SUPPLEMENT

Supplement to: Adgent MA, Umbach DM, Zemel BS, *et al*.

A longitudinal study of estrogen-responsive tissues and hormone concentrations in infants fed soy formula

Table of Contents Page

Methods: Additional Details

Study Design and Recruitment

IFED enrollment occurred between August 2010 and November 2013. Mothers were enrolled on the basis that they intended to feed their infants simple, exclusive diets throughout the course of study. They were enrolled in the 3rd trimester of pregnancy ("prenatal recruitment") or within 72 hours of the birth of their baby ("birth recruitment"). For prenatal recruitment, pregnant mothers who were planning to deliver at the Hospital of the University of Pennsylvania (Penn) were recruited from obstetric and midwifery practices, health clinics, and prenatal classes/activities, all within close proximity to the Children's Hospital of Philadelphia (CHOP) (November 2010 – April 2012). Birth recruitment began at PENN (August 2010), and later expanded to additional hospitals in order to expedite overall enrollment. Additional recruitment hospitals were introduced as follows: Pennsylvania Hospital (PH) (February 2012), Virtua Voorhees (September 2012), Holy Redeemer Hospital (October 2012), Virtua Memorial (November 2012), Cooper Memorial Hospital (January 2013), and Abington Memorial Hospital (February 2013). One subject self-referred to the study following birth at Lankenau Hospital (August 2012). Recruitment flyers were placed in Jefferson University Hospital (April 2013), but no enrollments were procured.

Pregnant women were initially screened for eligibility by medical record review/chart abstraction, and then provided verbal consent for a maternal screening interview. All mothers and infants were also screened at birth. Endocrine and thyroid conditions that prompted ineligibility included polycystic ovary syndrome, Cushing's syndrome, congenital adrenal hyperplasia, Addison's disease, Grave's disease, Hashimoto's thyroiditis, hyperthyroidism, and hypothyroidism. Mothers' intentions for feeding method were volunteered by mothers during this screening process. Due to their relative abundance,

recruitment of breast feeding mothers was completed before April 2012. Mothers feeding soy-protein formula were less common and were recruited through the end of the recruitment period.

Incentives were provided to families at each visit. Families using formula received all formula needed to feed their infant for the duration of the study in ready-to feed containers. Mothers were offered a choice of products from Nestlé's (Gerber Good Start [cow] and Gerber Good Start Soy); Mead Johnson (Enfamil Premium [cow] and Prosobee [soy]); or Abbott (Similac Advanced [cow] Similac Soy Isomil [soy]). Since most families were eligible for supplemental nutrition through Special Supplemental Nutrition Program for Women, Infants and Children (WIC), providing formula was not a financial benefit, but it was much more convenient. Breastfeeding families received an electric breast pump and a supplemental supply of diapers. The dollar value of the supplies to each feeding group family was about equal. With signed release forms, photos were taken of the infants at each study visit and compiled into an album, which was given to the mother at the last study visit.

Ultrasound

Birth ultrasounds were conducted at Penn (n = 223), Pennsylvania Hospital (n = 24), and Virtua Vorhees Hospital (n = 19); the remaining 17 subjects were scanned at one of the other enrollment hospitals using instrumentation and procedures described previously (1). Follow-up visits were conducted at CHOP using three Philips iU22 machines (Philips Healthcare; Bothell, WA) or at Virtua Vorhees using a GE LogiQ E9 (GE Ultrasound; Wauwatosa, WI), with the exception of four follow up visits from two subjects completed at Abington Memorial (Philips iU22).

Images were obtained by one of seven sonographers, all trained and certified in the study protocol that defined the order in which images were obtained as well as other organ-specific procedures (1). In addition, the sonographer completed a quality assurance (QA) image without calipers or measurements; the IFED radiologist later evaluated a sample of these unmeasured images (approximately 10% of

images per sonographer) and the radiologist's measurement was used as a 'gold-standard' in evaluating inter-rater variability.

Urinary Phytoestrogens

We analyzed phytoestrogen content in urine samples from a randomly selected subset of cow-milk formula-fed (n = 19) and soy formula-fed (n = 20) infants. All samples were collected at 12 weeks of age. Phytoestrogen concentrations, including genistein, were determined by use of high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry base on a method described previously (Centers for Disease Control and Prevention, Atlanta, GA)(2). The limit of detection for genistein was 0.2 ng/mL; intraassay and interassay CVs were 5.1-5.8% and 3.9-5.8%, respectively.

