EARLY LIFE SOY EXPOSURE AND CHILD DEVELOPMENT: AN ASSESSMENT OF LANGUAGE ACQUISITION, PLAY BEHAVIOR AND TIME-TO-MENARCHE

Margaret Adrian Adgent

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology.

Chapel Hill
2010

Approved by:

Julie Daniels

Walter Rogan

Linda Adair

Anna Maria Siega-Riz

Lloyd Edwards

ABSTRACT

MARGARET A. ADGENT: Early Life Soy Exposure and Child Development: An Assessment of Language Acquisition, Play Behavior, and Time-to-Menarche
(Under the direction of Julie Daniels)

Soy isoflavones are weak estrogenic compounds contained in products derived from soybeans, including soy-based infant formula. Exposure to these compounds in infancy may have lasting effects on later neurological and reproductive development and function. This study examined the association between early life soy-based feeding and developmental outcomes, including language acquisition, gender-role play behavior and time-to-menarche. Subjects were participants in the Avon Longitudinal Study of Parents and Children (ALSPAC). Subjects were classified into mutually exclusive infant feeding categories: *early soy, late soy, primarily breast,* and *early formula* (referent). Language acquisition, measured as word comprehension and production, was assessed using the MacArthur Communicative Development Inventory (MCDI) at 15 and 24 months of age. Gender-role play behavior was assessed using the Preschool Activities Inventory (PSAI) at 42 months of age. Time-to-menarche was assessed by self-report of age at menarche between ages 8 and 14.

Using generalized estimating equations (n = 3,384 boys; 3,176 girls), a small, imprecise increase in both word comprehension and word production was observed in girls with *early soy* exposure ($\beta_{\text{comprehension}}$ = 0.87, 95% CI: 0.36, 1.38; $\beta_{\text{production}}$ = 1.46, 95% CI: 0.56, 2.29) over time, as compared to the referent. No association was observed among boys.

Likewise, no association between PSAI score and infant feeding method was observed in boys, using multivariable linear regression models. However, the mean PSAI score among *early soy* exposed girls was slightly higher than the referent (β = 2.68, 95% CI: 0.20, 5.15), indicating slightly masculinized behavior.

Time-to-menarche was assessed in 2,884 girls using Cox proportional hazards modeling. The rate of menarche in early adolescence (before age 12.5) increased by 42% in the *early soy* fed girls (Hazard Ratio: 1.42, 95% CI: 0.92, 2.20). There was no association in later adolescence.

In this study, early life exposure to soy products was associated with slight, imprecise associations with developmental outcomes in girls, but not boys. Interpretation was limited by a small number of *early soy* exposed subjects. Despite the imprecision, similar patterns of association among girls in all outcomes support the need for additional studies to replicate these novel findings.

To my parents, Tony and Mary Beth, to my supportive and patient husband, Allen, and to the growing miracle we are waiting to meet.

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Julie Daniels for her exemplary advising and guidance throughout my experience at the University of North Carolina at Chapel Hill, and for her support in the completion of this project. I would also like to acknowledge the other members of my dissertation committee for their thoughtful contributions, UNC Department of Epidemiology faculty for providing quality training, and my fellow classmates for their support and company over the years.

In addition, I am extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council (Grant ref: 74882), the Wellcome Trust (Grant ref: 076467) and the University of Bristol provide core support for ALSPAC.

This research was specifically funded by the Training Program in Reproductive, Perinatal and Pediatric Epidemiology (NICHD/NIH (T32HD052468-01A2), 2008-2013).

TABLE OF CONTENTS

LIST OF TABLES……ix

LIST OF FIGURES……xi

LIST OF ABBREVIATIONS……xii

Chapter

I. INTRODUCTION……1

References……..3

II. REVIEW OF LITERATURE……4

a. Background: Soy Based Infant Formula……..4

b. Background: Sexually Dimorphic and Reproductive Outcomes…….14

i. Language Acquisition and Gender Role Behavior…….. 15

1. Background……..15

2. Critical Review of Soy Based Infant Formula Exposure and Cognitive and Behavioral Outcomes…… 25

3. Significance……..27

ii. Time-to-Menarche…….……...29

1. Background…….……29

2. Critical Review of Soy Based Infant Formula and Pubertal Onset……33

3. Significance……...37

References……..39

Figures and Tables……..51

III. SPECIFIC AIMS AND HYPOTHESES……... 53

IV. METHODS…… 55

a. Overview of The Avon Longitudinal Study of Parents And Children (ALSPAC)…… 55

b. Methods: Aim 1…….. 60

i. Hypothesis…….. 61

ii. Study Sample…….. 61

iii. Exposure, Outcome and Covariate Assessment……. 63

iv. Analysis…….. 73

v. Sample Size and Power……76

c. Methods: Aim 2……... 77

i. Hypothesis……... 77

ii. Study Sample…….. 77

iii. Exposure, Outcome and Covariate Assessment……. 78

iv. Analysis…….. 80

v. Sample Size and Power…….. 82

References…… 84

Figures and Tables……86

V. AIM 1.1 SUMMARY……96

References……..103

Figures and Tables…..…… 105

VI. AIM 1.2 MANUSCRIPT……... 111

References……..122

Figures and Tables…….125

VII. AIM 2 MANUSCRIPT……..130

References……..145

Figures and Tables…….149

VIII. CONCLUSIONS……... 156

a. Summary, Aim 1……. 157

b. Summary, Aim 2……. 160

c. Strengths…….. 162

d. Limitations……... 165

e. Implications……. 167

f. Future Directions…… 168

References……. 169

APPENDIX I: ADDITIONAL TABLES……. 170

Tables……171

Discussion…….177

LIST OF TABLES

Table

2.1. Isoflavone Concentrations in Infant Biological Samples (Adapted from Cao et al., 2009)…….. 52

4.1. Number of Respondents Per Questionnaire, According to ALSPAC Documentation……..……91

4.2. Power estimations for mean word comprehension and production differences…...…….. 92

4.3. Power estimations for mean PSAI score differences…… 93

4.4. Time-dependent hazard info for log-rank survival test…….. 94

4.5. Log-rank test of survival in two groups, simulation with specified rates (unequal n's)……...…….. 95

5.1. Study Sample Characteristics…… 106

5.2. Mean Word Comprehension MCDI Scores……107

5.3. Mean Word Production MCDI Scores……..…….. 108

5.4. Mean response estimates for MCDI word comprehension from 15 to 24 months…….109

5.5. Mean response estimates for MCDI word production from 15 to 24 months……..…. 110

6.1. Characteristics of study sample (n = 7,003) and eligible ALSPAC subjects (n = 12,931)…….127

6.2. Crude mean Preschool Activities Inventory (PSAI) (mean (SD)) scores and regression estimates (β (SE)) for exposure groups and select categorical and continuous covariates…….....128

6.3. Adjusted [a] change in mean Preschool Activities Inventory (PSAI) scores for boys and girls……. 129

7.1. Characteristics of eligible[a] ALSPAC study sample (n = 5,230), distinguished as those included in the present analysis (n = 2,884) and those excluded for missing exposure or outcome data (n = 2,346)……152

7.2. Adjusted[a] hazard ratios and median time-to-menarche estimates…….. 153

7.3. Proportion of censoring in study sample: lost to follow up (*LTF*) and administrative censored subjects [n (%)], by feeding group……154

7.4. Sensitivity analysis hazard ratios and median survival times……155

A.1. Age of introduction of soy product in early soy exposure group (n = 182)…...171

A.2. Covariate assessment by feeding group…… 172

A.3. Girls' BMI Z-scores for age, by feeding group……175

A.4. Distribution of covariate characteristics for lost to follow up subjects, across feeding group categories, and stratum-specific median follow up times……. 176

LIST OF FIGURES

Figure

2.1. Chemical Structure of 17β-estadiol, Genistein and Daidzein (Aglycone)…....51

4.1. Exposure Characterization……..86

4.2. 'Early Soy' Exposure Criteria: 4 to 6 Month Exposures……..……87

4.3. Conceptual diagram of association between infant feeding and language acquisition, noting confounding elationships…….....88

4.4. Conceptual diagram of association between infant feeding and gender-role play behavior, noting confounding relationships……89

4.5. Conceptual diagram of association between infant feeding and time-to-menarche, noting confounding relationships……..…… 90

5.1. Change in mean word comprehension and production over time, by feeding group…….105

6.1. Exposure Characterization for Infant Feeding Group..……...125

6.2. Distribution of PSAI Scores, by Gender and Feeding Group…….……126

7.1. Exposure Characterization……..149

7.2. Crude (a) and Adjusted (b) Kaplan Meier survival curves, by feeding group…150

7.3. Period specific hazard ratios by feeding group (*early formula* = referent).……151

LIST OF ABBREVIATIONS

ALSPAC: Avon Longitudinal Study of Parents and Children

BMI: Body Mass Index

CLR: Confidence limit ratio

CMF: Cow's milk formula

HR: Hazard Ratio

MCDI: MacArthur Communicative Development Inventory

PSAI: Preschool Activities Index

SBF: Soy based formula

I. INTRODUCTION

Soy-based infant formula (SBF) is a commonly used alternative to cow's milk based infant formula, particularly in instances of milk intolerance or preference for a vegan diet. It accounts for approximately 20% of the infant formula sold in the United States, and 7% in the United Kingdom (1, 2). SBF has been shown to be nutritionally adequate for term infants (2). However, it also contains high levels of phytoestrogens, plant compounds with structural and functional similarity to 17β-estradiol that may have long term effects on developing infants. Infants exclusively fed SBF are thought to receive doses of these compounds at approximately 4.5-8.0 mg/kg bodyweight, with plasma concentrations an order of magnitude higher than adults on a diet containing soy(3). Since endogenous estrogen levels are typically low during infancy, exposure to these estrogen-like compounds is of concern.

Animal models show that the phytoestrogens found in soy are capable of inducing disruptive effects on the endocrine system. Endocrine disrupting chemicals (EDCs) have been associated with a wide range of developmental outcomes. For example, EDCs have been shown to influence sexually dimorphic brain development in animals and consequently sex-specific animal behavior (4). Additionally, a popular hypothesis suggests that trends in early onset of puberty may be, in part, due to early life EDC exposure (5).

Few longitudinal studies of early life soy exposure have been conducted in human populations beyond infancy. There is a particular paucity of data with respect to neurological

and reproductive outcomes, which may be sensitive to changes in normal early life hormone concentrations (6). Therefore, the overall long term safety of SBF is unclear.

This epidemiologic investigation aims to assess the association between soy-based infant formula and milk use in infancy and hormonally sensitive outcomes related to neurological and reproductive development. Specifically, this study will examine the effects of SBF on 1) gender-specific language acquisition rates and gender-role behavior in preschool aged children, and 2) age at menarche in adolescent females. This investigation will be carried out in the Avon Longitudinal Study of Parents and Children (ALSPAC), a pregnancy cohort of over 14,000 live births recruited in 1991-1992 in the United Kingdom, and followed to the present day. Because SBF use is a relatively common source of infant nutrition in the United States, this study will be a much needed contribution to the understanding of effects of exposure to dietary estrogens on early child development.

REFERENCES

1. Essex C. Phytoestrogens and soy based infant formula. BMJ. 1996 Aug 31;313(7056):507-8.

2. Bhatia J, Greer F, American Academy of Pediatrics Committee on Nutrition. Use of soy protein-based formulas in infant feeding. Pediatrics. 2008 May;121(5):1062-8.

3. Setchell KD, Zimmer-Nechemias L, Cai J, Heubi JE. Exposure of infants to phyto-oestrogens from soy-based infant formula. Lancet. 1997 Jul 5;350(9070):23-7.

4. Li AA, Baum MJ, McIntosh LJ, Day M, Liu F, Gray LE,Jr. Building a scientific framework for studying hormonal effects on behavior and on the development of the sexually dimorphic nervous system. Neurotoxicology. 2008 May;29(3):504-19.

5. Buck Louis GM, Gray LE,Jr, Marcus M, Ojeda SR, Pescovitz OH, Witchel SF, et al. Environmental factors and puberty timing: Expert panel research needs. Pediatrics. 2008 Feb;121 Suppl 3:S192-207.

6. Irvine CH, Shand N, Fitzpatrick MG, Alexander SL. Daily intake and urinary excretion of genistein and daidzein by infants fed soy- or dairy-based infant formulas. Am J Clin Nutr. 1998 Dec;68(6 Suppl):1462S-5S.

II. REVIEW OF LITERATURE

a. Background: Soy Based Infant Formula

Prevalence and Indications for Soy Based Infant Formula Use: Soy based infant formula (SBF) was first introduced in the United States in the early 1900s, and has since become a popular feeding alternative to breast milk and cow's milk formula (CMF). SBF currently accounts for approximately 20-25% of the formula market in the United States (1). Internationally, SBF did not become available until the mid-1970s, and is much less prevalent than in the US. For example, only 2-3% of infants (7% of formula market) in the United Kingdom were reportedly using SBF in the mid 1990s (2, 3). SBF use in Asian countries is not typical, despite adult diets high in soy foods (4).

Parents choose to feed SBF for a variety of reasons, ranging from medical advice to personal preference. Recent American Academy of Pediatrics guidelines have identified galactosemia and hereditary primary lactase deficiency, which are rare autosomal recessive traits, as the only medical conditions that can be successfully controlled with an SBF diet (1). Historically, however, SBF has been recommended in instances when an infant's nutritional needs could not be met by breast or cow milk (5). Such instances might include colic, mild lactose intolerance or cow's milk allergy. Rash and family history of atopy are also potential motivations for SBF use.

