Decrease in Anogenital Distance among Male Infants with Prenatal Phthalate Exposure

*Shanna H. Swan,[1] Katharina M. Main,[2] Fan Liu,[3] Sara L. Stewart,[3] Robin L. Kruse,[3] Antonia M. Calafat,[4] Catherine S. Mao,[5] J. Bruce Redmon,[6] Christine L. Ternand,[7] Shannon Sullivan,[8] J. Lynn Teague,[9] and the Study for Future Families Research Team**

[1]Department of Obstetrics and Gynecology, University of Rochester, Rochester, New York, USA; [2]Department of Growth and Reproduction, University of Copenhagen, Copenhagen, Denmark; [3]Department of Family and Community Medicine, University of Missouri-Columbia, Columbia, Missouri, USA; [4]National Center for Environmental Health, Centers for Disease Control and Prevention, Division of Laboratory Sciences, Atlanta, Georgia, USA; [5]Department of Pediatrics, Division of Endocrinology, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Los Angeles, California, USA; Departments of [6]Pediatrics and [7]Medicine, University of Minnesota Medical School, Minneapolis, Minnesota, USA; [8]Department of Pediatrics, University of Iowa, Iowa City, Iowa, USA; [9]Departments of Surgery (Urology) and Child Health, University of Missouri-Columbia, Columbia, Missouri, USA

Prenatal phthalate exposure impairs testicular function and shortens anogenital distance (AGD) in male rodents. We present data from the first study to examine AGD and other genital measurements in relation to prenatal phthalate exposure in humans. A standardized measure of AGD was obtained in 134 boys 2–36 months of age. AGD was significantly correlated with penile volume (R = 0.27, p = 0.001) and the proportion of boys with incomplete testicular descent (R = 0.20, p = 0.02). We defined the anogenital index (AGI) as AGD divided by weight at examination [AGI = AGD/weight (mm/kg)] and calculated the age-adjusted AGI by regression analysis. We examined nine phthalate monoester metabolites, measured in prenatal urine samples, as predictors of age-adjusted AGI in regression and categorical analyses that included all participants with prenatal urine samples (n = 85). Urinary concentrations of four phthalate metabolites [monoethyl phthalate (MEP), mono-*n*-butyl phthalate (MBP), monobenzyl phthalate (MBzP), and monoisobutyl phthalate (MiBP)] were inversely related to AGI. After adjusting for age at examination, p-values for regression coefficients ranged from 0.007 to 0.097. Comparing boys with prenatal MBP concentration in the highest quartile with those in the lowest quartile, the odds ratio for a shorter than expected AGI was 10.2 (95% confidence interval, 2.5 to 42.2). The corresponding odds ratios for MEP, MBzP, and MiBP were 4.7, 3.8, and 9.1, respectively (all p-values < 0.05). We defined a summary phthalate score to quantify joint exposure to these four phthalate metabolites. The age-adjusted AGI decreased significantly with increasing phthalate score (p-value for slope = 0.009). The associations between male genital development and phthalate exposure seen here are consistent with the phthalate-related syndrome of incomplete virilization that has been reported in prenatally exposed rodents. The median concentrations of phthalate metabolites that are associated with short AGI and incomplete testicular descent are below those found in one-quarter of the female population of the United States, based on a nationwide sample. These data support the hypothesis that prenatal phthalate exposure at environmental levels can adversely affect male reproductive development in humans. *Key words:* anogenital distance, benzylbutyl phthalate, dibutyl phthalate, diethyl phthalate, monobenzyl phthalate, monoethyl phthalate, monoisobutyl phthalate, mono-*n*-butyl phthalate, phthalates, prenatal exposure. *Environ Health Perspect* 113:1056–1061 (2005). doi:10.1289/ehp.8100 available via *http://dx.doi.org/* [Online 27 May 2005]

Diesters of phthalic acid, commonly referred to as phthalates, are widely used in industry and commerce; they are used in personal care products (e.g., makeup, shampoo, and soaps), plastics, paints, and some pesticide formulations. Consistent toxicologic evidence indicates association between several of these phthalate esters and reproductive effects. In particular, dibutyl phthalate (DBP), benzylbutyl phthalate (BzBP), di-2-ethylhexyl phthalate (DEHP), and di-isononyl phthalate have been shown to disrupt reproductive tract development in male rodents in an antiandrogenic manner (Parks et al. 2000). Recent studies have reported significant reductions in anogenital distance (AGD) in Sprague-Dawley rats after prenatal exposure at high doses to BzBP (Nagao et al. 2000; Tyl et al. 2004), DBP (Barlow and Foster 2003; Foster et al. 2000), and DEHP (Gray et al. 2000; Parks et al. 2000).

Despite the growing body of literature on phthalate reproductive toxicity and data demonstrating extensive human exposure (Silva et al. 2004a), few studies have examined the effects of these chemicals on human reproductive development. Colón et al. (2000) reported elevated levels of several phthalates [including diethyl phthalate (DEP), DBP, and DEHP] in serum samples from young girls with premature breast development. However, the timing of exposure was unknown and high exposure levels may have reflected phthalate contamination of serum samples (McKee and Toxicology Research Task Group 2004). Until recently, the only study of humans to evaluate phthalate exposure and male reproductive toxicity measured phthalate diesters in semen. As with the Colón et al. study, contamination from diesters in laboratory equipment could not be excluded (Murature et al. 1987).