Statistical Analysis

We analyzed differences among feeding groups in maternal and infant characteristics using chi-squared statistics for categorical characteristics and t-tests for continuous ones. In both instances, we examined two orthogonal comparisons: 1) infants fed cow-milk formula compared to infants fed soy formula; 2) breastfed infants compared to all formula-fed infants together. For chi-squared tests, we used PROC FREQ (SAS v9.3, Cary, NC) to calculate likelihood ratio statistics and compute p values using exact, instead of asymptotic, methods. For t-tests, we included the three feeding groups as factors in an analysis of variance and constructed tests using contrasts; we used PROC GLM (SAS v9.3, Cary, NC). Examination of residuals indicated that assumptions needed for valid analysis of variance were met.

We used mixed-effects regression splines to examine age trajectories of cytological, organ-size, and hormonal outcomes. For the fixed-effects portion of the model, we modeled each feeding-group-specific trajectory as a natural cubic spline in the square-root of age. We considered splines with from three to six knots equally spaced in the square root of age, same knot locations for each feeding group. For the random portion of the mixed-effects model, we allowed a subject-specific random effect for

each spline coefficient and modeled those coefficients as mean-zero normal random variables with an unstructured covariance matrix. More specifically, with *k* knots, a natural cubic spline involves *k* regression coefficients; consequently, the subject-specific random effects had a *k* x *k* (symmetric) variance-covariance matrix common to all feeding groups. Our random specification estimated each unique entry in the matrix as a distinct parameter. For enhanced numerical stability and to guarantee a nonnegative estimated matrix, we parameterized this matrix through its Cholesky root. We used PROC GLIMMIX (SAS v9.3, Cary, NC) to fit these models.

We used the Bayesian information criterion (BIC) to select the number of knots that provided a parsimonious model with adequate fit. This criterion selected a four-knot model for each outcome. In addition, we examined residuals from each model to verify compliance with needed assumptions. We also checked whether an additional random-effect to allow for within-subject serial correlation would improve model fit and found that that additional term was unnecessary for every outcome.

To display these postnatal trajectory comparisons (denoted “relative trajectories”), we shifted the fitted feeding-group-specific trajectories (denoted “absolute trajectories”) vertically to have a common intercept. Postnatal relative trajectories were calculated by subtracting the predicted response at age zero from the predicted response at each subsequent age. This difference is zero at age zero. For maturation index, which we analyzed without transformation, relative trajectories for each feeding group have intercept zero. For other outcomes, which we analyzed as $\log_2$-tranformed, relative trajectories have intercept one, the anti-logarithm of zero. Because a difference in logarithms is the logarithm of a ratio, the relative trajectories for $\log_2$-tranformed responses can be interpreted as the ratio of the response at each subsequent age to the response at age zero.

Supplemental Table 1. Visit and measurement schedule for IFED infants

Procedure	Measurement	Birth	2 weeks	4 weeks	6 weeks	8 weeks	12 weeks	16 weeks	20 weeks	24 weeks	28 weeks	32 weeks	36 weeks
Girls													
Ultrasound	Breast	•		•		®		•		•		•	
	Uterus	•		•		®		•		•		•	
Blood Draw	E2, FSH		•	•	®	•	•	•	•	•	•	•	•
Vaginal Swab	Maturation Index	•	•	•	®	•	•	•	•	•	•	•	•
Physical Exam	Length, Weight, etc.	•	•	•	®	•	•	•	•	•	•	•	•
Boys													
Ultrasound	Breast	•		•		®		•		•			
Blood Draw	E2		•	•	®	•	•	•	•	•	•		
Urethral Swab	Maturation Index	•	•	•	®	•	•	•	•	•	•		
Physical Exam	Height, Weight, etc.	•	•	•	®	•	•	•	•	•	•		

® Procedure eliminated in June 2012; E2: estradiol; FSH: follicle stimulating hormone