However, SBF is not effective in the treatment or prevention of these most of these medical conditions. Controlled trials comparing SBF and CMF have not demonstrated any

benefit of soy on the relief of colic symptoms, including excessive spitting, vomiting, and fussiness. Additionally, SBF has not been shown to be effective in the prevention of development of allergy, nor has it been shown to be a suitable alternative for infants already experiencing CMF allergy (1). 10 -14% of infants with cow's milk allergy are also allergic to soy (6, 7). SBF fed infants with family history of atopy are no less likely to develop atopic responses, such as cow's milk allergy or eczema, than infants fed CMF. Therefore, maternal perception of an infant health problem, or of the effectiveness of SBF treatment may drive the initiation or continuation of SBF use over actual medical advice (8-10).

Maternal preference for SBF may be influenced by a number of other factors as well. These factors include vegan diet, perceived health benefit of soy products in adult populations, peer pressure from other mothers, or experience from previous births. Since most studies on SBF have been either controlled trials or small studies in homogenous communities, it is unclear if there are other socioeconomic, cultural or demographic characteristics related to preference of SBF over other types of formula.

Composition of Soy Based Infant Formula, Nutritional Content: Early formulations of SBF met concerns regarding digestibility and nutrient availability, and as a result, SBF has undergone numerous modifications over the past 50 years to enhance quality and nutritional content. In the mid-1960s, soy protein isolate, a form of soy protein that is highly digestible and high in essential amino acid content, replaced soy flour as the primary protein source in SBF. Around 1980, micronutrient availability was improved through mineral fortification and reduction of phytates, compounds capable of interfering with mineral absorption and iodine metabolism. Commercially available SBFs have since met American Academy of

Pediatrics recommendations for full-term infant feeding and U.S. Food and Drug Administration quality standards (1, 5, 11).

SBF composition has been well characterized in order to ensure nutritional adequacy, although content likely varies slightly between brands and batches because it is a complex mixture (12). In addition to soy protein isolate, L-methionine, L-carnitine, and taurine supplementation also contribute to SBF protein content (5). Fat content is derived primarily from vegetable oils, with routine addition of docosahexaenoic (DHA) and arachidonic (AA) acids. In lieu of lactose found in cow's milk, carbohydrate sources in SBF include corn maltodextrin, corn syrup solids, tapioca starch, and sucrose (1, 5).

Composition of Soy Based Infant Formula, Phytoestrogen Content : SBF also contains high levels of phytoestrogens, or plant based estrogen-like compounds, that are naturally found in soybeans. Specifically, SBF contains the compounds genistein and daidzein, which belong to a particular subclass of phytoestrogens known as isoflavones. SBF isoflavone content varies slightly between brand, formula type (powder vs. liquid) and country of origin. These compounds are typically measured in formula as aglycones (unconjugated genistein and daidzein) (Figure 2.1), and conjugates (daidzin, genistin, malonyl- and acetyl-glycosides), with the conjugates comprising the vast majority of formula composition.

Total isoflavone concentrations in common U.S. brands of SBF range from 32- 46 mg/L (13). The conjugate genistin comprises the highest proportion of the total concentration in all brands and types, followed by daidzin. Powdered formulas tend to have higher isoflavone content than liquid formulas. Soy formulas obtained between March 1996 and July 1997 from the U.K. had similar, but slightly lower isoflavone content ranging from 18-41 mg/L (14). Soy milk and other soy drinks, which also may be administered to infants, have even

higher isoflavone content than SBF, at 22.9 - 71.5 mg/L, based on commercial samples purchased in Australia in 1995 (15).

Isoflavone concentrations in SBF are markedly higher than in other sources of infant nutrition. Breast milk isoflavone content is minimal in comparison to SBF, at ~5 ng/ml in healthy women. This concentration increases approximately 10 fold following a soy dietary challenge (13, 16). Isoflavones genistein and daidzein have been detected at 0.1 – 2.0 ng/ml in cow's milk, while equol, an important metabolite of daidzein, has been found at slightly higher concentrations (5-30ng/ml) (17). Insensitive methods have prevented investigators from detecting isoflavones in cow's milk *formula* (13), but cow's milk formula isoflavone concentrations are presumed to be low.

It is important to note that the aglycone forms of these compounds are the more biologically available and biologically active isoflavone moieties, and yet typically only account for 3-5% of isoflavone content in SBF (12). The vast majority of ingested isoflavones are in the conjugate form. This brings into question, then, whether the majority of isoflavone intake is biologically relevant or capable of inducing a biological response. A more detailed discussion of isoflavone intake, metabolism and biologic activity, particularly with respect to aglycone and conjugated forms, will be provided in a later section. .

Biologic Availability of Isoflavones: Aglyconic and conjugated isoflavones differ in their initial capacity for digestive absorption into the body. While aglyconic isoflavones are rapidly absorbed through the intestine, conjugated isoflavones contain a bulky glycosidic bond that prevents such transport. Glycosidases must cleave the glycosidic bonds in order to facilitate absorption. This cleavage process readily occurs in both adult and infant humans (18, 19).

Bacteria in the gut may also further metabolize genistein and daidzein following glycoside bond cleavage. Metabolites of this process include dihydrogenistein and *O*-desmethylangeolensin, which are derived from genistein, and dihydrodaidzein, tetrahydrodaidzein, and equol, which are derived from daidzein (18). However, it is not likely that such bacteria are present in an immature infant digestive tract, so exposure to these metabolites is probably very limited following exposure to SBF.

Once absorbed, isoflavones are metabolized in the liver into glucuronide and sulphate conjugates. These metabolites are then excreted in urine. The half-life for genistein and daidzein detection in adult plasma is approximately 7-8 hours (20).

Detection of high levels of genistein and daidzein metabolites in both urine and serum of SBF fed infants provide evidence that infants efficiently absorb, metabolize and excrete soy isoflavones (Table 2.1). Infants exclusively fed SBF receive total isoflavone doses ranging from 4.5 – 8.0 mg/kg bodyweight, and have plasma isoflavone concentrations an order of magnitude higher than adults on a diet containing soy (21). Urinary concentrations of total genistein and daidzein have been shown to be 500 times the concentration detected in CMF fed infants (22). Total isoflavone concentrations in SBF fed infants are 13,000- 22,000 times the level of early life endogenous estradiol (13, 21).

It is important to note that low levels of equol have been detected in some CMF fed infants. In some instances, equol was detected in plasma of CMF infants at concentrations slightly higher than in SBF fed infants (21) (Table 2.1). This may be due to the presence of low levels of equol isomers in cow's milk (17, 23). Since equol has biologic activity that is similar to genistein (described below), it is important to consider how equol may influence a

comparison between CMF and SBF fed infants, as well as the interpretation of experimental animal studies of isoflavones.

Biologic Activity of Isoflavones: Much of the concern regarding genistein and daidzein exposure is attributed to their structural and functional similarities to 17β-estadiol (Figure 3.1). Typically, 17β –estradiol binds to estrogen receptors (ER), which ultimately allows for controlled gene expression and cell-specific protein synthesis (24). Estrogen receptors are distributed in tissues of the male and female reproductive system, breast, bone, cardiovascular system, and the hypothalamus region of the brain (25, 26). Estradiol-ER binding is thought to be necessary for the proper development and function of these various organ systems. However, like estradiol, aglyconic genistein and daidzein are also capable of binding to ERs, and are thus capable of disrupting normal developmental processes.

Comparatively, genistein has a higher ER binding affinity than daidzein, but one that is100 fold less than estradiol, so it is considered a "weak" estrogen (27). Both isoflavones have ER binding affinities that are also markedly less than the synthetic estrogen, diethylstibesterol (DES), but are more potent than other chemicals with known endocrine activity, such as bisphenol A and methoxychlor (28). In serum, these isoflavones have been shown to have greater access to ER binding in comparison to estradiol, since estradiol easily binds to serum proteins, such as α-fetoprotein, and isoflavones do not (29, 30). Estrogen receptors have been identified in two distinct sub-types, ERα and ERβ. Aglyconic genistein and daidzein can bind to both sub-types, although both have a much higher affinity for ERβ. Genistein and daidzein in the conjugate form are not capable of such activity, since the glycosidic bond occurs at the site of estrogen binding.

The activity of isoflavones appears to be at least partially dependent on the concentration of endogenous estradiol. In a high estradiol environment, these isoflavones tend to act as estrogen antagonists (inhibitors of estrogenic activity), whereas in a low estrogen environment, they can act as estrogen agonists (estrogen mimics) (18, 27, 28, 31). Given the capacity for such activity, these isoflavone compounds are appropriately categorized under a wider classification of exogenous hormonally active compounds known as endocrine disrupting chemicals (EDCs). A more detailed discussion of how isoflavone ER binding activity may affect brain and reproductive system development will be given in later sections.

Isoflavones may also be capable of affecting estrogen metabolism, as well as the metabolism of other steroid hormones, through a gene expression mediated mechanism. Dietary genistein was recently shown to reduce mRNA associated with the gene *Hsd3b* [7-Dehydrocholestrol reductase], which plays an important role in the synthesis of steroid hormones. Likewise, a gene associated with the degradation of these hormones, *Akr1d1* [Aldo-keto reductase family 1, member D1], was significantly upregulated by genistein and daidzein. Genistein was also shown to decrease serum cholesterol levels in rats, which may have additional implications for steroid hormone synthesis (32).

Isoflavones are also capable of exerting effects beyond those associated with steroid hormones. Genistein is a potent inhibitor of tyrosine kinase (21, 33), which may have estrogen-independent implications for cell proliferation and intracellular signaling pathways. Isoflavones are also capable of binding to peroxisome proliferator-activated receptors (PPAR), which are important in insulin regulation and adipogenesis(34-36).

Genistein has also been shown to effect thyroid function. Tyrosine peroxidase, which is required for thyroid hormone synthesis, is inactivated in rats exposed to dietary genistein

(37). Additionally, genistein can act as a T_3 antagonist(38). These findings have been documented at concentrations relevant to dietary soy exposure in humans, and are particularly relevant here since thyroid hormones, including T_3, are important for many developmental processes, including brain, ear, and bone development (38, 39).

Biological Activity in Humans: The biological activity of isoflavones has been explored extensively in experimental animal and *in vitro* models. Relevant findings have been described, in part, above, and additional animal studies specifically related to sexually dimorphic learning, memory, behavior, and reproductive development will be discussed in later sections.

Despite extensive animal evidence, however, the actual biologic activity of circulating isoflavones in SBF fed infants is unclear. Exposure to isoflavones has been well documented, but studies linking these exposures with sensitive developmental outcomes are limited. Approaches to clarify their biologic activity have included detailed biomarker measurement, and both observational and clinical studies of effects exposure to SBF. Additional evidence of biologic activity in humans can also be derived from studies of adults exposed to dietary soy products.

Given the current understanding of isoflavone ER binding potential, the extent of isoflavone biologic activity in infancy may be informed through the measurement of aglyconic, as opposed to conjugated, isoflavones *in vivo*. Accordingly, in a study of genistein exposure in rats, 31% of circulating isoflavones were aglyconic (40). This provides evidence that a relatively high level of active compound is circulating and available for ER binding. In contrast, Huggett et al., did not detect any circulating aglyconic isoflavones in plasma of four SBF fed infants. Instead, isoflavone composition consisted of glucoronide and

sulfide metabolites, with unknown biological significance (41). Other isoflavone exposure studies, such as those presented in Table 2.1, did not sufficiently document the relative concentrations of aglycone and conjugate forms. Therefore, it is not possible to derive substantial conclusions regarding biologic activity in infants based on the current biomarker literature.

Numerous clinical trials have examined the overall safety of SBF use in infancy since the mid-1960s. While these studies do not address the mechanistic behavior of SBF isoflavones, they can suggest biologic activity at the organism level. These studies have examined growth and bone mineralization in SBF fed infants as compared to CMF fed infants, gastrointestinal or allergic response to soy, responses to immunizations, and select childhood and early adult outcomes (1, 5, 42-44). SBF feeding has been shown, in general, to allow for normal growth and bone development in term infants, and normal immune response to immunizations. Nutritional recommendations based on this literature do specify that SBF is not appropriate for preterm infants. SBF may not support adequate growth and bone development in this population.

Distinct from these gross measures of development, such as growth, subtle effects on the endocrine system have been observed following exposure to SBF. For example, children with congenital hypothyroidism on a diet of SBF require higher than normal T_4 treatment dosages to regulate TSH levels (45-47). For these cases, SBF is likely inducing increased stool frequency, and thus a decreased capacity for T_4 absorption. An alternate explanation is that soy isoflavones may be adversely affecting thyroid function via thyroid peroxidase inhibition (37). Additionally, SBF may also influence cholesterol production in infants, as

infant urinary isoflavone concentrations were shown to be inversely related to cholesterol synthesis rates following exposure during the first four months of life (48).

In some instances, these subtle endocrine effects have not manifested until later in childhood. For example, Chinese children who consumed SBF at either 4-6 months of age, or 7-12 months of age were shown to have increased odds for Type I diabetes, as compared to CMF fed infants (ORs: 2.0 (95% CI: 1.1-3.4); 1.5 (95% CI: 1.0 – 2.1), respectively) (49). In an earlier study, Fort et al. reported a higher proportion of SBF use in infancy among children with Type I diabetes than in controls (50). Studies of other long term effects of SBF exposure, such as early life sex hormone status, sexually dimorphic cognition and behavior, and reproductive development are very limited, and will be discussed in later sections.