More recent studies have examined phthalate monoester metabolites in urine. Because urinary metabolites are not likely to be present as the result of contamination, these studies avoid this potential source of measurement error. Duty et al. (2003a) reported dose–response relationships between tertiles of monobutyl phthalate and sperm motility and sperm concentration, and between tertiles of monobenzyl phthalate (MBzP) and sperm concentration. They also reported inverse dose–response relationships between monoethyl phthalate (MEP) and sperm DNA damage

Address correspondence to S.H. Swan, University of Rochester, Department of Obstetrics and Gynecology, School of Medicine and Dentistry, 601 Elmwood Ave., Box 668, Rochester, NY 14642-8668 USA. Telephone: (585) 273-3521. Fax: (585) 275-7366. E-mail: shanna_swan@urmc.rochester.edu

*The Study for Future Families Research Team included, from the University of Missouri-Columbia: E.Z. Drobnis, B.S. Carter, D. Kelly, and T.M. Simmons. Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center: C. Wang, L. Lumbreras, S. Villanueva, M. Diaz-Romero, M.B. Lomeli, and E. Otero-Salazar. Cedars-Sinai Medical Center: C. Hobel and B. Brock. University of Minnesota: C. Kwong and A. Muehlen. University of Iowa: A. Sparks, A. Wolf, J. Whitham, M. Hatterman-Zogg, and M. Maifeld.

We thank the health care providers and study participants at University Physicians Clinic (Columbia, MO), Fairview Riverside Women's Clinic (Minneapolis, MN), Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center (Los Angeles, CA), Cedars-Sinai Medical Center (Los Angeles, CA), and University of Iowa Hospitals and Clinics (Iowa City, IA). We also thank M. Silva, J. Reidy, E. Samandar, and J. Preau for phthalate analyses and E. Gray, P. Foster, and D. Barr for their guidance.

This work was supported by grants from the U.S. Environmental Protection Agency and the National Institutes of Health (R01-ES09916 to the University of Missouri, MO1-RR00400 to the University of Minnesota, MO1-RR0425 to Harbor-UCLA Medical Center) and by grant 18018278 from the State of Iowa to the University of Iowa.

The authors declare they have no competing financial interests.

Received 7 March 2005; accepted 25 May 2005.

measured using the neutral single-cell gel electrophoresis (comet) assay (Duty et al. 2003b). In this population of men attending an infertility clinic, increased urinary concentration of MBzP was also associated with decreased follicle stimulating hormone, whereas increases in monobutyl phthalate were marginally associated with increased inhibin-B (Duty et al. 2005).

Newborn male rodents have no scrotum, and the external genitalia are undeveloped; only a genital tubercle is apparent for both sexes. The distance from the anus to the insertion of this tubercle, the AGD, is androgen dependent and about twice as long in males as in females. The AGD has been shown to be a sensitive measure of prenatal antiandrogen exposure (Rhees et al. 1997). Recently, Salazar-Martinez et al. (2004) studied AGD in 45 male and 42 female infants. They measured the distance from the anus to the base of the scrotum in males and from the anus to the base of the genitals (the fourchette) in females. By these measures, AGD was sexually dimorphic and about twice as long in males as in females. No other studies have examined AGD among human males, although two other studies have evaluated AGD in female infants (Callegari et al. 1987; Phillip et al. 1996).

Materials and Methods

Study participants. Women included in our study were originally recruited into the first phase of the Study for Future Families (SFFI), a multicenter pregnancy cohort study, at prenatal clinics in Los Angeles, California (Harbor-UCLA and Cedars-Sinai), Minneapolis, Minnesota (University of Minnesota Health Center), and Columbia, Missouri (University Physicians), from September 1999 through August 2002. Data collection is still ongoing in Iowa, where a center was added late in SFFI, so Iowa participants are not included in this analysis. Methods are described in detail elsewhere (Swan et al. 2003). Briefly, couples whose pregnancy was not medically assisted were eligible unless the woman or her partner was < 18 years of age, either partner did not read and speak Spanish or English, or the father was unavailable or unknown. All participants completed a questionnaire, most gave blood samples, and after urine collection was added midway through the study, most also gave a urine sample.

Eighty-five percent of SFFI participants agreed to be recontacted, and we invited these mothers to take part in our follow-up study. The family was eligible for the follow-up study (SFFII) if the pregnancy ended in a live birth, the baby was 2–36 months of age, and the mother lived within 50 mi of the clinic and could attend at least one study visit. Here we report on results from the first study visit only. Human subject committees at all participating institutions approved SFFI and SFFII, and all participants signed informed consents for each study.

Physical examination. After standard anthropometric measurements (height, weight, head circumference, and skin-fold thickness) were obtained, a detailed examination of the breast and genitals was conducted under the supervision of pediatric physicians who were trained in its administration. Every attempt was made to standardize the examination, which was developed specifically for this study. These methods included training sessions before and during the study and the use of standardized equipment. Neither the pediatric physicians nor the support staff had any knowledge of the mother's phthalate concentrations.

Boys' genital examinations included a description of the testes and scrotum, location and size of each testicle, and measurement of the penis. The placement of each testicle was initially coded in six categories; in the present analysis, boys are dichotomized into those with normal testicular descent (placement of both testes coded as normal or normal retractile) or with incomplete testicular descent (all other cases). The scrotum was categorized as distinct from surrounding tissue or not, and by size (small or not). Penile width and (stretched) length were recorded, and penile volume [proportional to (penile width/2)^2 × penile length] was calculated. We recorded the AGD, measured from the center of the anus to the anterior base of the penis. We also recorded the anoscrotal distance (ASD), measured from the center of the anus to the posterior base of the scrotum. This latter measurement was used by Salazar-Martinez et al. (2004), who refer to it as AGD.