119 Supplemental Table 2. Maternal and infant characteristics by feeding group.

	Feeding Groups				
	Soy	***Cow***	***Breast***	**Total Complete[a]**	**Incomplete/ Excluded[b]**
Number of Dyads	102	111	70	283	127
Maternal Race					
Black	80(78)	90(81)	29(41)	199(70)	77(65)
White	18(18)	15(14)	30(43)	63(22)	28(23)
Other/Multi/Unknown	4(4)	6(5)	11(16)	21(7)	14(12)
X^2 p- value[c]	*0.65*		*<0.0001*		*0.34*
Maternal Education					
≤ Some High School	25(25)	27(24)	0(0)	52(18)	25(21)
High School or GED	44(43)	50(45)	16(23)	110(39)	40(34)
Some College or Assoc.	27(26)	27(24)	20(28)	74(26)	34(29)
College or >	6(6)	7(6)	34(49)	47(17)	20(17)
X^2 p- value[c]	*0.99*		*<0.0001*		*0.78*
Maternal Age (years)					
Mean (Standard Deviation)	26.0(5.2)	25.5(6.0)	27.5(5.3)	26.2(5.6)	26.3(5.7)
Median	26	24	28	26	26
Minimum, Maximum	18, 43	18, 42	18, 38	18, 43	18, 43
p- value[d]	*0.47*		*0.02*		*0.79*
Sex					
Girls	48(47)	56(50)	32(46)	136(48)	60(50)
Boys	54(53)	55(50)	38(54)	147(52)	59(50)
X^2 p- value [c]	*0.68*		*0.68*		*0.74*
Neonatal Weight,[e] kg					
Mean (Standard Deviation)	3.22(0.4)	3.15(0.4)	3.28(0.4)	3.21(0.4)	3.19(0.4)
Median	3.25	3.11	3.28	3.18	3.19
Minimum, Maximum	2.40, 4.12	2.36, 4.28	2.51, 4.20	2.36, 4.28	2.49, 4.21
p- value[d]	*0.16*		*0.07*		*0.58*
Gestational Age, weeks					
37	8(8)	11(10)	4(6)	23(8)	10(8)
38	15(15)	25(23)	11(16)	51(18)	24(20)
39	50(49)	36(32)	20(28)	106(38)	43(36)
40	23(22)	27(24)	19(27)	69(24)	28(24)
41	6(6)	12(11)	16(23)	34(12)	14(12)
X^2 p- value [c]	*0.13*		*0.03*		*0.99*
Feeding Group at Enrollment					
Soy formula				102(36)	48(38)
Cow-milk formula				111(39)	40(31)
Breast milk				70(25)	39(31)
X^2 p- value [c]					*0.98*

120

121 Entries for categorical variables are number of infants (percentage of feeding-group total); entries for
122 continuous variables are named in the respective rows.

[a]Total Complete: IFED subjects who contributed a minimum of 3 (boys) to 4 (girls) blood samples and who were included in analysis

[b]Incomplete/Excluded: subjects who exited the study, did not contribute minimum blood samples and were not included in analysis. Baseline demographic data were available on 119 of 127 incomplete/excluded dyads.