Studies on the effects of soy products in the adult diet are also informative with respect to the potential biologic activity of soy isoflavones, despite obvious discordance in exposure timing. With an understanding that adult and infant metabolisms can differ in some respects, such as the capacity to metabolize daidzein into equol, these adult studies can provide some insight into the overall biological activity of soy isoflavones in humans, particularly with respect to sex hormone status. Accordingly, in a small study of premenopausal women, a diet supplemented with 45 g soy/day was shown to modify menstrual cycle characteristics, including suppression of mid-cycle hormone surges and delayed menstruation. Effects lasted for several months beyond the one month treatment period (51). Subsequent studies have reported similar findings, with a recent meta-analysis supporting that soy isoflavone exposure is associated with a reduction of circulating gonadotropins and increased menstrual cycle length in premenopausal women (52). In males, studies of the effects of adult exposure to soy protein on hormone levels are equivocal. Dietary soy consumption, as assessed by

questionnaire, has been associated with increased estradiol, and decreased testosterone among adult Asian men (53). However, concentrations of estradiol, testosterone, and other steroid hormones did not differ between men randomized to either lean meat or tofu diets (54), among men consuming daily soy milk (55), or on soy isoflavone supplements (56). No significant effects on adult testicular volume or sperm count have been noted following adult soy isoflavone exposure (56).

Data Gaps: The findings described above demonstrate that SBF is a potent source of soy isoflavone exposure in infancy, and that soy isoflavones have the biologic potential to interfere with normal developmental processes via endocrine disruption. However, to date, research of SBF exposure in infancy has primarily focused on nutritional adequacy and allergic response to soy. Little research has been done on the effects of soy on the developing neurologic and reproductive systems – two endpoints that are heavily influenced by endocrine function, including sex steroid concentrations. While professional pediatric guidelines have noted that SBF is an adequate source of nutrition for term infants, more research is needed to properly characterize risks to child development, particularly as they relate to long term health, following early life exposure to soy products.

b. Background: Sexually Dimorphic and Reproductive Outcomes

Early life is a period when exposure to endogenous estrogens is typically low (57). Since SBF dramatically increases the level of exposure to environmental estrogens with potential for endocrine disrupting activity, the present study intends to assess the relationship between SBF exposure and subtle changes developmental outcomes. The discussion below describes normal patterns and typical predictors of these developmental processes, followed

by a presentation of experimental and epidemiologic data relevant to the effects that EDCs, and isoflavones specifically, may have on these normal processes.

i. *Language Acquisition and Gender-Role Behavior*:

1. Background

Typical language acquisition rate and gender-role behavior characteristics are well described in child development literature. Both of these outcomes are largely influenced by age and gender, as well as a range of biological, social and environmental factors.

Language Acquisition: In a typical child, word comprehension begins around 9 months of age, and word production begins around 12 months of age. Following 12 months, a "fan effect" occurs in which rapid word producers and slow word producers begin to differentiate. At around16 to 18 months, a significant difference can be observed between males and females in both the number of words understood, and the number of words produced (58), with females consistently demonstrating higher word production than males (59-64). Language development continues with the initiation of word combinations around 18 to 20 months of age, and grammatical development around 24 to 30 months of age. By 30 months, the difference between rapid and slow producers begins to decrease, indicating a ceiling effect.

Language acquisition rates can be mediated by a number of influences in addition to age and gender. These may include social environmental factors, such as parental tendency to engage verbally with the child ("language input"), or presence of an older sibling. Physical factors such as oral development, hearing, and general health may influence language development. Chronic ear infections, for example, are thought to delay speech (61). Family history of speech problems, maternal vocabulary/education, non-English speaking

background (in English speaking countries), and temperament may also be predictive (64). The presence of other behavioral or cognitive development issues may delay speech as well, although they would not necessarily influence language development in a "normal" population.

Gender-Role Behavior: In the course of typical child development, differences in male and female behaviors are readily observable. While there is generally a fair amount of overlap between what is considered "male" and "female" behavior, one particularly prominent type of sex-specific behavior is a child's preference for certain toys or activities. Gender based differences in play activities can be apparent by as early as 12 months of age, and are clearly observable by age 3 years. At this time, it can be shown that boys tend to prefer items such as cars, trucks, and weapons. Girls prefer items such as dolls, dress-up toys, and domestic toys. Additionally, boys are also more likely than girls to engage in "rough and tumble" play. Children are also more likely to choose playmates of their same sex. Boys tend to play in large groups of other boys, while girls prefer to interact with only one or two other girls (65, 66).

Like language acquisition, gender-role behavior is influenced by a many factors, including both biology and social environment. Biologically, gender-role behavior has been shown to be influenced by early life brain development. A more detailed discussion of this developmental process is described below. With respect to the social environment, boys and girls learn to choose sex-specific toys simply because they are taught to use those toys. Influence from parents or older siblings may discourage "cross sex" play. Additionally, peer influence is also thought to contribute largely to gender-role behavior. Establishment of a

gender identity, or awareness of being a boy or a girl, will likely influence play choices, since children tend to model behaviors seen in other children of their same gender (65).

Biological Basis for Sexual Dimorphism in Language and Behavioral Development:

Both of the outcomes discussed above tend to show clear differences in performance and behavior on the basis of gender. The biological basis for observed sex differences in cognitive and behavioral development can be related to sexual differentiation of the brain in early development. Biologically, sex differences appear to be largely influenced by early life hormonal levels, regulated by the hypothalamic-pituitary-gonadal axis (HPG).

HPG activity is initiated by the central nervous systems signaling a release of gonadotropin releasing hormone (GnRH) from the hypothalamus. GnRH stimulates the pituitary to produce luteinizing hormone (LH) and follicular stimulating hormone (FSH) ("gonadotropins"), which then target the gonads, and initiate the release of sex steroids, including estrogen and testosterone (67). This series of events is mediated by negative feedback inhibition, where estrogen and testosterone at high enough concentrations act on the hypothalamus and pituitary to decrease gonadotropin secretion, and thus slow the production of estrogen and testosterone (68). "Normal" hormone concentrations vary throughout the full lifespan, as marked differences can be observed between the neonatal, childhood, reproductive and post-reproductive stages.

Hormone concentrations and bioavailability appear to influence sexually dimorphic brain development, and consequently, the manifestation of gender-specific learning and behavior characteristics. In developing males, circulating testosterone is converted to estrogen in the brain by the enzyme, aromatase. This estrogen, in turn, binds to estrogen receptors in the brain, and promotes synthesis of proteins that then allow for the development of male

neurological structures, and subsequently, behavioral characteristics (69). In females, however, testosterone is essentially absent. Furthermore, circulating estrogen is excluded from the brain because it tends to binds with α-fetoprotein, a steroid-binding protein in the serum (30). So, the estrogen-dependent brain development process that occurs in males does not occur in females. As a result, males and females have distinct, sexually dimorphic brains. The timing of these events span the late prenatal and early postnatal period in rats, but have been thought to occur primarily in the second trimester of gestation in the human (70).

Human infants are somewhat unique in that they also experience continued surges of sex hormones for the first several months of life, in addition to those seen in the prenatal stage. Male infants experience a testosterone surge that typically occurs between 1 and 5 months of age, a characteristic that is not observed in rodents and is shared only with other primate species (71). High levels of LH have also been detected in male infants through the first month (72), suggesting that testosterone production may be occurring through an LH mediated mechanism. The physiologic importance of the testosterone surge in both humans and non-human primates is unclear. Some evidence suggests that postnatal testosterone levels influence penile growth and later reproductive function (73, 74), but have little effect on sexually dimorphic behavior in infant rhesus monkeys (75-78). However, more recent evidence indicates that postnatal testosterone may play a role in postnatal brain organization in human infants, specifically with respect to language processing (79).

Female infants also experience low levels of certain hormones. Estradiol and FSH have been detected in 3-month old infant girls at levels comparable to those detected in early pubertal, premenarcheal females (age 6-12, Tanner Breast Stage I or II) (80, 81). As in males, the functionality of early life reproductive hormones are not well understood. However, their

presence demonstrates the potential for an additional or extended window of developmental susceptibility for endocrine related outcomes, such as brain development, beyond the prenatal period. Extended periods of susceptibility have been noted in other species as well. For example, prolonged postnatal exposure to exogenous estrogens has been shown to "defeminize" the networking of the HPG among ewe sheep, a species previously thought to experience sexual differentiation primarily in the prenatal period (82).

Influence of Endocrine Disruption on Language Acquisition and Behavior:

The effects of early life reproductive hormone concentrations and EDC exposure on sex-specific learning patterns and behaviors have been well studied in experimental animal models, and to a far lesser extent, in human populations. Such patterns and behaviors are, in fact, shown to be sensitive to early life hormone status. Accordingly, prenatal, neonatal, and long term exposures to SBF, genistein, and phytoestrogen-rich diets have been shown to elicit physiologic responses and behavioral changes that deviate from normal sexually dimorphic development.

The mechanism(s) by which EDC compounds may induce alterations in learning and behavioral development are unclear, but it is likely that EDCs disrupt the regulation of endogenous hormone concentrations via interference with the HPG axis. The neural components of the HPG axis are sexually differentiated by endogenous estradiol during the pre- and perinatal periods in rats (83, 84), so it is postulated that EDC exposure during these times could alter HPG organization. Both ERα and ERβ are present in the hypothalamus, making it a susceptible region to isoflavones and other disruptors (26, 85). In addition to perturbation of hypothalamus structure and function, isoflavones may also interfere with the

natural inhibition feedback mechanism of the HPG axis, and thus affect normal sex steroid production and concentrations.

Neonatal primates are ideal animals to study for early life endocrine disruption effects, since like humans, they experience neonatal hormone surges. Serum testosterone and LH levels at 3 months postpartum, for example, have been show to significantly decline following exposure to GnRH agonist in monkeys, due to blockage of the HPG axis pathway (73). Accordingly, in an early life feeding study of twin marmoset monkeys, twins that were fed SBF *ad-libitum*, much like human infants, had serum testosterone levels that were 55% lower than their CMF fed sibling at age 35-45 days ($p = 0.004$). Twelve out of 15 SBF fed monkeys had "low" testosterone levels (< 0.5 ng/ml), whereas this occurred in only 1 of the 15 CMF-fed monkeys ($p < 0.001$). In addition, an increase in the number of Leydig cells per testis was also observed among the SBF fed monkeys ($p < 0.001$) (86). Upon reexamination at age 3, testis weight, Sertoli and Leydig cell counts were elevated among SBF fed monkeys, where the highest Leydig cell concentrations were among the monkeys with the lowest testosterone concentrations, suggesting Leydig cell failure compensation (87). While this study addresses male reproductive health rather than neurodevelopment, it is important to note the effects that SBF has on early life hormone surges, as well as downstream effects of those surges in adulthood. Given the recognized importance of testosterone in sexually dimorphic brain development, this study suggests potential for overt effects of SBF on brain development.

In experimental rat models, the volume of the sexually dimorphic nucleus of the medial pre-optic area (SDN-POA) of the hypothalamus is often used as a marker of endocrine disruptive activity and changes in sexually dimorphic brain structure (69, 88). This brain

region is typically larger in males than in females, and is thought to play a role in male sex behaviors in rats (89-91). It expresses both ERα and β throughout the life span, and therefore is susceptible to isoflavone binding (26, 85). In developing males, testosterone that is converted to estradiol serves to maintain the SDN neurons throughout development. So, in the relative absence of estradiol, female SDN neurons undergo programmed cell death (apoptosis), which results in the characteristically smaller volume (92). However, exposure to testosterone, excess endogenous estradiol, or an exogenous estrogen-like compound may interrupt the apoptotic process to result in a "masculinized" female. A "feminized" male may result if an exogenous estrogen, acting as an ER antagonist, blocks endogenous estradiol from ER binding. Conversely, a "hyper-masculinized" male may result when an exogenous estrogen acts as an ER agonist (estrogen mimic), and any apoptosis is prevented more effectively than it would be in the normal male (69).

Studies examining the effects of isoflavones on SDN-POA volume in rats suggest that effects may be dose and time dependent (69). Larger SDN-POA volumes have been noted in female rats following postnatal exposure to high, but not low doses of genistein (93-95). However, a non-significant decrease was seen when rats were treated at similar doses only in the prenatal period (96), or across the prenatal and early postnatal period (97). Other studies show an SDN-POA volume decrease in exposed males, but not females, following a moderate, life-long dose of genistein (98). Male rats exposed to a phytoestrogen-rich diet for their first 80 days had significantly smaller SDN-POA volumes than their counterparts who were fed a phytoestrogen rich diet throughout life (99).

Another relevant outcome assessed in animals exposed to phytoestrogens is visual spatial memory (VSM). VSM refers to the ability to discern relationships between shapes and

objects, or to remember locations. This is an appropriate measure for examining effects on sexually dimorphic traits, since males consistently out-perform females on VSM tests. Earlier studies have demonstrated that exogenous estradiol will inhibit VSM in male rodents, but improve VSM in ovariectomized females. Similarly, in a study of life-long phytoestrogen diet exposure, exposed females performed VSM tasks significantly faster than unexposed females. The opposite effect was observed for males, with the unexposed performing better than the exposed rats (100, 101). Conversely, low dose of exogenous estrogen administered developmentally has also been shown to improve spatial learning in male rats (102).