Phthalate metabolite analysis. Urinary phthalate metabolite analyses were carried out by the Division of Laboratory Sciences, National Center for Environmental Health, Centers for Disease Control and Prevention (CDC), which had no access to participant data. The analytical approach for the analysis of urinary phthalate metabolites (Silva et al. 2004b) is a modification of previously published methods (Silva et al. 2003). The analysis involves the enzymatic deconjugation of the phthalate metabolites from their glucuronidated form, automated on-line solid-phase extraction, separation with high-performance liquid chromatography, and detection by isotope-dilution tandem mass spectrometry. This high-throughput method allows for the simultaneous quantification in human urine of the nine phthalate metabolites reported in this work. Limits of detection (LOD) are in the low nanogram per milliliter range. Isotopically labeled internal standards were used along with conjugated internal standards to increase precision and accuracy of the measurements. The method is accurate (spiked recoveries are near 100%), and precise with between-day relative standard deviations of < 10%. Quality control (QC) samples and laboratory blanks were analyzed along with unknown samples to monitor performance of the method. The metabolite concentrations reported here are from 85 prenatal maternal urine samples of a total of 214 that also included postnatal maternal and baby samples from the same mothers and their children. The 214 samples were analyzed for phthalate metabolites in six batches, none of which had to be re-extracted for QC failures. Of the 214 samples, seven were re-extracted using < 1 mL of urine because concentrations of MEP calculated using 1 mL were above the linear range of the method.

Statistical analysis. After examining descriptive and summary statistics for all study variables, we explored models for AGD. We fit several alternative measures of body size (weight, height, and body mass index) and both additive and multiplicative functions of these. We defined the anogenital index [AGI = AGD/weight (mm/kg)] as a weight-normalized index of AGD.

AGD and AGI were modeled as both linear and quadratic functions of age. For babies born at < 38 weeks, age at examination in the first year was calculated from the estimated date of conception instead of the birth date. Once the best fitting model was identified, we plotted the expected AGI and its 25th and 75th percentiles as a function of age. We categorized boys in two ways: We dichotomized boys into those with AGI smaller than or at least as large as expected, and we used the difference between observed and expected AGI to define three groups of boys, short (AGI < 25th percentile for age), intermediate (25th percentile ≤ AGI < 75th percentile), and long (AGI ≥ 75th percentile for age) AGI. We also calculated the proportion of boys in these three groups with normal testicular descent (both testes normal or normal retractile) and normal scrotal (scrotum of normal size and distinct from surrounding tissue). We calculated the correlations between AGD and AGI and penile volume, testicular placement and scrotal parameters (size and distinctness from surrounding tissue). Our decision to use AGI as the measure of genital development was made, and cut points for categorical analyses of outcomes were selected, before obtaining phthalate metabolite values.

We used general linear models to explore the relationships between phthalate metabolite concentration (unadjusted for urine concentration) and genital parameters. Most metabolite concentrations were above the LOD; those below the LOD were assigned the value LOD divided by the square root of 2, which has been recommended when the data are not highly skewed, as was the case here (Hornung and

Reed 1990). Metabolite concentrations were logarithmically transformed to normalize distributions. We examined several potentially confounding factors including mother's ethnicity and smoking status, time of day and season in which the urine sample was collected, gestational age at sample collection, and baby's weight at examination.

We also categorized metabolite concentrations into low (< 25th percentile), intermediate (between the 25th and 75th percentiles), and high (≥ 75th percentile) categories and examined the odds ratio (OR) for smaller than expected AGI for babies with high compared with low exposure, and medium compared with low. On the basis of these regression and categorical analyses, we identified the phthalate metabolites most strongly associated with AGI. We refer to these as AGI-associated phthalates.

Because phthalate metabolite concentrations are highly correlated, and because our limited sample size prohibited us from examining multiway interactions, we constructed a summary phthalate score to examine the effect of joint exposure to more than one AGI-associated phthalate. For this purpose, we used quartiles of metabolite concentration; values in the lowest quartile did not contribute to the sum, whereas higher values increased the sum one unit per quartile. We divided this sum into three categories: low (0–1, reflecting little or no exposure to AGI-associated phthalates), intermediate (2–10), and high (11–12, reflecting high exposure to all, or almost all, AGI-associated phthalates). We examined the magnitude of the residual (observed – expected) AGI as a function of this summary phthalate score.

Results

The population for the present analysis was identified from families recruited in California, Minnesota, or Missouri for whom data entry was complete by 17 December 2004, the cutoff date for the present analysis. At that time, 654 participants from these three centers had completed SFFI and given permission to be recontacted. Of these, 477 (72.9%) were eligible for SFFII and 346 (72.5%) participated (Table 1). SFFII participants were demographically similar to nonparticipants except that nonparticipants were more likely to be Hispanic because of a lower eligibility rate (60%) in CA, where most participants were Hispanic. Of the 172 boys born to these mothers, we excluded 5 boys in twin births, 10 boys with incomplete data, and 23 boys for whom AGD was not recorded [two whose mothers declined the genital exam, with the remainder older boys (mean age, 19.6 months), for whom the study examiner felt the measurement was not reliable, usually because of the boys' activity level]. The remaining 134 boys comprise the sample used for the analysis of AGD and other genital measurements. Among the 134 boys for whom we have genital measurements, no frank genital malformations or disease were detected, and no parameters appeared grossly abnormal. The mean age at first examination was 15.9 months, and mean weight was 10.5 kg (Table 2). Mean (± SD) AGD was 70.3 ± 11.0 mm, with a distribution that was well approximated by a normal curve. Overall, 86.6% of boys had both testes classified as normal or normal-retractile.

A prenatal urine sample was assayed for phthalate metabolites for mothers of 85 of these boys. These mother–son pairs comprise the data set for the analysis of AGD and phthalate metabolite concentration. Because urine collection began midway through SFFI, mothers with a stored urine sample were recruited later in the study, and their sons tended to be younger at examination (mean age, 12.6 months; interquartile range, 5–16 months). Summary statistics for all boys included in the analysis of physical measurements, and the subset of boys for whom mothers' prenatal phthalate concentrations were also available are shown separately in Table 2.

