoconazole) 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. Toxicol Ind Health. 1999; 15(1-2):94–118. [PubMed: 10188194]

3. Foster PM. Disruption of reproductive development in male rat offspring following in utero exposure to phthalate esters. Int J Androl. 2006; 29(1):140–7. discussion 181-5. [PubMed: 16102138]

4. Hsieh MH, Breyer BN, Eisenberg ML, Baskin LS. Associations among hypospadias, cryptorchidism, anogenital distance, and endocrine disruption. Curr Urol Rep. 2008; 9(2):137–42. [PubMed: 18419998]

5. Hsieh MH, Eisenberg ML, Hittelman AB, Wilson JM, Tasian GE, Baskin LS. Caucasian male infants and boys with hypospadias exhibit reduced anogenital distance. Hum Reprod. 2012; 27(6): 1577–80. [PubMed: 22434852]

6. Jain VG, Singal AK. Shorter anogenital distance correlates with undescended testis: a detailed genital anthropometric analysis in human newborns. Hum Reprod. 2013; 28(9):2343–9. [PubMed: 23838161]

7. Thankamony A, Lek N, Carroll D, Williams M, Dunger DB, Acerini CL, et al. Anogenital distance and penile length in infants with hypospadias or cryptorchidism: comparison with normative data. Environ Health Perspect. 2014; 122(2):207–11. [PubMed: 24316680]

8. Callegari C, Everett S, Ross M, Brasel JA. Anogenital ratio: measure of fetal virilization in premature and full-term newborn infants. J Pediatr. 1987; 111(2):240–3. [PubMed: 3612396]

9. Eisenberg ML, Jensen TK, Walters RC, Skakkebaek NE, Lipshultz LI. The relationship between anogenital distance and reproductive hormone levels in adult men. J Urol. 2012; 187(2):594–8. [PubMed: 22177168]

10. Eisenberg ML, Hsieh MH, Walters RC, Krasnow R, Lipshultz LI. The relationship between anogenital distance, fatherhood, and fertility in adult men. PLoS One. 2011; 6(5):e18973. [PubMed: 21589916]

11. Mendiola J, Stahlhut RW, Jørgensen N, Liu F, Swan SH. Shorter anogenital distance predicts poorer semen quality in young men in Rochester, New York. Environ Health Perspect. 2011; 119(7):958–63. [PubMed: 21377950]

12. Mendiola J, Roca M, Mínguez-Alarcón L, Mira-Escolano MP, López-Espín JJ, Barrett ES, et al. Anogenital distance is related to ovarian follicular number in young Spanish women: a cross-sectional study. Environ Health. 2012; 11:90. [PubMed: 23217457]

13. Barrett ES, Sathyanarayana S, Janssen S, Redmon JB, Nguyen RHN, Kobrosly R, Swan SH. and the TIDES study team. Environmental health attitues and behaviors: findings from a large pregnancy cohort study. Eur J Obstet Gynecol Reprod Biol. 2014; 176:119–25. [PubMed: 24647207]

14. 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. Environ Health Perspect. 2005; 113(8):1056–61. [PubMed: 16079079]

15. Sathyanarayana S, Beard L, Zhou C, Grady R. Measurement and correlates of ano-genital distance in healthy, newborn infants. Int J Androl. 2010; 33(2):317–23. [PubMed: 20132349]

16. Centers for Disease Control and Prevention, N.C.f.H.S. [cited 2014 11/18/14] WHO Growth Standards are Recommended for Use in the US for Infants and Children 0-2 Years of Age. 2010. http://www.cdc.gov/growthcharts/who_charts.htm

17. Papadopoulou E, Vafeiadi M, Agramunt S, Basagaña X, Mathianaki K, Karakosta P, et al. Anogenital distances in newborns and children from Spain and Greece: predictors, tracking and reliability. Paediatr Perinat Epidemiol. 2013; 27(1):89–99. [PubMed: 23215716]

18. Thankamony A, Ong K, Dunger DB, Acerini CL, Hughes IA. Anogenital Distance from Birth to Two Years: A Population Study. Environ Health Perspect. 2009; 117(11):1786–90. [PubMed: 20049133]

19. Salazar-Martinez E, Romano-Riquer P, Yanez-Marquez E, Longnecker MP, Hernandez-Avila M. Anogenital distance in human male and female newborns: a descriptive, cross-sectional study. Environ Health. 2004; 3(1):8. [PubMed: 15363098]
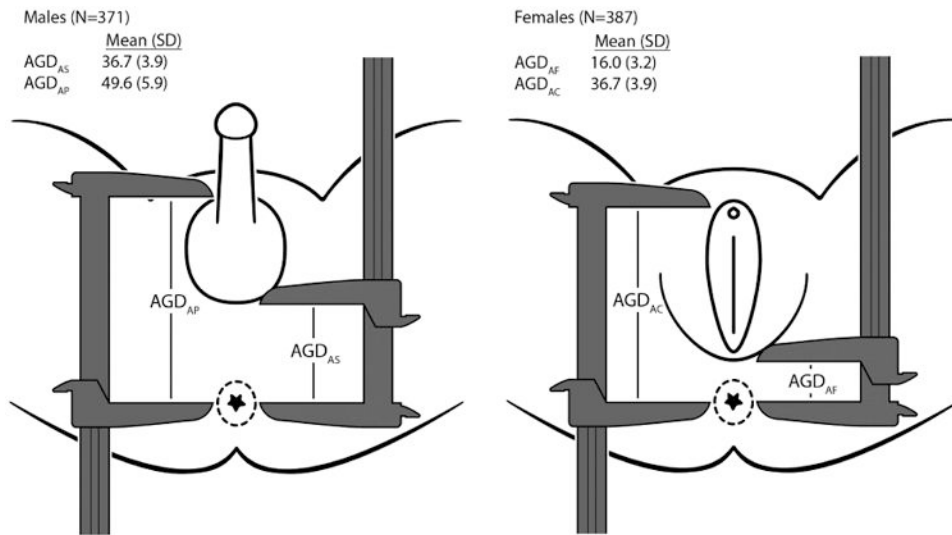
Figure 1. Anogenital Distance Measurements and Values Within the TIDES cohort