Male rats exposed to low doses (but not high doses) of genistein from conception through lactation exhibited decreased offensive behaviors and increased defensive behaviors in adulthood, which is considered to be indicative of 'demasculinizing' effects (103). Adult exposure to phytoestrogen rich diets increased aggression in male monkeys and hamsters (104, 105). Post-weaning exposure in the hamster increased aggressive behavior, but not significantly over controls. In the female rat, neonatal genistein exposure affected estrus cycling. It also interfered with lordic posturing, an important adult reproductive behavior, which may be indicative of "defeminizing" effects (106). This is similar to effects seen in animals exposed to other EDC chemicals. For example, male rats have shown increased aggressive behavior during adulthood following prenatal exposure to the EDC, bisphenol A (107).

Most of these experiments have focused on exposures occurring in the prenatal, perinatal and neonatal stages of rodent development. However, the timing of these exposures may have limited relevance to the developing human. Typically, the neonatal rodent is underdeveloped in comparison to the neonatal human, and thus may be more representative

of the prenatal human in terms of developmental stage. With this understanding, these rodent models are regarded as informative with respect to general biologic processes, but limited with respect to the amount of extrapolation that can be carried out between species. Another important caveat with respect to interpreting these animal studies is that, in studies of "phytoestrogen" exposure, one should note that animals tend to produce equol from daidzein much more rapidly and consistently than humans. Therefore, it is possible that observed effects may be attributable to equol, and not other more relevant phytoestrogen components (108).

Human Studies on the Influence of Hormonal and EDC Exposures: Human studies of hormonal influence of language and behavior have largely focused on the effects of exposure to androgens, such as testosterone. As described previously, testosterone is influential in normal sexual differentiation of the male brain. Additional studies of exogenous estrogens, such as DES and polycholobiphenyls (PCBs), have also provided evidence that early life endocrine disruption can affect learning and behavior.

The behavioral effects of androgens have perhaps been most widely studied among subjects with congenital adrenal hyperplasia (CAH). CAH is a disorder characterized by 21-hydroxylase deficiency that results in elevated levels of androgens. Accordingly, females with CAH tend to show masculinized behaviors, including masculine toy play, increased physical aggression, greater spatial ability, and decreased interest in parenting. Masculine effects are associated most highly with the most severe cases of CAH. Severe cases are typically diagnosed and treated before 6 months of age, suggesting that the prenatal and early postnatal windows of exposure to androgens may be the most influential on behavior (109).

Additionally, "normal" female infants born to mothers with high prenatal testosterone levels were shown to exhibit masculine toy preference at age 3.5 years. No effect on toy preference was observed in boys born to women with high prenatal testosterone (110).

Postnatally, neonatal brain activity associated with phonological discrimination has been shown to vary by both biological sex, and by testosterone level. When exposed to a series of similar and dissimilar syllable sounds, the brain activity of 4-week old female infants demonstrated a clear response to the "mismatch" of sounds. A discrimination response was also observed in male infants with low testosterone (< 168 ng/dl), but not in male infants with high testosterone levels (>168 ng/dl). The response in low testosterone males was primarily lateralized to the left hemisphere, whereas the response was bilateral in females (median testosterone = 17 ng/dl) (79). This study provides preliminary evidence that brain organization and early life language processing may be driven, at least in part, by testosterone.

Prenatal exposure to diethylstilbestrol (DES), a potent estrogen-like compound, is associated with lowered spatial cognitive ability among males, evidence for a feminizing effect (111). Sexually dimorphic behavioral characteristics among adult women exposed to DES prenatally are varied. DES exposure among women has been associated with both decreased interest in parenting, which is a masculine trait, and decreased levels of "rough and tumble" play as girls, a more feminine trait (112). Other studies have shown no effect (113).

Lastly, in a study of the effects of environmental contaminiant exposure via maternal fish consumption, duration of breast feeding among highly exposed women was associated with more feminized (hyper-feminization) behavior in girls. In boys, masculine behavior increased with age and years of maternal fish consumption (114).

2. Critical Review of Soy Based Infant Formula Exposure and Cognitive and Behavioral Outcomes

Studies examining the effects of SBF on cognitive and behavioral development are limited in many respects. Only three studies have been identified that explore "cognitive function," and each assesses this outcome in a different way: brain activity, intelligence quotients, and academic achievement. Sexually dimorphic *behaviors* are not assessed in any of these studies, nor are subtle effects of SBF on gender-modified developmental characteristics. The results, limitations, and overall relevance of these studies are discussed below.

Jing et al. (115) recently investigated the potential differences in electroencephalographic (EEG) spectral power in multiple brain areas of 46 CMF and 39 SBF fed infants, at 3 and 6 months of age. When EEG signals were compared between the feeding groups, no differences were detected. However, spectral edge frequencies were significantly different between genders, ages, and brain areas. No significant interaction was observed between feeding group and gender, suggesting an overall null effect of soy feeding on EEG-measured brain activity. However, this study does have a limited sample size, which restricts its power to detect modest significant findings, particularly with respect to interactions. Additionally, this study does not capture more latent effects of soy exposure on brain development and function, highlighting a need for assessment of long-term, not immediate, effects of early life soy exposure on neurological outcomes.

Malloy and Berendes (116) explored the effects of breastfeeding on IQ in a population of 9 and 10 year old subjects exposed to either soy formula, or a combination of soy formula and breastfeeding. Six hundred fifty three school-aged subjects were recruited into a

retrospective study, intended to examine the effects of exposure to a chloride-deficient soy formula, as compared to other soy formulas, in infancy. Formula exposure was determined through parental recall questionnaires. IQs were measured using the Weschler Intelligence Scale (WISC) at ages 9 and 10. When no difference was detected between the chloride-deficient and other soy formulas, a follow up was conducted to assess the effects of partial breastfeeding versus no breastfeeding (exclusive soy feeding). One hundred eighty eight subjects were exclusively soy fed, while 466 were fed some combination of soy formula and breast milk. In crude analyses, the breastfed group had significantly higher IQ scores than the exclusively soy-fed group. However, following adjustment for parental education and income, the difference in IQ score was no longer distinguishable between the groups. More importantly, no difference was seen when the comparison was limited to exclusively soy-fed infants, and infants reportedly exclusively breast-fed for the first 60 days. No report on IQ differences with respect to gender is reported, nor is it supported that IQs among boys and girls at this age are distinct enough to use as measures of sexually-dimorphic development. This study is also limited by the lack of an adequate, non-soy fed reference group. Given that all subjects were fed soy at some point in infancy, results reported in this study are likely to be diluted by poor exposure contrast, and thus biased towards the null.

Finally, Strom et al. (117) suggested a null effect of SBF on cognitive function in a population of adults that had been exposed to either SBF or CMF in infancy as part of a clinical trial. They reported no difference in the level of education attained among the SBF and CMF groups. Education level was defined dichotomously as less than or equal to high school, and greater than high school. The null effect was reported for both males and females. This study is informative, in that it suggests no gross cognitive effects in SBF

exposed infants. However, this study does not allow for investigation of subtle cognitive or behavioral effects, such as those that should be observable in the proposed investigation.

3. Significance

SBF is an alternative source of infant nutrition for parents not otherwise satisfied with breast feeding or cow's milk infant formulas. It is widely used, and is generally considered to be a safe infant feeding method. Previous studies have suggested that it does not have gross effects on developmental outcomes, including growth, immune function, and to some extent, cognitive function. However, no study to date has examined the EDC-like properties of SBF, particularly with respect to outcomes related to "masculinzing" or "feminizing" behaviors or characteristics.

Experimental animal research provides compelling evidence that soy phytoestrogens, and genistein in particular, are capable of affecting sexually dimorphic developmental characteristics. Gender-modified responses to these compounds have been observed in the study of both physiologic and behavioral outcomes. These findings are supported by plausible mechanisms of action, involving isoflavone-estrogen receptor binding and alterations of normal estrogen-dependent activity. Human studies have consistently demonstrated that SBF-fed infants are exposed to remarkably high levels of isoflavones. Combined, these elements support that improved evaluation of SBF exposure effects in humans is warranted, with particular emphasis on gender-modified sexually dimorphic behaviors.

Comparison of within-gender language acquisition rates and gender-role behavior among SBF exposed and unexposed individuals will provide insight into the subtle effects that SBF may have on normal child development. Although subtle changes in these developmental

measures may not present a large public health burden, such changes may indicate that a biological response to isoflavones is occurring in the developing child exposed to SBF. With the understanding that early life is generally a time of low hormonal activity, such a response may be cause for concern. It is possible that isoflavone-induced changes in development may have greater health consequences in the long term. Such effects may include adverse effects on reproductive function, or the development of certain cancers. For example, perinatal estrogen exposure has been shown to affect prostate development in men (118), while neonatal genistein can induce uterine adenocarcinoma in rodents (119).

It is important to acknowledge that soy products, in general, have become somewhat ubiquitous in the human diet in the form of oils, meat/dairy substitutes, and other processed foods and drinks. The popularity of soy is driven, in part, by economic forces and ease of soybean production and growth, and also by common popular culture perceptions that soy-based products are "health foods." However, this perception is increasingly falling under question as more research attention is focused on the potential adverse effects of estrogen-like isoflavones. Exploring the true health effects of soy products in adult populations is beyond the scope of this study, but studying the subtle health effects in infants who are highly exposed to soy isoflavones through SBF is an ideal approach for assessing their biologic activity in a susceptible human population.

Finally, exposure to SBF and soy isoflavones is modifiable, and is thus appropriate for public health research and intervention. If developmental effects of SBF exposure are discovered, parents and pediatricians can be encouraged to seek out alternative formulas, such as hydrolyzed formulas. Soy formula manufacturers may also pursue opportunities to improve the safety of their products by decreasing or eliminating the isoflavone content.

ii. Time-to-Menarche

1. Background

Characterization and Trends of Puberty in Adolescent Females: Puberty is a transitional period in development that encompasses the change from a non-reproductive to a reproductive state. It is characterized by a series of biological changes including the development of secondary sex characteristics, accelerated growth, behavioral changes, and menarche (67). The physiological properties that regulate the onset of puberty are complex, and potentially have origins in the fetal and neonatal periods.

Developmental changes in puberty are largely regulated by the hypothalamic-pituitary-gonadal axis (HPG) and the hypothalamic-pituitary-adrenal axis (HPA). Activation of both axes requires central nervous system signaling to the hypothalamus. HPG activity has been described previously. Specifically in the female, LH and FSH target the ovaries, and initiate the release of both estrogens and androgens. LH and FSH signaling eventually lead to ova production and menarche. The HPG is of particular interest not only to early life learning and behavioral outcomes described above, but also to research questions involving effects of early life exposures on puberty. The HPG is active in mid-fetal, neonatal, infant, and prepubertal/pubertal stages of human development (67), so early life modifications can certainly influence outcomes later (120).

For the HPA, a similar cascade of signaling events from the hypothalamus, to the pituitary, to the adrenal cortex occurs. The adrenal cortex then produces hormones that interact with androgens to initiate growth of armpit and pubic hair, and skin changes. The hypothalamus and pituitary also can influence the production and release of hormones from the thyroid, including thyroid stimulating hormone (TSH). TSH can trigger the release of

other thyroid hormones (T_3 and T_4). Experimental evidence suggests that thyroid hormones can interact with gonadotropins in the HPG cascade described above, and consequently influence ovarian development. Hypothyroidism, for instance, has been shown to decrease FSH and LH concentrations, and block the event of first ovulation in rats (121).

The onset of puberty can be measured in a variety of ways. Commonly, breast budding, pubic hair development, and age at menarche are used. Age at menarche is ideal for use in studies that do not allow for physical examination. It can be reliably reported by a girl or her parent, provided that the interview occurs shortly after the menarche event. In a study of 88 adolescent girls, Koo et al. (122) determined that 59.1% of girls could correctly recall the exact month and year of menarche, and 77.3% could recall menarche correctly within one month, after a mean 430 day interval between initial reporting of menarche and recall reporting (range = 266-698 days). Exact recall reliability was higher (66%) for girls interviewed after a shorter interval (mean = 323 days) than it was for girls interviewed after a longer interval (mean = 649 days) (44.8%). Other studies have assessed age at menarche recall in adult populations, and have also shown that reliability of recall tends to decrease over time (123), or that adult recall is only accurate with respect to year of age (124, 125). Therefore, this measure is best and most precisely used in adolescent, not adult, populations.

Populations have historically recorded age at menarche data for many years. Data from the late 1800s through the mid-1900s indicate that the average age at menarche has been declining over time. This decline is largely attributed to improved health and nutrition. More controversial is the proposed decline in pubertal onset that has been seen in more recent years. Reports from the 1950s and 1960s suggest the average age of menarche was

approximately 12.8-12.9 years. However, more recent data report an earlier age at menarche in the US population, at 12.4-12.5 years (126).

Age at pubertal onset, and age at menarche specifically, differ by race. National Health and Nutrition Examination Survey (NHANES) III data indicate that the average age for initiation of breast development was approximately 10.3 years for Caucasian girls, and 9.5 years for African American girls in a 1988-1994 US population of 8 – 16 year olds. Average age at menarche was approximately 12.6 for Caucasians, and approximately 12.1 for African Americans (127).

Other factors that appear to affect the age at onset of puberty include body fat composition or body mass index (BMI), birth weight and genetics. High childhood BMI and low birth weight have been associated with advanced age at menarche, while poor nutrition may result in a delay of pubertal onset. Poor nutrition may be the result of socioeconomic environment, an eating disorder, or the result of excessive exercise. Maternal age at menarche is also predictive of a daughter's age at puberty.

Influence of Endocrine Disruption on Onset of Puberty: Timing of puberty and normalcy of pubertal development are measured in animal studies in a variety of ways. Perhaps the most commonly reported outcome is age at vaginal opening (VO) in rats. This event is thought to correspond to breast development in humans. It typically occurs around postnatal day 35, approximately 4.2 days prior to the animals' first ovulation and estrus (analogous to menarche in humans) (128). Thyroid hormones, uterine and ovary weights, and vaginal cytology are also frequently reported outcomes when studying the effects of environmental compounds on pubertal development in rats (121).