All phthalate metabolites tested were above the LOD in > 49% of women, and most tested were above the LOD in > 90% of the samples (Table 3). Concentrations spanned four orders of magnitude, from below the LOD (estimated value = 0.71 ng/mL) to 13,700 ng/mL for MEP. Means ranged from 2.68 for mono-3-carboxypropyl phthalate (MCPP) to 629.8 for MEP. Three of the four AGI-associated metabolites (other than MEP) were significantly correlated ($p < 0.005$).

Regression analyses. We initially modeled AGD as a linear function of age and weight, but this model fit poorly (adjusted $R^2 = 0.22$). We found that using AGI (AGD/weight) as a function of age provided the best fit, as has been shown in rodent models (Vandenbergh and Huggett 1995). The best-fitting model for AGI includes linear and quadratic terms for age and is given by AGI = 10.8835 – 0.3798 (age) + 0.0068 (age^2) (adjusted $R^2 = 0.61$). Using this model, we calculated mean AGI and its 5th, 25th, 75th, and 95th percentiles (Figure 1).

We then examined models that included individual phthalate metabolites. Other than age and age squared, no covariates altered regression coefficients for the phthalate metabolites by > 15%, and none were included in final models. All regression coefficients for individual metabolites (logarithmically transformed to normalize distributions) were negative (Table 4). MEP, mono-*n*-butyl phthalate (MBP), MBzP, and monoisobutyl phthalate (MiBP) were (inversely) related to AGI; *p*-values for regression coefficients were between 0.007 and 0.097. We also measured three metabolites of DEHP. Although the hydrolytic monoester metabolite mono-2-ethylhexyl phthalate (MEHP) was unrelated to AGI [regression coefficient = –0.05; 95% confidence interval (CI), –0.53 to 0.43], regression

Table 1. Participants included in present analysis.

	No.	Percent potential participants	Percent male babies
All pregnancy outcomes (CA, MN, and MO)			
Potential participants[a]	654	100	—
Eligible for SFFII	477	72.9	—
SFFII participant	346	72.5	—
Male babies only (CA, MN, and MO)			
SFFII participant	172	—	100
With AGD, age, and weight[b]	134	—	78
Prenatal urine sample[c]	85	—	49

[a]A potential participant is an SFFI participant from CA, MO, or MN who gave permission to be recontacted for future studies and for whom all study data were entered by 17 December 2004. [b]Boys in twin births and boys with missing data or AGD measurements considered unreliable by pediatricians excluded. [c]Urine collection began midway through SFFI.

Table 2. Characteristics of boys with complete physical examination.

		Percentile		
Characteristic	Mean ± SD	25th	50th	75th
All boys (*n* = 134)				
Age (months)	15.9 ± 8.6	11.0	15.0	23.0
Height (cm)	79.1 ± 10.6	72.6	80.0	87.2
Weight (kg)	10.5 ± 2.7	8.7	10.7	12.3
AGD (mm)	70.3 ± 11.0	63.9	70.3	76.6
AGI (mm/kg)	7.1 ± 1.9	5.8	6.7	7.8
ASD (mm)	37.4 ± 7.5	31.2	36.8	43.4
Boys whose mother's prenatal urine was assayed for phthalate metabolites (*n* = 85)				
Age (months)	12.6 ± 6.9	5.0	14.0	16.0
Height (cm)	75.6 ± 9.5	66.5	77.6	82.0
Weight (kg)	9.7 ± 2.4	8.4	10.0	11.1
AGD (mm)	68.0 ± 9.7	61.7	66.7	74.4
AGI (mm/kg)	7.4 ± 1.8	6.1	7.0	8.2
ASD (mm)	35.9 ± 7.1	30.4	35.6	41.4

coefficients for the oxidative monoester metabolites of DEHP, mono-2-ethyl-5-oxohexyl phthalate (MEOHP), and mono-2-ethyl-5-hydroxyhexyl phthalate (MEHHP) were of a magnitude comparable with those for MEP and MBzP (p-values = 0.114 and 0.145 for MEOHP and MEHHP, respectively). AGI appeared to be independent of the concentrations of monomethyl phthalate (MMP) and MCPP, metabolites of dimethyl phthalate and di-*n*-octyl phthalate, respectively.

Categorical analyses. The 25 boys with AGI below the 25th percentile for age were classified as having a short AGI. This group had an AGI that was, on average, 18.3% (range, 10–32%) shorter than expected based on the final regression model. Boys with AGI ≥ 75th percentile of expected were classified as having a long AGI, and boys with AGI between the 25th and 75th percentile of expected were considered intermediate. Boys' weight and age did not differ appreciably among these groups.

Table 5 shows mean and median values for the AGI-associated metabolites for boys in the short, intermediate, and long categories of AGI. We calculated the ORs for short AGI for each monoester metabolite (Table 6). For high compared with low concentration of MBP, the OR for a short AGI was 10.2 (95% CI, 2.5 to 42.2), whereas for medium concentration compared with low the OR was 3.8 (95% CI, 1.2 to 12.3). The corresponding ORs for high compared with low concentration of MEP, MBzP, and MiBP were 4.7, 3.8, and 9.1, respectively (all p-values < 0.05).

Other genital parameters. Degree of testicular descent was associated with AGD (R = 0.20, p = 0.02). The proportions of boys with one or both testicles incompletely descended were 20.0, 9.5, and 5.9% for boys classified as having short, intermediate, and long AGI (p-value for short AGI compared with all other boys < 0.001). The proportion of boys with a scrotum categorized as small and/or "not distinct from surrounding tissue" was also elevated for boys with short AGI (p < 0.001). AGD was significantly associated with penile volume (R = 0.27, p = 0.001), and penile volume divided by weight was correlated with AGI (R = 0.43, p = 0.001). Testicular volume, which was measured by orchidometer, is not shown here because participating physicians considered the measurement to be unreliable—a decision made before analyses of phthalate exposure.