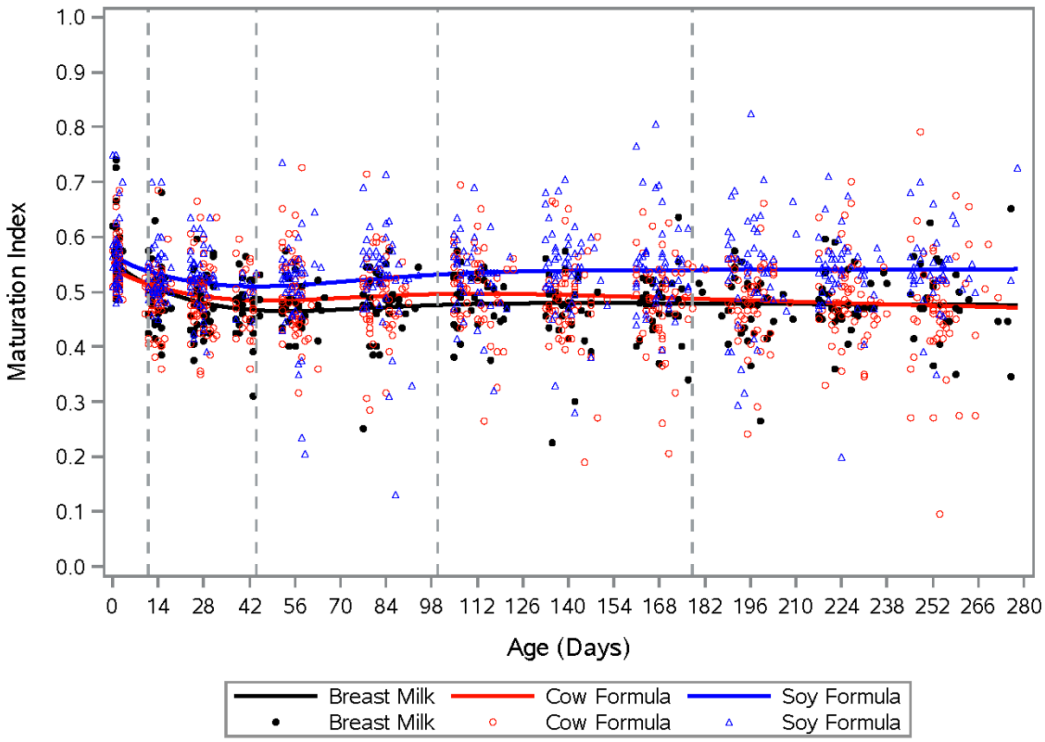
[c] X^2 p- values are from exact tests based on likelihood ratio statistics. Entry between Cow and Soy columns compares those two feeding groups; entry under Breast compares the breast-fed group with the cow- and soy-formula-fed groups combined; entry under Incomplete/Excluded compares Total Complete (included IFED subjects who completed study) to Incomplete/Excluded (subjects that did not complete the study).

[d] p-values are from t-tests of contrasts comparing means in a one-way analysis of variance. Entries are as described for X^2 p- values.

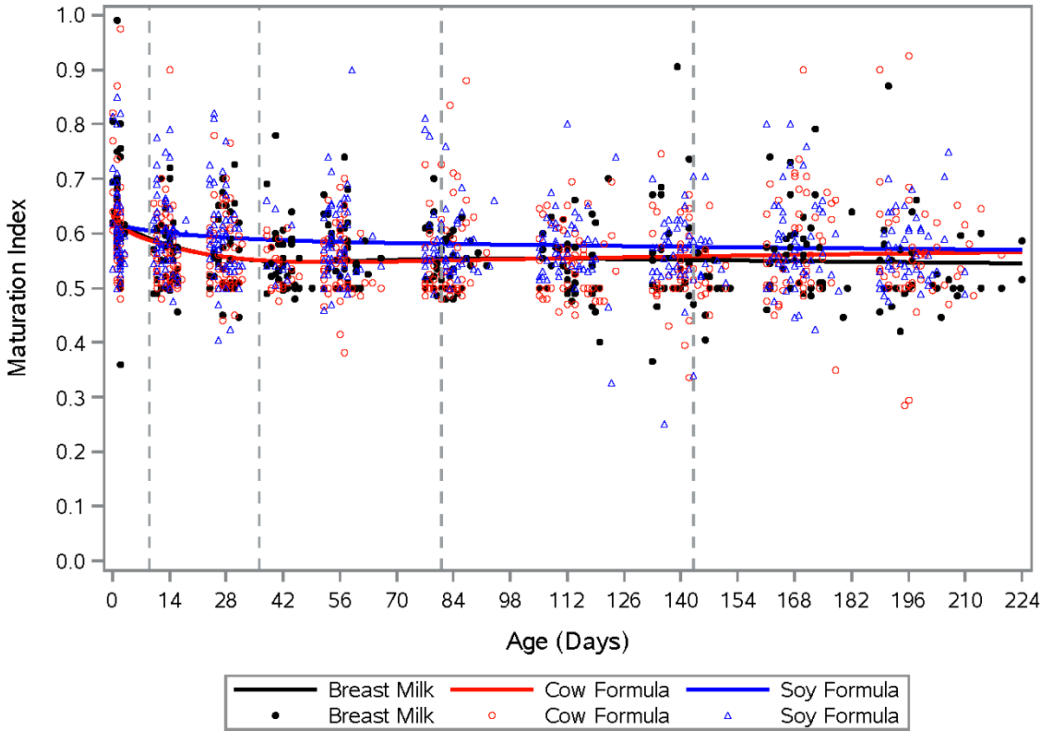
[e] Neonatal weight is based on birth visit weight measurement, conducted within first 72 hours after birth. Neonatal weight is reported for 117 incomplete/excluded subjects.