Accordingly, genistein and other isoflavones have been shown to affect VO in multiple studies. Advancement of VO was observed in association with isoflavone exposure following early neonatal (birth – PND 3) exposure (83), persistent (birth – PND 21) and post-weaning (PND21-27) postnatal exposure (93, 129), and life-long exposure (gestation – puberty)(130). A similar advancement effect was seen when soy protein isolate (SPI) was administered from gestation through adulthood (131).

Another study showed a linear trend toward accelerated onset of VO across doses ranging from 4 to 1250 ppm diet, along with abdominal cellular maturation of vaginal cells at the highest doses (gestation – PND 50) (132). Delayed VO was observed on one study of prenatal genistein exposure (96), and one study assessing lifelong exposure to a phytoestrogen-rich diet (99). No effect on VO was observed in another study, but differences in estrus cycling were observed after exposure during PND 1-5, instead (133). Masutomi et al. (97), however, found no effect on endocrine dependent outcomes, apart from a decrease in body weight at week 11 following exposure from gestational day 15 to PND 10.

The mechanism(s) by which EDC compounds may induce alterations in pubertal development are unclear, but again, are likely to be partially regulated by the hypothalamic-pituitary-gonadal (HPG) axis,. As described previously, the HPG axis regulates GnRH secretion and subsequent hormone release that ultimately leads to menarche. The functionality of the HPG axis is regulated by endogenous estradiol during development, and so it thought to be particularly susceptible to endocrine disruption. As evidence, exogenous estradiol exposure in rats has consistently induced effects on HPG dependent outcomes, including advanced age at VO and advanced age at first estrus (128). These effects were seen following exposures that started as early as postnatal day 1, and as late as postnatal day 10

and 20. Since the "critical" window of effect on HPG development in the rat is thought to be around postnatal day 10 or 12, these later exposures again provide evidence that there is a relatively wide time frame for susceptibility.

Human Studies, Effects of Endocrine Disruptors on Pubertal Onset: The effects of EDCs on pubertal development have been studied in several human populations. Both the advancement and delay of pubertal onset have been seen following perinatal exposure to certain pesticides and other organic compounds (67, 120). Interestingly, a study of brominated flame retardant chemicals in maternal serum demonstrated that girls who were exposed in utero and breastfed reported earlier menarche than girls exposed in utero and not breastfed, and earlier pubic hair development compared to both unexposed girls and girls who were not breastfed (134). While the role of brominated flame retardants in this scenario is not precisely clear, this study is informative with respect to the role of the postnatal human environment influencing pubertal onset later in life.

2. Critical Review of Soy Based Infant Formula and Pubertal Onset

Only one longitudinal study has assessed the effects of SBF on outcomes related to adolescence or adulthood. Several other studies have, however, assessed estrogen dependent outcomes in younger children exposed to SBF in infancy, such as early breast bud development and retention, hormone levels, and vaginal cytology.

Strom et al.(117) investigated approximately 30 outcomes related to overall health, onset of puberty and reproductive function. The study population consisted of individuals who had participated in SBF and CMF clinical trials as infants who were contacted again as adults, aged 20 to 34 years (n = 248 SBF, 563 CMF). All were healthy term infants, with birth weights greater than 2500 g. Outcomes among females included height, weight, education

attainment (discussed previously), age at menarche, age a bra was first worn, menstrual cycle length, number of days of menstrual bleeding, and many others. Data was collected through a telephone questionnaire. Among women, those fed SBF were slightly younger than those fed CMF, they were more likely to use allergy or asthma medication, and reported marginally less time spent engaged in "sedentary activities." No difference was seen in outcomes related to pubertal onset, with a mean age at menarche of 12.6 for SBF and 12.7 for CMF fed infants. Additionally, no difference was seen in height or weight. SBF fed women did report significantly more days of menstrual bleeding per cycle, and greater discomfort with menstruation. All estimates were adjusted for birth weight, current age, "usual" BMI, parent's stature, hormone disorders, alcohol and cigarette use, current soy food consumption, current physical activity, and others.

The soy exposure in this study is well classified due to clinical trial participation. However, since subjects consumed soy produced in the 1960s and 1970s, it is possible that the formulations administered to them differ slightly from what is available today. Authors do note that few subjects (n = 12) used a formula containing soy flour. Findings suggest that no severe long health complications result from early life soy exposure. However, the ability for women to accurately and precisely recall past events, such as age at menarche or breast development, is likely limited, and it is not clear whether validation of report was attempted (or possible) from medical records. There is also limited generalizablity associated with this population. Subjects were from the Midwestern US, were of high socioeconomic status, and were a predominantly white population.

Two studies have reported positive associations between SBF and breast development at or before the age of 2. In a case control study of girls with early thelarche in Puerto Rico, no

association was found among 120 matched pairs and any environmental or familial factors (135), when the population was limited to children under 2 years of age, a significant association with SBF consumption in infancy was found. Maternal ovarian cysts and consumption of various meat products were also associated with early thelarche. However, despite statistical findings, over 50% of the study population did not report having these risk factors. The findings of this study suggest that soy exposure may be part of a larger, multi-factorial causal mechanism leading to early thelarche, but is not the sole attributable exposure.

A cross sectional study of 694 Israeli girls, age 3 to 24 months, reported that those fed SBF for 3 months or more had a higher prevalence of breast tissue at age 2 compared to a combined CMF/breast milk comparison group (136). There was no difference in breast bud prevalence in the first year of life across the feeding groups. There was no statistical difference in breast bud development across a range of soy feeding regimes (exclusive soy, soy with breast and cow milk, soy with cow milk, soy with breast milk). Additionally, whereas CMF and breast fed girls experienced a decline in breast bud prevalence between 1 and 2 years of age, the SBF fed girls were more likely to retain breast tissue into the second year. Since this is not a population based study, it has limited generalizability. There is also potential differential exposure misclassification due to misreporting of feeding patterns by parents, as well as potential outcome misclassification of breast bud measurement due to difficulties in standardizing breast measurement.

Giampietro et al. (137) assessed a range physical, hormonal and metabolic factors in a study of 48 SBF fed children (age range 7 – 96 months) and 18 controls. 27 boys and 21 girls comprised the SBF fed group, and all had been exclusively fed SBF for at least 6

months due to either a family history of allergy or a documented CMF allergy. The 20 SBF subjects under the age of 24 months were still consuming soy regularly. Outcomes of interest included bone age, urinary markers of bone metabolism, serum levels of bone alkaline phosphatase, oestocalcin, 17B-estradiol, intact parathyroid hormone, "precocious puberty" in females and gynocomastia in males. Testosterone concentrations were not measured. No significant difference was reported for any outcome, with the exception of lower levels of calciuria and higher levels of phosphaturia in SBF fed children. Estradiol was below the level of detection in all subjects. This study uses a range of biomarkers and physical assessments to assert that SBF induces no long term hormonal effect in children. However, given the small sample size and wide range of ages and developmental stages captured in the study population, these results should be interpreted with caution. The study was grossly underpowered to study the outcomes that are most relevant to timing of puberty, such as estradiol levels, since the study included only 8 subjects at a relevant age (73 to 96 months). Lastly, this study was not designed to look at interactions of SBF and gender, so sex-specific effects, apart from precocious puberty in females and gynocomastia in males, were not assessed.

Finally, in a recent pilot study of 37 male and 35 female infants, at various age ranges between birth and 6 months, the effects of SBF, CMF and breastfeeding were assessed for a range of developmental measures (138). These included breast adipose tissue, breast buds, testicular volume, genital development, vaginal cell cytology, and information on vaginal discharge. The objective of this study was to assess feasibility of sample collection and study protocol, and thus does not provide reliable information on associations between exposure

and outcome. Interestingly, though, the study found that, in girls older than 30 days, vaginal wall cells tended to show higher maturity in SBF fed girls than in other girls.

Apart from the Strom et al., study, the epidemiologic literature offers little information on the effects of SBF on reproductive maturity. However, several studies do suggest that SBF exposure has endocrine related activity in females, as evidenced in studies of early thelarche and vaginal cytology. While little can be inferred from these studies regarding the association between SBF exposure and onset of menarche, they do highlight the need for additional studies of soy exposure and female reproductive health.

3. Significance

SBF is a common alternative to other infant feeding methods, such as cow's milk formula and breast feeding. For several decades, it has been considered a generally safe nutrition source in term infant populations. However, increased attention is currently focusing on how soy isoflavones found in SBF, and other soy products, may be associated with subtle effects on health and development.

Much is not known about the effects of early life exposure to soy isoflavones that can be found in SBF and other soy products. Since human endocrine and reproductive development is active from the prenatal period through childhood, exposure to biologically active soy isoflavones may cause a disruption or alteration in normal development. It is therefore the obligation of the public health community to thoroughly assess these possible effects in developing females.

The timing of the onset of puberty is widely posited as a sensitive endpoint for perinatal exposure to endocrine disruptors. Although it is a somewhat subtle outcome, it does have large public health significance. Early puberty can put children at higher risk for obesity,

psychosocial abuse, behavior disorders, and potential sexual abuse. Furthermore, early menarche has been associated with increased risk for breast cancer (139, 140).

The study of SBF as it relates to age at menarche will contribute to a growing literature on endocrine disruptors and age at menarche. In contrast to exposures to other environmental EDCs, which often occur passively, ubiquitously, and at low levels, isoflavone exposure is isolated to those knowingly given SBF as part of the diet. Infants with predominantly SBF-based diets have documented high exposures to soy isoflavones, and infants fed other diets have comparatively little exposure. This wide exposure contrast makes SBF an ideal EDC exposure to study.

Additionally, unlike other environmentally EDCs, exposure to soy products in infancy is highly modifiable. If this study identifies SBF as a potential risk for advanced age at menarche, infant formula product manufacturers and parents have the opportunity to adjust their behaviors accordingly.

REFERENCES

1. Bhatia J, Greer F, American Academy of Pediatrics Committee on Nutrition. Use of soy protein-based formulas in infant feeding. Pediatrics. 2008 May;121(5):1062-8.

2. Essex C. Phytoestrogens and soy based infant formula. BMJ. 1996 Aug 31;313(7056):507-8.

3. Hughes I, Woods HF, Committee on Toxicity Working Group on Phytoestrogens. Phytoestrogens and health. London: The Food Standards Agency; 2003.

4. Badger TM, Ronis MJ, Hakkak R, Rowlands JC, Korourian S. The health consequences of early soy consumption. J Nutr. 2002 Mar;132(3):559S-65S.

5. American Academy of Pediatrics. Committee on Nutrition. American academy of pediatrics. committee on nutrition. soy protein-based formulas: Recommendations for use in infant feeding. Pediatrics. 1998 Jan;101(1 Pt 1):148-53.

6. Zeiger RS, Sampson HA, Bock SA, Burks AW,Jr, Harden K, Noone S, et al. Soy allergy in infants and children with IgE-associated cow's milk allergy. J Pediatr. 1999 May;134(5):614-22.

7. Klemola T, Vanto T, Juntunen-Backman K, Kalimo K, Korpela R, Varjonen E. Allergy to soy formula and to extensively hydrolyzed whey formula in infants with cow's milk allergy: A prospective, randomized study with a follow-up to the age of 2 years. J Pediatr. 2002 Feb;140(2):219-24.

8. Nevo N, Rubin L, Tamir A, Levine A, Shaoul R. Infant feeding patterns in the first 6 months: An assessment in full-term infants. J Pediatr Gastroenterol Nutr. 2007 Aug;45(2):234-9.

9. Berger-Achituv S, Shohat T, Romano-Zelekha O, Ophir E, Rachmani S, Malovizky D, et al. Widespread use of soy-based formula without clinical indications. J Pediatr Gastroenterol Nutr. 2005 Nov;41(5):660-6.

10. Forsyth BW, Leventhal JM, McCarthy PL. Mothers' perceptions of problems of feeding and crying behaviors. A prospective study. Am J Dis Child. 1985 Mar;139(3):269-72.

11. Thygarajan A, Burks AW. American academy of pediatrics recommendations on the effects of early nutritional interventions on the development of atopic disease. Curr Opin Pediatr. 2008 Dec;20(6):698-702.

12. Chen A, Rogan WJ. Isoflavones in soy infant formula: A review of evidence for endocrine and other activity in infants. Annu Rev Nutr. 2004;24:33-54.

13. Setchell KD, Zimmer-Nechemias L, Cai J, Heubi JE. Isoflavone content of infant formulas and the metabolic fate of these phytoestrogens in early life. Am J Clin Nutr. 1998 Dec;68(6 Suppl):1453S-61S.

14. Ministry of Agriculture, Forestries and Food. Plant oestrogens in soya-based infant formulae. UK: MAFF; 1998.

15. Knight DC, Eden JA, Huang JL, Waring MA. Isoflavone content of infant foods and formulas. J Paediatr Child Health. 1998 Apr;34(2):135-8.

16. Franke AA, Halm BM, Custer LJ, Tatsumura Y, Hebshi S. Isoflavones in breastfed infants after mothers consume soy. Am J Clin Nutr. 2006;84:406-13.

17. Antignac J, Carou R, Le Bizec B, Cravedi J, Andre F. Identification of phytoestrogens in bovine milk using liquid chromatography/electrospray tandem mass spectrometry. Rapid Commun Mass Spectrom. 2003;17:1256-64.