ASD was, on average, 47% as long as AGD, and these two measurements were correlated (R = 0.47, p < 0.0001). However, the model predicting ASD as a function of baby's age and weight fit poorly (adjusted R^2 = 0.10). The fit for the model using ASD/weight as a function of age and age squared was better (adjusted R^2 = 0.47) but did not fit as well as the model using AGI (R^2 = 0.61). ASD/weight was associated with MEP concentration (regression coefficient = –0.429; 95% CI, –0.722 to –0.137). For the other phthalate metabolites, regression coefficients were less significant (all p-values between 0.11 and 0.97).

Summary phthalate score. We used the summary phthalate score as defined in "Materials and Methods" to study the effect of joint exposure to more than one AGI-associated phthalate. The summary phthalate score was directly related to the proportion of boys with short AGI (p = 0.001). Of the 10 boys whose phthalate scores were high (score = 11–12), all but one had a short AGI. Conversely, of the 11 boys whose scores were low (score = 0 or 1), only one had a short AGI. The ORs for having a short AGI for high summary phthalate score compared with low (OR = 90.0; 95% CI, 4.88 to 1,659), and high compared with medium (29.4; 95% CI, 3.4 to 251) were large and significant, although the confidence intervals were very wide. These data are shown graphically in Figure 1.

Discussion

In the recent National Health and Nutrition Examination Survey (NHANES 1999–2000), most of the general population in the United States had measurable exposure to multiple phthalates (CDC 2003; Silva et al. 2004a). The samples in the present study and in NHANES were both analyzed using comparable methods and standards by the same laboratory, although the specific metabolites that were measured in the two studies differed somewhat. We compared the medians and 75th percentiles of the AGI-associated phthalate metabolite concentrations among two groups of mothers in our study (those whose boys fell in the short AGI group and all others) with those of females in the NHANES sample (Table 7). In the analysis of the NHANES samples, monobutyl phthalate includes both MBP and MiBP, which were measured separately in our study. Metabolite concentrations for mothers of boys with short AGI were consistently higher than those of other mothers. Compared with women in the NHANES sample, metabolite concentrations for our population were somewhat lower. However, our population cannot be directly compared with

Table 3. Percentiles of phthalate monoester metabolites.

Monoester metabolite	Percentile (ng/mL) 25th	50th	75th	Percent > LOD[a]
Phthalate monoester metabolite				
MBP	7.2	13.5	30.9	96.5
MBzP	3.5	8.3	23.5	94.1
MCPP	0.7	2.1	3.6	69.4
MEP	53.3	128.4	436.9	97.6
MiBP	0.7	2.5	5.1	74.1
MMP	0.7	0.7	3.2	49.4
Metabolites of DEHP				
MEHHP	6.0	11.4	20.1	97.6
MEHP	1.3	3.3	9.0	77.6
MEOHP	5.1	11.1	19.0	94.1

[a]LOD for all metabolites was between 0.95 and 1.07 ng/mL.

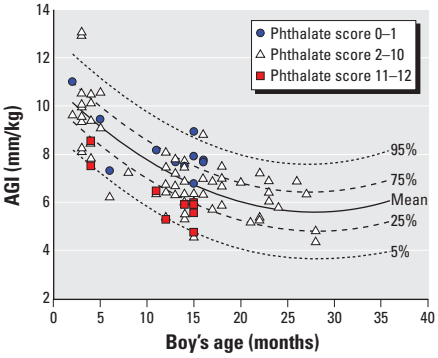

Figure 1. Mean AGI (mm/kg) in relation to boys' age at examination (months).

Table 4. Regression analyses of AGI on $\log_{10}$ monoester metabolite concentration, controlling for age and age squared.

Monoester metabolite	$\log_{10}$ monoester metabolite concentration Coefficient (SE)	p-Value (95% CI)
MBP	–0.592 (0.269)	0.031 (–1.126 to –0.057)
MBzP	–0.390 (0.232)	0.097 (–0.851 to 0.072)
MCPP	–0.264 (0.356)	0.461 (–0.973 to 0.445)
MEHHP	–0.398 (0.270)	0.145 (–0.935 to 0.140)
MEHP	–0.051 (0.241)	0.833 (–0.530 to 0.428)
MEOHP	–0.412 (0.258)	0.114 (–0.925 to 0.101)
MEP	–0.400 (0.164)	0.017 (–0.726 to –0.074)
MiBP	–0.765 (0.274)	0.007 (–1.309 to –0.220)
MMP	–0.283 (0.323)	0.383 (–0.924 to 0.359)
Phthalate score[a]	–0.0951 (0.035)	0.009 (–0.165 to –0.025)

[a]Phthalate score measures joint exposure to MBP, MBzP, MEP, and MiBP; see "Statistical analysis."

NHANES: the proportion of pregnant women in the NHANES sample is unknown, and age distributions differ. Nonetheless, these data demonstrate that the four AGI-associated phthalate metabolites are prevalent in the U.S. female population, and levels were not unusually high among mothers whose sons had a short AGI.

Although not identical, AGD in pups is most similar to AGD as we defined it in this study. In rodents, AGD has been shown to be one of the most sensitive end points for phthalates such as DBP (Mylchreest et al. 2000) and other antiandrogens such as flutamide (Barlow and Foster 2003; McIntyre et al. 2001) and finasteride (Bowman et al. 2003). It is difficult to compare the dose to humans from low-level, ongoing, environmental exposure with that delivered to rodents experimentally in a narrow window of gestation. Nonetheless, it is likely that the doses to which our participants were exposed are lower than those used in toxicologic settings, suggesting that humans may be more sensitive to prenatal phthalate exposure than rodents. This greater sensitivity in humans has been observed for other toxicants. For example, humans are more sensitive to trenbolone by an order of magnitude (Neumann 1976). This greater sensitivity is thought to be a result of rodents' higher metabolic rate and more rapid inactivation of toxicants, both of which have been shown to be inversely related to body size (White and Seymour 2005).