Table 1

Characteristics of pregnant women and infants within TIDES cohort (*n*=758)

Characteristic	No.[a]	%	Mean (SD)
Maternal age (years)			31.1 (5.5)
20	23	3.03	
21-30	275	36.28	
31-40	423	55.80	
>40	33	4.35	
Study center			
San Francisco, CA	187	24.67	
Minneapolis, MN	206	27.18	
Rochester, NY	213	28.10	
Seattle, WA	152	20.05	
Race/ethnicity			
White	500	65.96	
Hispanic/Latina	87	11.48	
Asian	67	8.84	
Black/African American	48	6.33	
Other	56	7.39	
Education			
High school or some college	57	7.52	
Tech. school or college/tech. school graduate	135	17.81	
Some graduate school	235	31.00	
Graduate degree	323	42.61	
Gestational age at birth (weeks)			39.3 (1.8)
37 weeks	691	91.16	
<37 weeks	67	8.84	
Infant sex			
Male	371	48.94	
Female	387	51.06	
Infant age at exam (days)			6.4 (15.5)
3 days	626	82.59	
> 3 days	132	17.41	
Infant birthweight (g)			3356 (558)
2500	41	5.42	
>2500 and 4000	633	83.73	
>4000	82	10.85	

[a] Some rows sum to <758 because of missing values.

Table 2

Anogenital distance and penile width (mm) in all TIDES cohort

	Mean (SD)	25%	50%	75%
Males (*N*=371)				
AGD$_{AS}$	24.7 (4.5)	21.7	24.3	27.2
AGD$_{AP}$	49.6 (5.9)	45.6	49.0	53.2
PW	10.8 (1.3)	9.9	10.8	11.6
Females (*N*=387)				
AGD$_{AF}$	16.0 (3.2)	13.9	15.7	18.2
AGD$_{AC}$	36.7 (3.9)	34.1	36.8	39.2

Table 3

Regression coefficients (β) and 95% confidence intervals (CI) from multivariable models of anogenital distance and penile width in TIDES infants

	Males = 369			Females = 376		
	AGD_{AS}	AGD_{AP}	Penile width	AGD_{AF}	AGD_{AC}	
	β (CI)	β (CI)	β (CI)	β (CI)	β (CI)	
Age at exam (days)	0.12 (-1.04, 0.14)	0.20 (0.17, 0.23)	0.04 (0.03, 0.04)	0.03 (0.01, 0.05)	0.07 (0.04, 0.09)	
Gestational age (weeks)	0.59 (0.35, 0.84)	0.92 (0.63, 1.21)	0.07 (-0.01, 0.14)	0.17 (-0.01, 0.35)	0.54 (0.32, 0.76)	
Weight-for-length Z-Score	0.58 (0.24, 0.91)	0.83 (0.43, 1.23)	0.10 (-0.01, 0.20)	0.22 (-0.01, 0.45)	0.62 (0.33, 0.91)	
Maternal age (years)	0.13 (0.03, 0.22)	0.18 (0.07, 0.29)	-0.02 (-0.05, 0.01)	-0.07 (0.13, -0.01)	-0.05 (-0.12, 0.03)	
Study center (relative to UMN)						
UCSF	0.14 (-1.04, 1.32)	-1.07 (-2.47, 0.34)	0.11 (0.24, 0.46) -	1.71 (-2.56, -0.87)	0.32 (-0.74, 1.37)	
URMC	-0.05 (-1.21, 1.12)	-1.81 (-3.20, -0.42)	0.18 (-0.16, 0.53) -	1.09 (-1.92, -0.25)	-0.81 (-1.85, 0.24)	
UW/SCH	1.08 (-0.13, 2.28)	-2.11 (-3.55, -0.67)	-0.37 (0.73, -0.02)	1.92 (1.06, 2.79)	-0.29 (-1.38, 0.81)	

Table 4

Measures of within- and between-examiner variability across all centers, examiners, and study visits

	Males			Females	
	AGD_{AS}	AGD_{AP}	Penile width	AGD_{AF}	AGD_{AC}
Intra-examiner ICC	.95 (.87, .99)	.94 (.85, .99)	.89 (.71, .98)	.92 (.78, .98)	.92 (.78, .99)
Inter-examiner ICC	.84 (.77, .88)	.82 (.75, .87)	.69 (.60, .77)	.73 (.64, .80)	.79 (.71, .85)

ICC = Intraclass correlation coefficient.