18. Hwang CS, Kwak HS, Lim HJ, Lee SH, Kang YS, Choe TB, et al. Isoflavone metabolites and their in vitro dual functions: They can act as an estrogenic agonist or antagonist depending on the estrogen concentration. J Steroid Biochem Mol Biol. 2006 Nov;101(4-5):246-53.

19. ESPGHAN Committee on Nutrition, Agostoni C, Axelsson I, Goulet O, Koletzko B, Michaelsen KF, et al. Soy protein infant formulae and follow-on formulae: A commentary by the ESPGHAN committee on nutrition. J Pediatr Gastroenterol Nutr. 2006 Apr;42(4):352-61.

20. Setchell KD. Phytoestrogens: The biochemistry, physiology, and implications for human health of soy isoflavones. Am J Clin Nutr. 1998 Dec;68(6 Suppl):1333S-46S.

21. Setchell KD, Zimmer-Nechemias L, Cai J, Heubi JE. Exposure of infants to phyto-oestrogens from soy-based infant formula. Lancet. 1997 Jul 5;350(9070):23-7.

22. Cao Y, Calafat AM, Doerge DR, Umbach DM, Bernbaum JC, Twaddle NC, et al. Isoflavones in urine, saliva, and blood of infants: Data from a pilot study on the estrogenic activity of soy formula. J Expo Sci Environ Epidemiol. 2009 Feb;19(2):223-34.

23. Bannwart C, Adlercreutz H, Wahala K, Kotiaho T, Hesso A, Brunow G, et al. Identification of the phyto-oestrogen 3',7-dihydroxyisoflavan, an isomer of equol, in human urine and cow's milk. Biomed Environ Mass Spectrom. 1988 Jul;17(1):1-6.

24. Hadley ME, Levine JE. Endocrinology. 6th ed. San Francisco: Benjamin Cummings; 2006.

25. Couse JF, Lindzey J, Grandien K, Gustafsson JA, Korach KS. Tissue distribution and quantitative analysis of estrogen receptor-alpha (ERalpha) and estrogen receptor-beta

(ERbeta) messenger ribonucleic acid in the wild-type and ERalpha-knockout mouse. Endocrinology. 1997 Nov;138(11):4613-21.

26. Shughrue PJ, Lane MV, Merchenthaler I. Comparative distribution of estrogen receptor-alpha and -beta mRNA in the rat central nervous system. J Comp Neurol. 1997 Dec 1;388(4):507-25.

27. Santell RC, Chang YC, Nair MG, Helferich WG. Dietary genistein exerts estrogenic effects upon the uterus, mammary gland and the hypothalamic/pituitary axis in rats. J Nutr. 1997 Feb;127(2):263-9.

28. Patisaul HB. Phytoestrogen action in the adult and developing brain. J Neuroendocrinol. 2005 Jan;17(1):57-64.

29. Nagel SC, vom Saal FS, Welshons WV. The effective free fraction of estradiol and xenoestrogens in human serum measured by whole cell uptake assays: Physiology of delivery modifies estrogenic activity. Proc Soc Exp Biol Med. 1998 Mar;217(3):300-9.

30. Milligan SR, Khan O, Nash M. Competitive binding of xenobiotic oestrogens to rat alpha-fetoprotein and to sex steroid binding proteins in human and rainbow trout (oncorhynchus mykiss) plasma. Gen Comp Endocrinol. 1998 Oct;112(1):89-95.

31. Dixon RA, Ferreira D. Genistein. Phytochemistry. 2002 Jun;60(3):205-11.

32. Takahashi Y, Odbayar TO, Ide T. A comparative analysis of genistein and daidzein in affecting lipid metabolism in rat liver. J Clin Biochem Nutr. 2009 May;44(3):223-30.

33. Akiyama T, Ishida J, Nakagawa S, Ogawara H, Watanabe S, Itoh N, et al. Genistein, a specific inhibitor or tyrosine-specific protein kinases. J Biol Chem. 1987;262(12):5592-5.

34. Dang ZC. Dose-dependent effects of soy phyto-oestrogen genistein on adipocytes: Mechanisms of action. Obes Rev. 2009 Jan 15.

35. Dang ZC, Audinot V, Papapoulos SE, Boutin JA, Lowik CW. Peroxisome proliferator-activated receptor gamma (PPARgamma) as a molecular target for the soy phytoestrogen genistein. J Biol Chem. 2003 Jan 10;278(2):962-7.

36. Wagner JD, Zhang L, Shadoan MK, Kavanagh K, Chen H, Tresnasari K, et al. Effects of soy protein and isoflavones on insulin resistance and adiponectin in male monkeys. Metabolism. 2008 Jul;57(7 Suppl 1):S24-31.

37. Doerge DR, Sheehan DM. Goitrogenic and estrogenic activity of soy isoflavones. Environ Health Perspect. 2002 Jun;110 Suppl 3:349-53.

38. Hofmann PJ, Schomburg L, Kohrle J. Interference of endocrine disrupters with thyroid hormone receptor-dependent transactivation. Toxicol Sci. 2009 Jul;110(1):125-37.

39. Oetting A, Yen PM. New insights into thyroid hormone action. Best Pract Res Clin Endocrinol Metab. 2007 Jun;21(2):193-208.

40. Doerge DR, Twaddle NC, Banks EP, Jefferson WN, Newbold RR. Pharmacokinetic analysis in serum of genistein administered subcutaneously to neonatal mice. Cancer Lett. 2002 Oct 8;184(1):21-7.

41. Huggett AC, Pridmore S, Malnoe A, Haschke F, Offord EA. Phyto-oestrogens in soy-based infant formula. Lancet. 1997 Sep 13;350(9080):815-6.

42. Rozman KK, Bhatia J, Calafat AM, Chambers C, Culty M, Etzel RA, et al. NTP-CERHR expert panel report on the reproductive and developmental toxicity of soy formula. Birth Defects Res B Dev Reprod Toxicol. 2006 Aug;77(4):280-397.

43. Mendez MA, Anthony MS, Arab L. Soy-based formulae and infant growth and development: A review. J Nutr. 2002 Aug;132(8):2127-30.

44. Merritt RJ, Jenks BH. Safety of soy-based infant formulas containing isoflavones: The clinical evidence. J Nutr. 2004 May;134(5):1220S-4S.

45. Jabbar MA, Larrea J, Shaw RA. Abnormal thyroid function tests in infants with congenital hypothyroidism: The influence of soy-based formula. J Am Coll Nutr. 1997 Jun;16(3):280-2.

46. Chorazy PA, Himelhoch S, Hopwood NJ, Greger NG, Postellon DC. Persistent hypothyroidism in an infant receiving a soy formula: Case report and review of the literature. Pediatrics. 1995 Jul;96(1 Pt 1):148-50.

47. Conrad SC, Chiu H, Silverman BL. Soy formula complicates management of congenital hypothyroidism. Arch Dis Child. 2004 Jan;89(1):37-40.

48. Cruz ML, Wong WW, Mimouni F, Hachey DL, Setchell KD, Klein PD, et al. Effects of infant nutrition on cholesterol synthesis rates. Pediatr Res. 1994 Feb;35(2):135-40.

49. Strotmeyer ES, Yang Z, LaPorte RE, Chang YF, Steenkiste AR, Pietropaolo M, et al. Infant diet and type 1 diabetes in china. Diabetes Res Clin Pract. 2004 Sep;65(3):283-92.

50. Fort P, Lanes R, Dahlem S, Recker B, Weyman-Daum M, Pugliese M, et al. Breast feeding and insulin-dependent diabetes mellitus in children. J Am Coll Nutr. 1986;5(5):439-41.

51. Cassidy A, Bingham S, Setchell KD. Biological effects of a diet of soy protein rich in isoflavones on the menstrual cycle of premenopausal women. Am J Clin Nutr. 1994 Sep;60(3):333-40.

52. Hooper L, Ryder JJ, Kurzer MS, Lampe JW, Messina MJ, Phipps WR, et al. Effects of soy protein and isoflavones on circulating hormone concentrations in pre- and post-menopausal women: A systematic review and meta-analysis. Hum Reprod Update. 2009 Jul-Aug;15(4):423-40.

53. Nagata C, Inaba S, Kawakami N, Kakizoe T, Shimizu H. Inverse association of soy product intake with serum androgen and estrogen concentrations in japanese men. Nutr Cancer. 2000;36(1):14-8.

54. Habito RC, Montalto J, Leslie E, Ball MJ. Effects of replacing meat with soyabean in the diet on sex hormone concentrations in healthy adult males. Br J Nutr. 2000 Oct;84(4):557-63.

55. Nagata C, Takatsuka N, Shimizu H, Hayashi H, Akamatsu T, Murase K. Effect of soymilk consumption on serum estrogen and androgen concentrations in japanese men. Cancer Epidemiol Biomarkers Prev. 2001 Mar;10(3):179-84.

56. Mitchell JH, Cawood E, Kinniburgh D, Provan A, Collins AR, Irvine DS. Effect of a phytoestrogen food supplement on reproductive health in normal males. Clin Sci (Lond). 2001 Jun;100(6):613-8.

57. Cooke PS, Selvaraj V, Yellayi S. Genistein, estrogen receptors, and the acquired immune response. J Nutr. 2006 Mar;136(3):704-8.

58. Fenson L, Dale PS, Reznick JS, Bates E, Thal DJ, Pethick SJ. Variability in early communicative development. Monographs of the Society for Research in Child Development. 1994;59(5, Serial No. 242):i-185.

59. Bauer DJ, Goldfield ba, Reznick JS. Alternative approaches to analyzing individual differences in the rate of early vocabulary development. 2002;23:313-35.

60. Westerlund M, Lagerberg D. Expressive vocabulary in 18-month-old children in relation to demographic factors, mother and child characteristics, communication style and shared reading. Child Care Health Dev. 2008 Mar;34(2):257-66.

61. Stokes SF, Klee T. Factors that influence vocabulary development in two-year-old children. J Child Psychol Psychiatry. 2009 Apr;50(4):498-505.

62. Berglund E, Eriksson M, Westerlund M. Communicative skills in relation to gender, birth order, childcare and socioeconomic status in 18-month-old children. Scand J Psychol. 2005 Dec;46(6):485-91.

63. Bornstein MH, Haynes MO, Painter KM. Sources of child vocabulary competence: A multivariate model. J Child Lang. 1998 Jun;25(2):367-93.

64. Reilly S, Wake M, Bavin EL, Prior M, Williams J, Bretherton L, et al. Predicting language at 2 years of age: A prospective community study. Pediatrics. 2007;120(6):e1441-9.

65. Hines M. Sex steroids and human behavior: Prenatal androgen exposure and sex-typical play behavior in children. Ann N Y Acad Sci. 2003 Dec;1007:272-82.

66. Golombok S, Rust J, Zervoulis K, Croudace T, Golding J, Hines M. Developmental trajectories of sex-typed behavior in boys and girls: A longitudinal general population study of children aged 2.5-8 years. Child Dev. 2008 Sep-Oct;79(5):1583-93.

67. Buck Louis GM, Gray LE,Jr, Marcus M, Ojeda SR, Pescovitz OH, Witchel SF, et al. Environmental factors and puberty timing: Expert panel research needs. Pediatrics. 2008 Feb;121 Suppl 3:S192-207.

68. Meethal SV, Atwood CS. The role of hypothalamic-pituitary-gonadal hormones in the normal structure and functioning of the brain. Cell Mol Life Sci. 2005;62:257-70.

69. Scallet AC, Divine RL, Newbold RR, Delclos KB. Increased volume of the calbindin D28k-labeled sexually dimorphic hypothalamus in genistein and nonylphenol-treated male rats. Toxicol Sci. 2004 Dec;82(2):570-6.

70. Udry JR. Putting prenatal effects on sex-dimorphic behavior in perspective: An absolutely complete theory. Epidemiology. 2003 Mar;14(2):135-6.

71. Cohen-Bendahan CC, van de Beek C, Berenbaum SA. Prenatal sex hormone effects on child and adult sex-typed behavior: Methods and findings. Neurosci Biobehav Rev. 2005 Apr;29(2):353-84.

72. Schmidt H, Schwarz HP. Serum concentrations of LH and FSH in the healthy newborn. Eur J Endocrinol. 2000 Aug;143(2):213-5.

73. Mann DR, Gould KG, Collins DC, Wallen K. Blockade of neonatal activation of the pituitary-testicular axis: Effect on peripubertal luteinizing hormone and testosterone secretion and on testicular development in male monkeys. J Clin Endocrinol Metab. 1989 Mar;68(3):600-7.

74. Boas M, Boisen KA, Virtanen HE, Kaleva M, Suomi AM, Schmidt IM, et al. Postnatal penile length and growth rate correlate to serum testosterone levels: A longitudinal study of 1962 normal boys. Eur J Endocrinol. 2006 Jan;154(1):125-9.

75. Brown GR, Dixson AF. Investigation of the role of postnatal testosterone in the expression of sex differences in behavior in infant rhesus macaques (macaca mulatta). Horm Behav. 1999 Apr;35(2):186-94.

76. Nevison CM, Brown GR, Dixson AF. Effects of altering testosterone in early infancy on social behaviour in captive yearling rhesus monkeys. Physiol Behav. 1997 Dec;62(6):1397-403.

77. Mann DR, Akinbami MA, Gould KG, Paul K, Wallen K. Sexual maturation in male rhesus monkeys: Importance of neonatal testosterone exposure and social rank. J Endocrinol. 1998 Mar;156(3):493-501.

78. Wallen K, Maestripieri D, Mann DR. Effects of neonatal testicular suppression with a GnRH antagonist on social behavior in group-living juvenile rhesus monkeys. Horm Behav. 1995 Sep;29(3):322-37.