In light of the toxicologic literature for MBP, MBzP, and MiBP (Ema et al. 2003; Foster et al. 1980, 1981; Gray et al. 2000; Nakahara et al. 2003), our data suggest that the end points affected by these phthalates are quite consistent across species. A boy with short AGI has, on average, an AGI that is 18% shorter than expected based on his age and weight as well as an increased likelihood of testicular maldescent, small and indistinct scrotum, and smaller penile size. These changes in AGD and testicular descent are consistent with those reported in rodent studies after high-dose phthalate exposure (Ema et al. 2003; Gray et al. 2000; Mylchreest et al. 2000). The lack of association for MCPP and MMP, which have not been widely studied, is not inconsistent with the toxicologic literature.

With respect to DEP and its metabolite MEP, we note that there are three other human studies suggesting reproductive toxicity (Colón et al. 2000; Duty et al. 2003b; Main KM, unpublished data). It is therefore uncertain whether the absence of data in rodents showing reproductive toxicity is the result of failure to detect it, unmeasured confounding in human studies, or interspecies differences in response to these compounds.

DEHP has been shown to shorten AGD (Gray et al. 2000) and reduce testosterone (Parks et al. 2000). Although MEHP was not associated with AGD in our data, the associations for the oxidative metabolites of DEHP (MEOHP and MEHHP) were of comparable magnitude with those for metabolites of DBP and BzBP, although not statistically significant. Thus, it is unclear whether MEOHP and MEHHP are (inversely) associated with AGI, although associations are of borderline statistical significance because of our sample size, or whether human and rodent responses to this phthalate and its metabolites differ.

Masculinization of external male genitalia, represented by longer AGD, is controlled by dihydrotestosterone (Clark et al. 1990). Ema and Miyawaki (2001) demonstrated that this metabolite of testosterone is markedly decreased by prenatal administration of MBP, suggesting that MBP acts as an antiandrogen. AGD in male rodents is associated with other adverse developmental effects (Foster and McIntyre 2002) and some phthalate-induced changes have been shown to be permanent. For example, Barlow et al. (2004) report that prenatal exposure to 500 mg/kg/day DBP resulted in permanently decreased AGD and testicular dysgenesis. They also report that *in utero* DBP exposure induced proliferative Leydig cell lesions. Follow-up of exposed children until adulthood will be required to determine whether long-term effects, including testicular dysgenesis, are seen in humans after prenatal phthalate exposure.

Several recent studies of the variability of phthalate monoester concentration in human samples suggest that phthalate concentration in humans is fairly stable, perhaps reflecting habitual use of phthalate-containing household and consumer products (Colón et al. 2000; Hauser et al. 2004; Hoppin et al. 2002). These studies lend support to the use of a single sample for exposure assessment. We obtained only a single prenatal urine sample from each woman, and most samples were obtained quite late in pregnancy (mean = 28.3 weeks). Therefore, the measured phthalate metabolite levels may not reflect exposure during the most sensitive developmental window, resulting in some degree of exposure misclassification. However, unless this misclassification varied systematically with outcome, such errors would bias the effect estimate toward the null. In fact, the categorical analysis, which should be less sensitive to such misclassification, showed

Table 5. Mean (median) phthalate monoester metabolite levels by AGI category.

Monoester metabolite	AGI category [mean (median; ng/mL)]		
	Long[a] (n = 17)	Intermediate[b] (n = 43)	Short[c] (n = 25)
MBP	13.1 (11.5)	22.2 (13.1)	38.7 (24.5)
MBzP	10.6 (6.6)	15.1 (7.7)	25.8 (16.1)
MEP	124 (47.1)	592 (112)	1,076 (225)
MiBP	2.3 (1.5)	3.3 (2.1)	7.7 (4.8)

[a]Long, AGI ≥ 75th percentile of expected AGI. [b]Intermediate, 25th percentile ≤ AGI < 75th percentile of expected AGI. [c]Short, AGI < 25th percentile of expected AGI.

Table 6. OR (95% CI) for AGI less than expected from regression model, by monoester metabolite level.

Monoester metabolite	Level (percentile)	AGI < expected	AGI ≥ expected	OR (95% CI)
MBP	Low	5	15	Referent
	Medium	24	19	3.8 (1.2 to 12.3)
	High	17	5	10.2 (2.5 to 42.2)
MBzP	Low	6	13	Referent
	Medium	26	18	3.1 (1.002 to 9.8)
	High	14	8	3.8 (1.03 to 13.9)
MEP	Low	7	14	Referent
	Medium	25	19	2.6 (0.9 to 7.8)
	High	14	6	4.7 (1.2 to 17.4)
MiBP	Low	6	16	Referent
	Medium	23	18	3.4 (1.1 to 10.5)
	High	17	5	9.1 (2.3 to 35.7)

Low, < 25th percentile; medium, ≥ 25th and < 75th percentile; high, ≥ 75th percentile.

Table 7. Concentrations of four phthalate metabolites in three groups of women (ng/mL).

Monoester metabolite	Percentile	This study		NHANES[a]
		Short AGI	Others	
MBP	50th	24.5	12.1	30.0
	75th	44.8	28.0	59.5
MBzP	50th	16.1	7.2	16.0
	75th	27.5	17.8	35.8
MEP	50th	225	90.4	174
	75th	551	281	425
MiBP	50th	4.8	2.1	—[b]
	75th	12.1	4.3	—[b]

[a]Females only (CDC 2003). [b]MBP in the NHANES analysis includes both MBP and MiBP; in this study these metabolites were measured separately.

stronger associations than did the continuous analysis.