Supplemental Figure 1: Absolute trajectories for maturation index: observed values and fitted trajectories versus age by feeding group. Panels: a) girls; b) boys. Vertical reference lines indicate location of knots for fitted spline. Though not explicitly depicted, each girl contributes up to 12 measurements through time; each boy contributes up to 10.

a) Girls

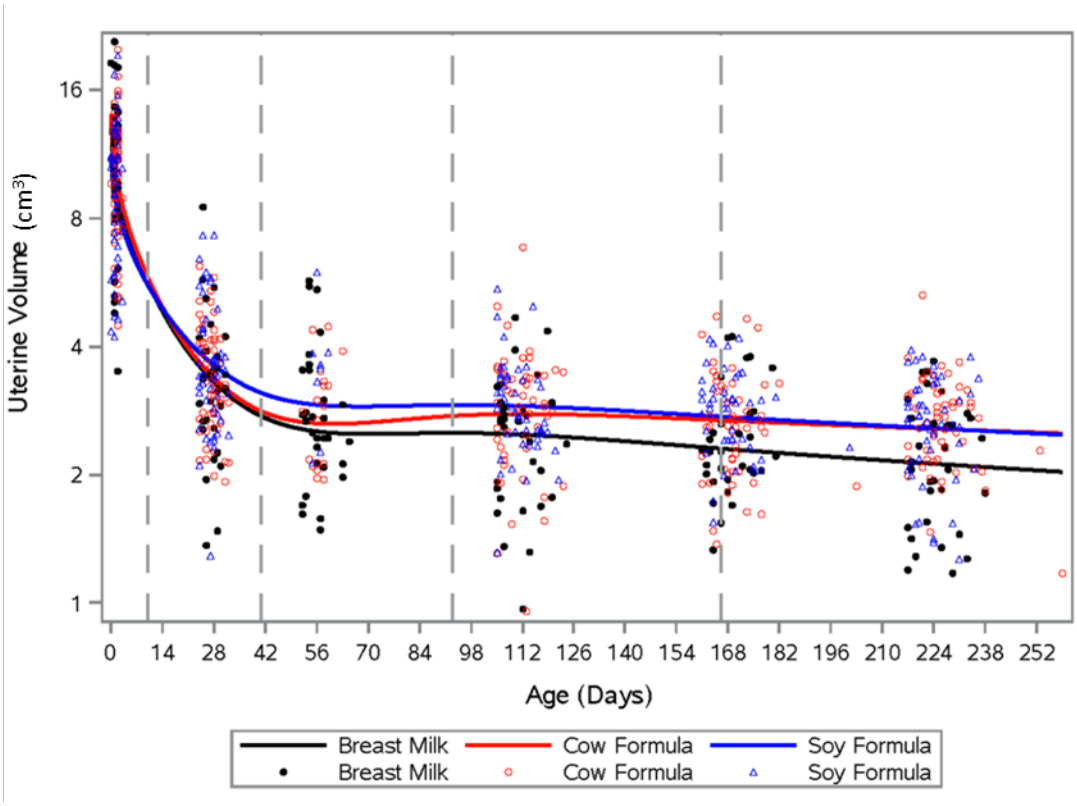


b) Boys

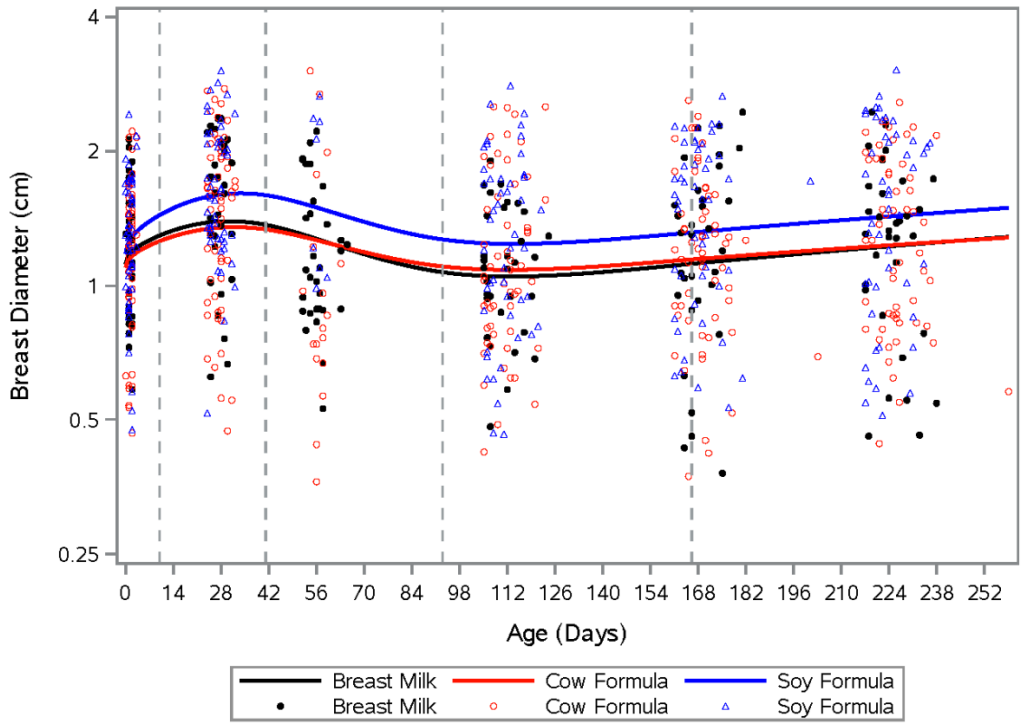


Supplemental Figure 2: Absolute trajectories for organ size: observed values and fitted trajectories versus age by feeding group. Panels: a) uterine volume in girls; b) breast-bud diameter in girls; c) breast-bud diameter in boys. Vertical reference lines indicate location of knots for fitted spline. Though not explicitly depicted, each girl contributes up to six measurements through time and each boy contributes up to five.