79. Friederici AD, Pannekamp A, Partsch CJ, Ulmen U, Oehler K, Schmutzler R, et al. Sex hormone testosterone affects language organization in the infant brain. Neuroreport. 2008 Feb 12;19(3):283-6.

80. Chellakooty M, Schmidt IM, Haavisto AM, Boisen KA, Damgaard IN, Mau C, et al. Inhibin A, inhibin B, follicle-stimulating hormone, luteinizing hormone, estradiol, and sex hormone-binding globulin levels in 473 healthy infant girls. J Clin Endocrinol Metab. 2003 Aug;88(8):3515-20.

81. Sehested A, Juul AA, Andersson AM, Petersen JH, Jensen TK, Muller J, et al. Serum inhibin A and inhibin B in healthy prepubertal, pubertal, and adolescent girls and adult women: Relation to age, stage of puberty, menstrual cycle, follicle-stimulating hormone, luteinizing hormone, and estradiol levels. J Clin Endocrinol Metab. 2000 Apr;85(4):1634-40.

82. Robinson J. Prenatal programming of the female reproductive neuroendocrine system by androgens. Reproduction. 2006 Oct;132(4):539-47.

83. Bateman HL, Patisaul HB. Disrupted female reproductive physiology following neonatal exposure to phytoestrogens or estrogen specific ligands is associated with decreased GnRH activation and kisspeptin fiber density in the hypothalamus. Neurotoxicology. 2008 Nov;29(6):988-97.

84. Cooke B, Hegstrom CD, Villeneuve LS, Breedlove SM. Sexual differentiation of the vertebrate brain: Principles and mechanisms. Front Neuroendocrinol. 1998 Oct;19(4):323-62.

85. Perez SE, Chen EY, Mufson EJ. Distribution of estrogen receptor alpha and beta immunoreactive profiles in the postnatal rat brain. Brain Res Dev Brain Res. 2003 Oct 10;145(1):117-39.

86. Sharpe RM, Martin B, Morris K, Greig I, McKinnell C, McNeilly AS, et al. Infant feeding with soy formula milk: Effects on the testis and on blood testosterone levels in marmoset monkeys during the period of neonatal testicular activity. Hum Reprod. 2002 Jul;17(7):1692-703.

87. Tan KA, Walker M, Morris K, Greig I, Mason JI, Sharpe RM. Infant feeding with soy formula milk: Effects on puberty progression, reproductive function and testicular cell numbers in marmoset monkeys in adulthood. Hum Reprod. 2006 Apr;21(4):896-904.

88. Patisaul HB, Fortino AE, Polston EK. Differential disruption of nuclear volume and neuronal phenotype in the preoptic area by neonatal exposure to genistein and bisphenol-A. Neurotoxicology. 2007 Jan;28(1):1-12.

89. Patisaul HB, Polston EK. Influence of endocrine active compounds on the developing rodent brain. Brain Res Rev. 2008 Mar;57(2):352-62.

90. Gorski RA. Sexual dimorphisms of the brain. J Anim Sci. 1985;61 Suppl 3:38-61.

91. Gorski RA, Gordon JH, Shryne JE, Southam AM. Evidence for a morphological sex difference within the medial preoptic area of the rat brain. Brain Res. 1978 Jun 16;148(2):333-46.

92. Davis EC, Popper P, Gorski RA. The role of apoptosis in sexual differentiation of the rat sexually dimorphic nucleus of the preoptic area. Brain Res. 1996;734(1-2):10-8.

93. Lewis RW, Brooks N, Milburn GM, Soames A, Stone S, Hall M, et al. The effects of the phytoestrogen genistein on the postnatal development of the rat. Toxicol Sci. 2003 Jan;71(1):74-83.

94. Faber KA, Hughes CL,Jr. The effect of neonatal exposure to diethylstilbestrol, genistein, and zearalenone on pituitary responsiveness and sexually dimorphic nucleus volume in the castrated adult rat. Biol Reprod. 1991 Oct;45(4):649-53.

95. Faber KA, Hughes CL,Jr. Dose-response characteristics of neonatal exposure to genistein on pituitary responsiveness to gonadotropin releasing hormone and volume of the sexually dimorphic nucleus of the preoptic area (SDN-POA) in postpubertal castrated female rats. Reprod Toxicol. 1993;7(1):35-9.

96. Levy JR, Faber KA, Ayyash L, Hughes CL,Jr. The effect of prenatal exposure to the phytoestrogen genistein on sexual differentiation in rats. Proc Soc Exp Biol Med. 1995 Jan;208(1):60-6.

97. Masutomi N, Shibutani M, Takagi H, Uneyama C, Takahashi N, Hirose M. Impact of dietary exposure to methoxychlor, genistein, or diisononyl phthalate during the perinatal period on the development of the rat endocrine/reproductive systems in later life. Toxicology. 2003 Nov 5;192(2-3):149-70.

98. Slikker W,Jr, Scallet AC, Doerge DR, Ferguson SA. Gender-based differences in rats after chronic dietary exposure to genistein. Int J Toxicol. 2001 May-Jun;20(3):175-9.

99. Lund TD, Rhees RW, Setchell KD, Lephart ED. Altered sexually dimorphic nucleus of the preoptic area (SDN-POA) volume in adult long-evans rats by dietary soy phytoestrogens. Brain Res. 2001 Sep 28;914(1-2):92-9.

100. Lephart ED, Setchell KD, Handa RJ, Lund TD. Behavioral effects of endocrine-disrupting substances: Phytoestrogens. ILAR J. 2004;45(4):443-54.

101. Lephart ED, West TW, Weber KS, Rhees RW, Setchell KD, Adlercreutz H, et al. Neurobehavioral effects of dietary soy phytoestrogens. Neurotoxicol Teratol. 2002 Jan-Feb;24(1):5-16.

102. Corrieri L, Della Seta D, Canoine V, Fusani L. Developmental exposure to xenoestrogen enhances spatial learning in male rats. Horm Behav. 2007 May;51(5):620-5.

103. Wisniewski AB, Klein SL, Lakshmanan Y, Gearhart JP. Exposure to genistein during gestation and lactation demasculinizes the reproductive system in rats. J Urol. 2003 Apr;169(4):1582-6.

104. Simon NG, Kaplan JR, Hu S, Register TC, Adams MR. Increased aggressive behavior and decreased affiliative behavior in adult male monkeys after long-term consumption of diets rich in soy protein and isoflavones. Horm Behav. 2004 Apr;45(4):278-84.

105. Moore TO, Karom M, O'Farrell L. The neurobehavioral effects of phytoestrogens in male syrian hamsters. Brain Res. 2004 Jul 30;1016(1):102-10.

106. Kouki T, Kishitake M, Okamoto M, Oosuka I, Takebe M, Yamanouchi K. Effects of neonatal treatment with phytoestrogens, genistein and daidzein, on sex difference in female rat brain function: Estrous cycle and lordosis. Horm Behav. 2003 Aug;44(2):140-5.

107. Farabollini F, Porrini S, Della Seta D, Bianchi F, Dessì-Fulgheri F. Effects of perinatal exposure to bisphenol A on sociosexual behavior of female and male rats. Environ Health Perspect. 2002;110(Suppl 3):409-14.

108. Lephart ED, Porter JP, Lund TD, Bu L, Setchell KD, Ramoz G, et al. Dietary isoflavones alter regulatory behaviors, metabolic hormones and neuroendocrine function in long-evans male rats. Nutr Metab (Lond). 2004 Dec 23;1(1):16.

109. Berenbaum SA, Duck SC, Bryk K. Behavioral effects of prenatal versus postnatal androgen excess in children with 21-hydroxylase-deficient congenital adrenal hyperplasia. J Clin Endocrinol Metab. 2000 Feb;85(2):727-33.

110. Hines M, Golombok S, Rust J, Johnston KJ, Golding J, Avon Longitudinal Study of Parents and Children Study Team. Testosterone during pregnancy and gender role behavior of preschool children: A longitudinal, population study. Child Dev. 2002 Nov-Dec;73(6):1678-87.

111. Reinisch JM, Sanders SA. Effects of prenatal exposure to diethylstilbestrol (DES) on hemispheric laterality and spatial ability in human males. Horm Behav. 1992 Mar;26(1):62-75.

112. Ehrhardt AA, Meyer-Bahlburg HF, Rosen LR, Feldman JF, Veridiano NP, Elkin EJ, et al. The development of gender-related behavior in females following prenatal exposure to diethylstilbestrol (DES). Horm Behav. 1989 Dec;23(4):526-41.

113. Lish JD, Ehrhardt AA, Meyer-Bahlburg HF, Rosen LR, Gruen RS, Veridiano NP. Gender-related behavior development in females exposed to diethylstilbestrol (DES) in utero: An attempted replication. J Am Acad Child Adolesc Psychiatry. 1991 Jan;30(1):29-37.

114. Sandberg DE, Vena JE, Weiner J, Beehler GP, Swanson M, Meyer-Bahlburg HF. Hormonally active agents in the environment and children's behavior: Assessing effects on children's gender-dimorphic outcomes. Epidemiology. 2003 Mar;14(2):148-54.

115. Jing H, Pivik RT, Gilchrist JM, Badger TM. No difference indicated in electroencephalographic power spectral analysis in 3- and 6-month-old infants fed soy- or milk-based formula. Matern Child Nutr. 2008 Apr;4(2):136-45.

116. Malloy MH, Berendes H. Does breast-feeding influence intelligence quotients at 9 and 10 years of age? Early Hum Dev. 1998 Jan 9;50(2):209-17.

117. Strom BL, Schinnar R, Ziegler EE, Barnhart KT, Sammel MD, Macones GA, et al. Exposure to soy-based formula in infancy and endocrinological and reproductive outcomes in young adulthood. JAMA. 2001 Aug 15;286(7):807-14.

118. Harkonen PL, Makela SI. Role of estrogens in development of prostate cancer. J Steroid Biochem Mol Biol. 2004 Nov;92(4):297-305.

119. Newbold RR, Padilla-Banks E, Jefferson WN, Heindel JJ. Effects of endocrine disruptors on obesity. Int J Androl. 2008 Apr;31(2):201-8.

120. Schoeters G, Den Hond E, Dhooge W, van Larebeke N, Leijs M. Endocrine disruptors and abnormalities of pubertal development. Basic Clin Pharmacol Toxicol. 2008 Feb;102(2):168-75.

121. Goldman JM, Laws SC, Balchak SK, Cooper RL, Kavlock RJ. Endocrine-disrupting chemicals: Prepubertal exposures and effects on sexual maturation and thyroid activity in the female rat. A focus on the EDSTAC recommendations. Crit Rev Toxicol. 2000 Mar;30(2):135-96.

122. Koo MM, Rohan TE. Accuracy of short-term recall of age at menarche. Ann Hum Biol. 1997 Jan-Feb;24(1):61-4.

123. Cooper R, Blell M, Hardy R, Black S, Pollard TM, Wadsworth ME, et al. Validity of age at menarche self-reported in adulthood. J Epidemiol Community Health. 2006 Nov;60(11):993-7.

124. Casey VA, Dwyer JT, Coleman KA, Krall EA, Gardner J, Valadian I. Accuracy of recall by middle-aged participants in a longitudinal study of their body size and indices of maturation earlier in life. Ann Hum Biol. 1991 Mar-Apr;18(2):155-66.

125. Must A, Phillips SM, Naumova EN, Blum M, Harris S, Dawson-Hughes B, et al. Recall of early menstrual history and menarcheal body size: After 30 years, how well do women remember? Am J Epidemiol. 2002 Apr 1;155(7):672-9.

126. Euling SY, Herman-Giddens ME, Lee PA, Selevan SG, Juul A, Sorensen TI, et al. Examination of US puberty-timing data from 1940 to 1994 for secular trends: Panel findings. Pediatrics. 2008 Feb;121 Suppl 3:S172-91.

127. Wu T, Mendola P, Buck GM. Ethnic differences in the presence of secondary sex characteristics and menarche among US girls: The third national health and nutrition examination survey, 1988-1994. Pediatrics. 2002 Oct;110(4):752-7.

128. Rasier G, Toppari J, Parent AS, Bourguignon JP. Female sexual maturation and reproduction after prepubertal exposure to estrogens and endocrine disrupting chemicals: A review of rodent and human data. Mol Cell Endocrinol. 2006 Jul 25;254-255:187-201.

129. Gallo D, Cantelmo F, Distefano M, Ferlini C, Zannoni GF, Riva A, et al. Reproductive effects of dietary soy in female wistar rats. Food Chem Toxicol. 1999 May;37(5):493-502.

130. Casanova M, You L, Gaido KW, Archibeque-Engle S, Janszen DB, Heck HA. Developmental effects of dietary phytoestrogens in sprague-dawley rats and interactions of genistein and daidzein with rat estrogen receptors alpha and beta in vitro. Toxicol Sci. 1999 Oct;51(2):236-44.

131. Badger TM, Ronis MJ, Hakkak R. Developmental effects and health aspects of soy protein isolate, casein, and whey in male and female rats. Int J Toxicol. 2001 May-Jun;20(3):165-74.

132. Delclos KB, Bucci TJ, Lomax LG, Latendresse JR, Warbritton A, Weis CC, et al. Effects of dietary genistein exposure during development on male and female CD (sprague-dawley) rats. Reprod Toxicol. 2001 Nov-Dec;15(6):647-63.