Our analysis is based on a single measure of AGD, and the reliability of this measurement in humans has not been established. During two training sessions, three study physicians each measured AGD in four male infants (mean age, 8.1 months). The mean AGD for these measurements was 58.6 mm, SD was (within infant) 4.2 mm, and coefficient of variation of 7.2%, suggesting that AGD can be measured reliably. Use of this measurement in larger studies in a range of diverse populations, with many more such training sessions, will be needed to obtain normative data.

Although it might have been ideal to examine babies shortly after birth, the timing of grant funding did not allow this. Babies were born to SFFI mothers as early as January 2000, and the first baby visits did not occur until April 2002. To maximize the number of children participating, we allowed recruitment over a range of ages. On the other hand, because the use of AGD in humans is new, the optimal timing for this measurement is not known. Our data suggest that measurements are reliable and informative in young children at least until 18 months, when AGD becomes more difficult to obtain reliably. Its value in adolescents and adults has yet to be determined.

We note that phthalate metabolite levels were highly correlated, and most women were exposed to all metabolites at detectable levels. Gray et al. (2000) suggested that risk assessments for phthalate-induced reproductive toxicity should consider phthalates as a group and include exposures from multiple sources. The score we use reflects joint exposure to the four AGI-associated phthalates, and our results suggest that joint exposure may convey greater than additive risk, but larger sample sizes are needed to confirm this.

Gray and Foster (2003) refer to a "phthalate syndrome" characterized by testicular, epididymal, and gubernacular cord agenesis as well as decreased AGD, and stress the importance of evaluating all components of a syndrome so that affected animals are not misidentified. It has recently been suggested (Fisher 2004) that this "phthalate syndrome" shares many features with the human testicular dysgenesis syndrome proposed by Skakkebaek et al. (2001) to follow chemically induced disruption of embryonic programming and gonadal development during fetal life. The present findings, though based on small numbers, provide the first data in humans linking measured levels of prenatal phthalates to outcomes that are consistent with this proposed syndrome.

This is the first study to look at subtle patterns of genital morphology in humans in relation to any prenatal exposure. It was motivated by toxicologic studies showing that genital morphology is altered by antiandrogens, including some phthalates. We report that AGD, the most sensitive marker of antiandrogen action in toxicologic studies, is shortened and testicular descent impaired in boys whose mothers had elevated prenatal phthalate exposure. These changes in male infants, associated with prenatal exposure to some of the same phthalate metabolites that cause similar alterations in male rodents, suggest that commonly used phthalates may undervirilize humans as well as rodents.

References

Barlow NJ, Foster PM. 2003. Pathogenesis of male reproductive tract lesions from gestation through adulthood following *in utero* exposure to di(*n*-butyl) phthalate. Toxicol Pathol 31:397–410.

Barlow NJ, McIntyre BS, Foster PM. 2004. Male reproductive tract lesions at 6, 12, and 18 months of age following *in utero* exposure to di(*n*-butyl) phthalate. Toxicol Pathol 32:79–90.

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

Callegari C, Everett S, Ross M, Brasel JA. 1987. Anogenital ratio: measure of fetal virilization in premature and full-term newborn infants. J Pediatr 111:240–243.

CDC. 2003. Second National Report on Human Exposure to Environmental Chemicals. Atlanta, GA:Centers for Disease Control and Prevention, National Center for Environmental Health, Division of Laboratory Sciences.

Clark RL, Antonello JM, Grossman SJ, Wise LD, Anderson C, Bagdon WJ, et al. 1990. External genitalia abnormalities in male rats exposed in utero to finasteride, a 5 alpha-reductase inhibitor. Teratology 42:91–100.

Colón I, Caro D, Bourdony CJ, Rosario O. 2000. Identification of phthalate esters in the serum of young Puerto Rican girls with premature breast development. Environ Health Perspect 108:895–900.

Duty SM, Calafat AM, Silva MJ, Ryan L, Hauser R. 2005. Phthalate exposure and reproductive hormones in adult men. Hum Reprod 20:604–610.

Duty SM, Silva MJ, Barr DB, Brock JW, Ryan L, Chen Z, et al. 2003a. Phthalate exposure and human semen parameters. Epidemiology 14:269–277.

Duty SM, Singh NP, Silva MJ, Barr DB, Brock JW, Ryan L, et al. 2003b. The relationship between environmental exposures to phthalates and DNA damage in human sperm using the neutral comet assay. Environ Health Perspect 111:1164–1169.

Ema M, Miyawaki E. 2001. Adverse effects on development of the reproductive system in male offspring of rats given monobutyl phthalate, a metabolite of dibutyl phthalate, during late pregnancy. Reprod Toxicol 15:189–194.

Ema M, Miyawaki E, Hirose A, Kamata E. 2003. Decreased anogenital distance and increased incidence of undescended testes in fetuses of rats given monobenzyl phthalate, a major metabolite of butyl benzyl phthalate. Reprod Toxicol 17:407–412.

Fisher JS. 2004. Environmental anti-androgens and male reproductive health: focus on phthalates and testicular dysgenesis syndrome. Reproduction 127:305–315.

Foster PM, Cattley RC, Mylchreest E. 2000. Effects of di-*n*-butyl phthalate (DBP) on male reproductive development in the rat: implications for human risk assessment. Food Chem Toxicol 38:S97–S99.

Foster PM, Lake BG, Thomas LV, Cook MW, Gangolli SD. 1981. Studies on the testicular effects and zinc excretion produced by various isomers of monobutyl-*o*-phthalate in the rat. Chem Biol Interact 34:233–238.