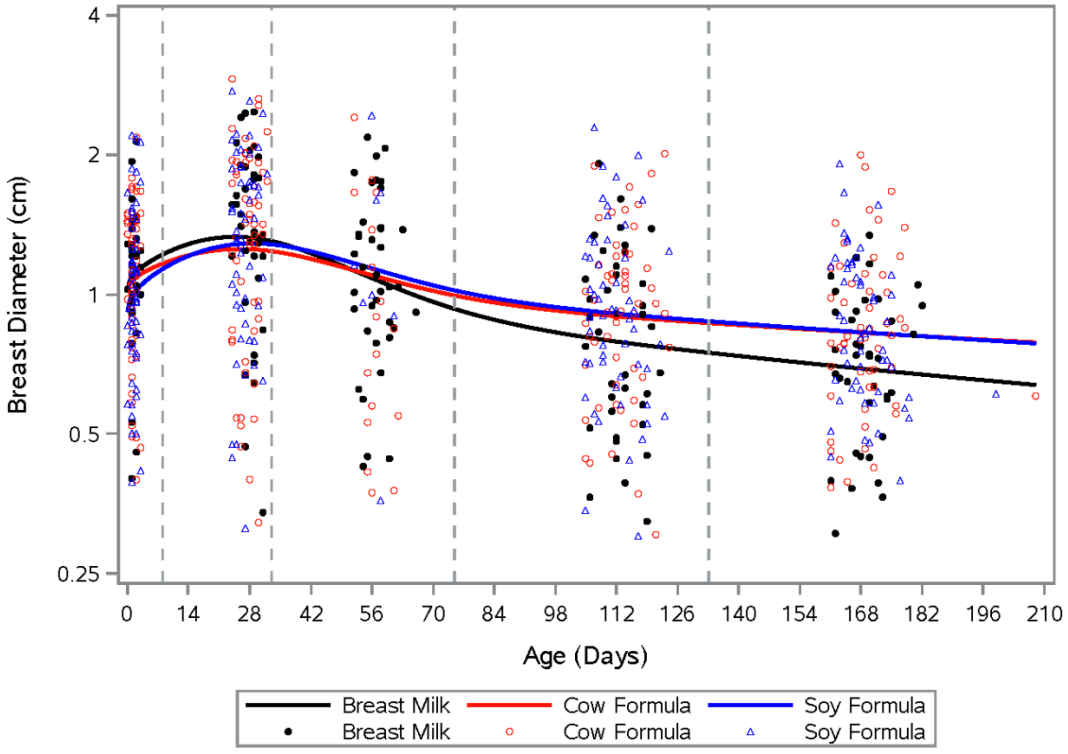
a) Uterine volume



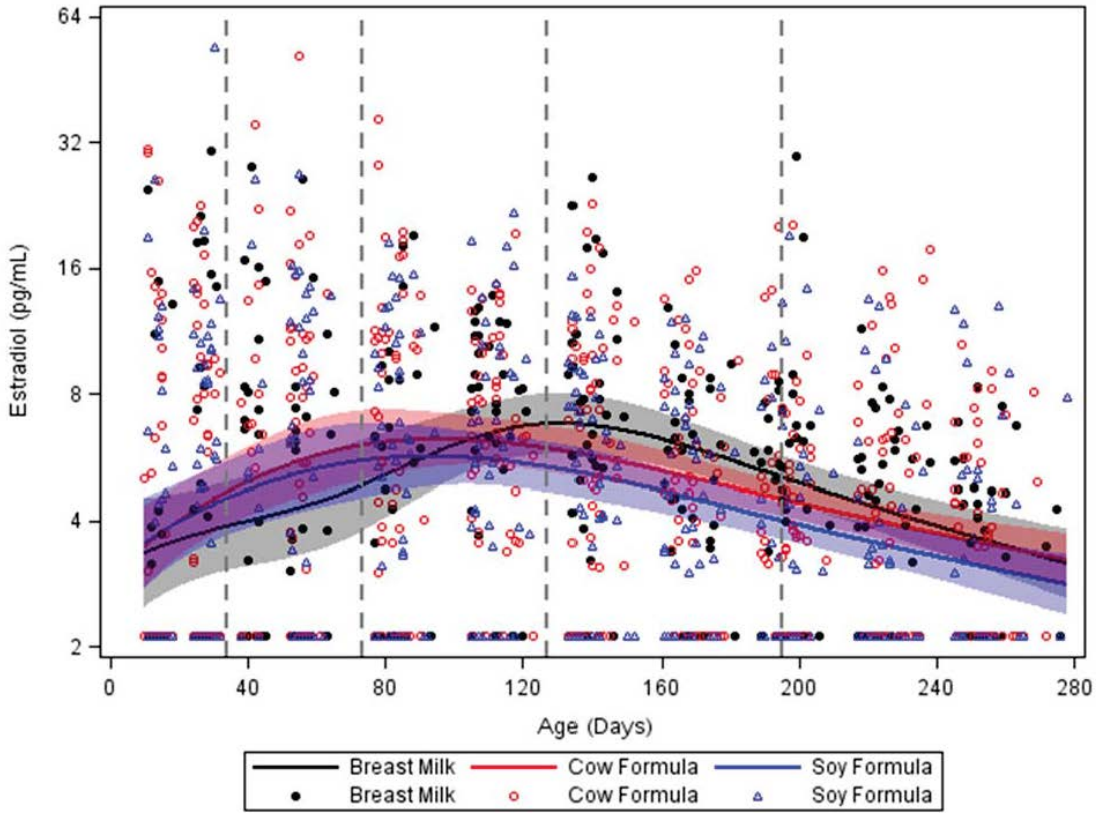
b) Breast Bud Diameter, Girls



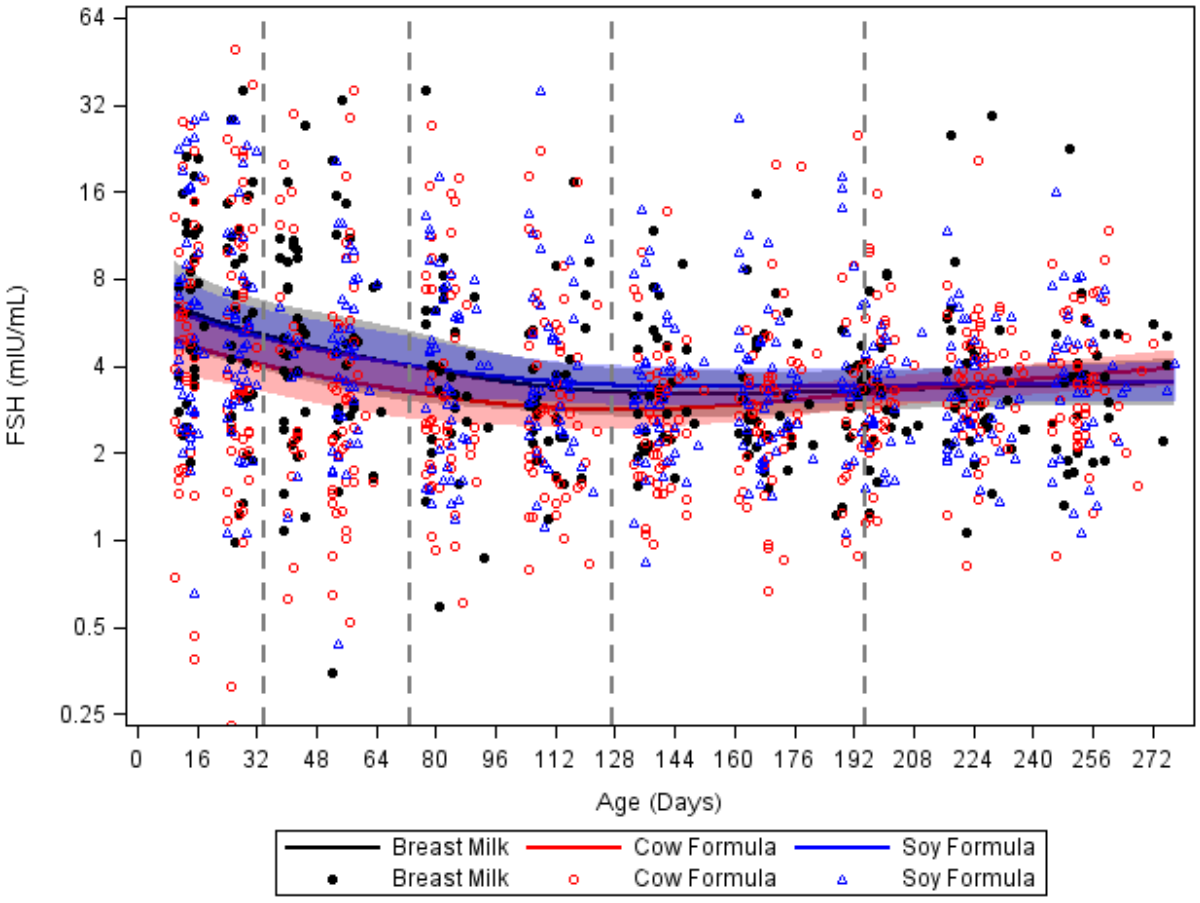
c) Breast Bud Diameter, Boys



Supplemental Figure 3. Observed E2 values and fitted E2 trajectory versus age by feeding group in girls. Colors code feeding groups: black, breast milk; red, cow-milk formula; blue, soy formula. Values below the limit of detection (2.99 pg/mL) are plotted at the limit of detection divided by the square root of 2. Shaded bands represent 95% pointwise confidence limits for each feeding-group-specific trajectory. Vertical reference