Foster PM, McIntyre BS. 2002. Endocrine active agents: implications of adverse and non-adverse changes. Toxicol Pathol 30:59–65.

Foster PM, Thomas LV, Cook MW, Gangolli SD. 1980. Study of the testicular effects and changes in zinc excretion produced by some *n*-alkyl phthalates in the rat. Toxicol Appl Pharmacol 54:392–398.

Gray LE Jr, Foster PMD. 2003. Significance of experimental studies for assessing adverse effects of endocrine-disrupting chemicals. Pure Appl Chem 75:2125–2141.

Gray LE Jr, Ostby J, Furr J, Price M, Veeramachaneni DNR, Parks L. 2000. Perinatal exposure to the phthalates DEHP, BBP, and DINP, but not DEP, DMP, or DOTP, alters sexual differentiation of the male rat. Toxicol Sci 58:350–365.

Hauser R, Meeker JD, Park S, Silva MJ, Calafat AM. 2004. Temporal variability of urinary phthalate metabolite levels in men of reproductive age. Environ Health Perspect 112:1734–1740.

Hoppin JA, Brock JW, Davis BJ, Baird DD. 2002. Reproducibility of urinary phthalate metabolites in first morning urine samples. Environ Health Perspect 110:515–518.

Hornung RW, Reed LD. 1990. Estimation of average concentration in the presence of nondetectable values. Appl Occup Environ Hyg 5:46–51.

McIntyre BS, Barlow NJ, Foster PM. 2001. Androgen-mediated development in male rat offspring exposed to flutamide *in utero*: permanence and correlation of early postnatal changes in anogenital distance and nipple retention with malformations in androgen-dependent tissues. Toxicol Sci 62:236–249.

McKee RH, Toxicology Research Task Group. 2004. Phthalate exposure and early thelarche. Environ Health Perspect 112:A541–A543.

Murature DA, Tang SY, Steinhardt G, Dougherty RC. 1987. Phthalate esters and semen quality parameters. Biomed Environ Mass Spectrom 14:473–477.

Mylchreest E, Wallace DG, Cattley RC, Foster PMD. 2000. Dose-dependent alterations in androgen-regulated male reproductive development in rats exposed to di(*n*-butyl) phthalate during late gestation. Toxicol Sci 55:143–151.

Nagao T, Ohta R, Marumo H, Shindo T, Yoshimura S, Ono H. 2000. Effect of butyl benzyl phthalate in Sprague-Dawley rats after gavage administration: a two-generation reproductive study. Reprod Toxicol 14:513–532.

Nakahara H, Shono T, Suita S. 2003. Reproductive toxicity evaluation of dietary butyl benzyl phthalate (BBP) in rats. Fukuoka Igaku Zasshi 94:331–337.

Neumann F. 1976. Pharmacological and endocrinological studies on anabolic agents. Environ Qual Saf 5(suppl)253–264.

Parks LG, Ostby JS, Lambright CR, Abbott BD, Klinefelter GR, Barlow NJ, et al. 2000. The plasticizer diethylhexyl phthalate induces malformations by decreasing fetal testosterone synthesis during sexual differentiation in the male rat. Toxicol Sci 58:339–349.

Phillip M, De Boer C, Pilpel D, Karplus M, Sofer S. 1996. Clitoral and penile sizes of full term newborns in two different ethnic groups. J Pediatr Endocrinol Metab 9:175–179.

Rhees RW, Kirk BA, Sephton S, Lephart ED. 1997. Effects of prenatal testosterone on sexual behavior, reproductive morphology and LH secretion in the female rat. Dev Neurosci 19:430–437.

Salazar-Martinez E, Romano-Riquer P, Yanez-Marquez E, Longnecker MP, Hernandez-Avila M. 2004. Anogenital distance in human male and female newborns: a descriptive, cross-sectional study. Environ Health 3:8; doi:10.1186/1476-069X-3-8 [Online 13 September 2004].

Silva MJ, Barr DB, Reidy JA, Malek NA, Hodge CC, Caudill SP, et al. 2004a. Urinary levels of seven phthalate metabolites in the U.S. population from the National Health and Nutrition Examination Survey (NHANES) 1999–2000. Environ Health Perspect 112:331–338.

Silva MJ, Malek NA, Hodge CC, Reidy JA, Kato K, Barr DB, et al. 2003. Improved quantitative detection of 11 urinary phthalate metabolites in humans using liquid chromatography-atmospheric pressure chemical ionization tandem mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 789:393–404.

Silva MJ, Slakman AR, Reidy JA, Preau JL Jr, Herbert AR, Samandar E, et al. 2004b. Analysis of human urine for fifteen phthalate metabolites using automated solid-phase extraction. J Chromatogr B Analyt Technol Biomed Life Sci 805:161–167.

Skakkebaek NE, Rajpert-De Meyts E, Main KM. 2001. Testicular dysgenesis syndrome: an increasingly common developmental disorder with environmental aspects. Hum Reprod 16:972–978.

Swan SH, Brazil C, Drobnis EZ, Liu F, Kruse RL, Hatch M, et al. 2003. Geographic differences in semen quality of fertile U.S. males. Environ Health Perspect 111:414–420.

Tyl RW, Myers CB, Marr MC, Fail PA, Seely JC, Brine DR, et al. 2004. Reproductive toxicity evaluation of dietary butyl benzyl phthalate (BBP) in rats. Reprod Toxicol 18:241–264.

Vandenbergh JG, Huggett CL. 1995. The anogenital distance index, a predictor of the intrauterine position effects on reproduction in female house mice. Lab Anim Sci 45:567–573.

White CR, Seymour RS. 2005. Allometric scaling of mammalian metabolism. J Exp Biol 208:1611–1619.