lines indicate location of knots for fitted spline. Each subject contributes up to 11 measurements through time (not depicted). Soy formula vs. cow-milk formula p=0.44; soy formula vs. breast milk p=0.02; breast milk vs. cow-milk formula p=0.17.



Supplemental Figure 4. Follicle stimulating hormone (FSH): observed values and fitted trajectories versus age by feeding group (girls only). Colors code feeding groups: black, breast milk; red, cow-milk formula; blue, soy formula. Shaded bands represent 95% pointwise confidence limits for each feeding-group-specific trajectory. Vertical reference lines indicate location of knots for fitted spline. Each girl contributes up to 11 measurements through time (not depicted).



205 References

206

207 1. Kaplan SL, Edgar JC, Ford EG, Adgent MA, Schall JI, Kelly A, Umbach DM, Rogan WJ, Stallings VA,

208 and Darge K. Size of testes, ovaries, uterus and breast buds by ultrasound in healthy full-term

209 neonates ages 0-3 days. *Pediatr Radiol.* 2016;46(13):1837-47.

210 2. Parker DL, Rybak ME, and Pfeiffer CM. Phytoestrogen biomonitoring: an extractionless LC-

211 MS/MS method for measuring urinary isoflavones and lignans by use of atmospheric pressure

212 photoionization (APPI). *Anal Bioanal Chem.* 2012;402(3):1123-36.

213

214