133. Nagao T, Yoshimura S, Saito Y, Nakagomi M, Usumi K, Ono H. Reproductive effects in male and female rats of neonatal exposure to genistein. Reprod Toxicol. 2001 Jul-Aug;15(4):399-411.

134. Blanck HM, Marcus M, Tolbert PE, Rubin C, Henderson AK, Hertzberg VS, et al. Age at menarche and tanner stage in girls exposed in utero and postnatally to polybrominated biphenyl. Epidemiology. 2000 Nov;11(6):641-7.

135. Freni-Titulaer LW, Cordero JF, Haddock L, Lebron G, Martinez R, Mills JL. Premature thelarche in puerto rico. A search for environmental factors. Am J Dis Child. 1986 Dec;140(12):1263-7.

136. Zung A, Glaser T, Kerem Z, Zadik Z. Breast development in the first 2 years of life: An association with soy-based infant formulas. J Pediatr Gastroenterol Nutr. 2008 Feb;46(2):191-5.

137. Giampietro PG, Bruno G, Furcolo G, Casati A, Brunetti E, Spadoni GL, et al. Soy protein formulas in children: No hormonal effects in long-term feeding. J Pediatr Endocrinol Metab. 2004 Feb;17(2):191-6.

138. Bernbaum JC, Umbach DM, Ragan NB, Ballard JL, Archer JI, Schmidt-Davis H, et al. Pilot studies of estrogen-related physical findings in infants. Environ Health Perspect. 2008 Mar;116(3):416-20.

139. Feigelson HS, Henderson BE. Estrogens and breast cancer. Carcinogenesis. 1996 Nov;17(11):2279-84.

140. Jacobson-Dickman E, Lee MM. The influence of endocrine disruptors on pubertal timing. Curr Opin Endocrinol Diabetes Obes. 2009 Feb;16(1):25-30.

Figure 2.1. Chemical Structure of 17β-estadiol, Genistein and Daidzein (Aglycone)(18)

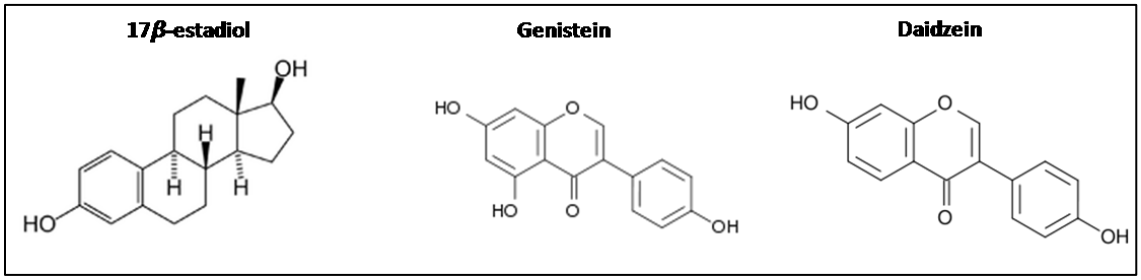


Table 2.1. Isoflavone Concentrations in Infant Biological Samples (Adapted from Cao et al., 2009(22))

Reference	Fluid	N	Age	Isoflavone Concentrations* [% < LOD]		
				Genistein	Daidzein	Equol
SBF Fed						
Cruz et al., 1994	Urine	13	4 mo	~390 [0]	~700 [0]	[~100]
Setchell et al., 1997	Plasma	7	4 mo	684 ± 443 [0]	295 ±59.5 [0]	~ 2.0 [42.9]
Irvine et al., 1998	Urine	4	2 – 16w	2.9-3.8 $mg \cdot kg^{-1} \cdot d^{-1}$ [0]		--
Hoey et al., 2004[a]	Urine	7	4 -6 mo	[a]65 ± 53 [0]	[a]60±32 [0]	[a]0.05±0.10[75]
Cao et al., 2009	Urine	125	0-12 mo	5891 x/÷ 3.1 [0]	5096 x/÷ 2.5 [0]	2.3 x/÷ 4.6 [95]
	Saliva	119	0-12 mo	11.6 x/÷ 5.1 [9]	5.2 x/÷ 5.8 [17]	[100]
	Whole Blood	27	0-12 mo	757 x/÷ 3.0 [4]	256 x/÷ 2.8 [4]	[100]
CMF Fed						
Cruz et al., 1994	Urine	8	4 mo	~15 [0]	~20 [0]	[~100]
Setchell et al., 1997	Plasma	7	4 mo	3.2 ± 0.7 [0]	2.1 ± 0.3 [0]	4.1± 0.5 [0]
Irvine et al., 1998	Urine	25	2 – 16w	[100]	[100]	--
Hoey et al., 2004	Urine	7	4-6 mo	[100]	[100]	[100]
Cao et al., 2009	Urine	128	0-12 mo	11.8 x/÷ 5.7 [9]	8.2 x/÷ 5.0 [22]	2.4 x/÷ 2.1 [78]
	Saliva	120	0-12 mo	0.7 x/÷ 1.3 [95]	0.4 x/÷ 1.4 [93]	[100]
	Whole Blood	30	0-12 mo	14.2 x/÷ 1.5 [90]	5.5 x/÷ 1.5 [97]	[100]
Breast Milk Fed						
Cruz et al., 1994	Urine	12	4 mo	[~100]	[~100]	[~100]
Setchell et al., 1997	Plasma	7	4 mo	2.8 ± 0.7 [0]	1.5 ± 0.1 [0]	~0.5 [85.7]
Franke et al., 2006	Urine	7	2-45 w	29.8 ±11.6 nmol/mg creatine		
	Urine[b]	7	2-45 w	[b]111.6 ± 18.9 nmol/mg creatine		
	Plasma[b]	11	2-45 w	[b,c]19.7 ±13.2 nmol/L		
Cao et al., 2009	Urine	128	0-12 mo	1.5 x/÷ 4.8 [51]	1.5 x/÷ 2.9 [70]	1.7 x/÷ 1.2 [98]
	Saliva	120	0-12 mo	0.7 x/÷ 1.3 [98]	0.4 x/÷ 1.5 [93]	[100]
	Whole Blood	20	0-12 mo	10.8 x/÷ 2.7 [95]	5.3 x/÷ 1.2 [95]	[100]

*Concentration units are reported in ng/ml unless otherwise specified, and represent either mean ± standard deviation or geometric mean x/÷ geometric standard deviation. LOD = limit of detection.

[a] units = ug/mg creatine.

[b] following 2-4 day maternal soy challenge, consisting of one serving of soy beverage per day

[a] approximately equivalent to 5.2 +/- 3.5 ng/ml, using conversion unit 3.8 x 10^{-3} mol/g (Cao et al., 2009)

III. SPECIFIC AIMS AND HYPOTHESES

Aim1: To assess the effects of soy product exposure on early life, sexually dimorphic cognitive and behavioral function.

- Sub Aim 1.1. To assess the effect of early life soy product use on language acquisition within strata of gender
- Sub Aim 1.2. To assess the effect of early life soy product use on gender-role behavior within strata of gender

Hypothesis

The hypothesis is that soy fed males will demonstrate more feminized behaviors, characterized by both a more rapid acquisition of vocabulary, and increased tendency toward feminine play behavior. Given that a rise in testosterone occurs in the postnatal male, it is possible that isoflavone effects on elements of the HPG axis could disrupt this normal trend, resulting in a suppression of the testosterone rise. Such effects on postnatal programming could result in more feminized behaviors in males. It is more difficult to predict how postnatal isoflavone exposure might affect a female's capacity to fully undergo early life sexual differentiation. However, de-feminizing effects of genistein seen in postnatally exposed animal models, such as changes in SDN-POA volume and lordic posturing, suggest that masculinzing effects are possible.

Aim 2: To assess the effect of early life soy product use on time to menarche in adolescent females.

Hypothesis

In accordance with a substantial animal literature on genistein, other phytoestrogens, and other exogenous estrogens, it is hypothesized that SBF exposure will advance the age at menarche in this population.

IV. METHODS

The aims of this study will be investigated using data from the Avon Longitudinal Study of Parents and Children (ALSPAC). This study is a large, population based study of children followed from pregnancy to adolescence. ALSPAC's longitudinal study design and substantial sample size offer a unique and efficient opportunity for addressing the research questions of interest here.

a. Overview of The Avon Longitudinal Study of Parents and Children (ALSPAC)

The Avon Longitudinal Study of Parents and Children (ALSPAC) is an ongoing longitudinal study of parents and children born in the Avon region of the United Kingdom (UK) between April 1991 and December 1992. Over 14,000 pregnancies were recruited into the study, representing approximately 85% of all eligible pregnancies in the region. Live births were followed throughout childhood. Several thousand subjects and their families are still being studies in their late teenage years. Ongoing study activities, as well as management of existing data, are overseen by staff at the University of Bristol, Bristol, UK.

The study area, formerly known as the county of Avon, is the region in western England currently known as Greater Bristol. Recruitment of ALSPAC participants took place within this defined geographic region, exclusive of the area of Bath. Avon has a total population of approximately 1 million, and is comprised of the city of Bristol (population 0.5 million) and surrounding communities, including both rural areas and mid-sized villages and towns {{}} (1, 2). Prior to initiating the ALSPAC study, births in this area were evaluated in comparison to other areas of Great Britain. Children in Avon were shown to be similar to children in

other areas of the country with respect various demographic and descriptive factors, including the proportion of children living with a single parent at age 5 (4%), being non-European non-Caucasian (5.1%), having a parent with a University degree (14.0%), living in an apartment (10.2%), living in a rural area (15%), and having a smoking parent (36.7% for mothers, 49.7% for fathers). There was no difference in prevalence of preterm birth, low birth weight, mental, physical or behavioral problems, or respiratory or allergic disease. Children in Avon were less likely to live in a rented home (26.7%), or to have a father in a manual occupation (51.6%)(1). Generally, though, this region was considered to be well representative of other areas of the country, and thus suitable for a population based study of child development.

Population Eligibility and Recruitment: Efforts to enroll eligible subjects began in September 1990. Eligible subjects included women who were pregnant, residing in Avon, and expecting to deliver between April 1, 1991 and December 31, 1992. Women who enrolled, but moved out of the area in early pregnancy, were not included in the study. Women that moved following completion of a third trimester questionnaire, however, were eligible for continued enrollment, regardless of their place of residence at the time of delivery.

Recruitment followed a combination of clinic and population based methods. Posters advertising the "Children of the Nineties" (ALSPAC) study were displayed in pharmacies, libraries, preschool play groups, mother-toddler play areas, physician waiting rooms, prenatal clinics, and other areas pregnant women may be likely to visit. Local and national news media also frequently discussed the study, and provided information for enrollment. In clinic settings, community midwives were encouraged to discuss the study during first visits with

newly expectant mothers. ALSPAC staff also approached eligible women during routine ultrasound visits, and hospitals distributed information about the study to women along with the mother's 'booking information.' Women were also approached by ALSPAC staff at delivery (1).

14,893 pregnancies were recruited into the study, representing approximately 85% of eligible pregnancies in Avon at the time of recruitment (ALSPAC documentation, Feb. 1996). Of these, 13,995 pregnancies were followed into the postnatal period, along with 14,138 children who survived the first 12 months. 13,978 twin or singleton children are included in data available for analysis. Only term singletons will be included in this analysis (n = 12,931).

Protocols and Response Rates: Data was collected from early pregnancy onward through a variety of methods, including self-completed questionnaires for mothers, their partners, and children; medical and educational record examination; biological specimen collection; and clinic visits and household environmental assessments for sub-sets of the study population. Since the present study will be using data collected from the self-completed questionnaires and medical records, the following discussion will be limited to those instruments. Questionnaires of interest are documented in Table 4.1, along with corresponding response rates.

Pregnancy Questionnaires

A series of four questionnaires were administered during pregnancy. They were completed at various time points depending on the stage of pregnancy at enrollment. Two of these questionnaires, "Having a Baby" and "Your Pregnancy," were intended to be completed at 18 and 32 weeks' gestation, respectively. They aimed to capture attitudes,

perceptions and environments at specific points in pregnancy. However, women enrolled later in pregnancy were allowed to complete these questionnaires as late as 23 weeks' and 40 weeks', as needed. A questionnaire specifically interested in early pregnancy exposures, "Your Environment," was only administered to women who enrolled prior to 23 weeks' gestation. A separate questionnaire, "Your Home and Lifestyle," was given in lieu of "Your Environment" and "Having a Baby" to women who enrolled past 23 weeks'. The "Your Home and Lifestyle" questionnaire was designed to capture early life environment and lifestyle characteristics that could still be validly reported later in pregnancy. For subjects that could not complete the 32 week "Your Pregnancy" Questionnaire, several questions regarding race, education, and other descriptive variables were presented again in a follow up questionnaire at 12 months postpartum, "Filling the Gaps." Finally, "About Yourself" was administered without respect to time in pregnancy to all women. Since this involved questions about past medical, social, and environmental history, it was not sensitive to time of enrollment. It was thus administered throughout a wide range of gestational ages, and in some cases, it was administered after delivery (1).

Following a questionnaire mailing, if a response from the mother was not received within 7 days, a reminder letter was sent. A second reminder was mailed if a response was not received within 10 days of the first reminder. If no response was received 1 month after the second mailed reminder, study staff would either call or visit the mother to encourage her to complete the questionnaire(s), and provide assistance if necessary.

Delivery Outcomes

The study obtained information about live births and deaths using government mandated birth notifications made available to health authorities. Notifications of live births included

date of delivery, birth weight, sex of child, and singleton versus multiple birth status. This notification also alerted study staff to the time at which they should send the first post-natal questionnaire. For births occurring outside the Avon area, a questionnaire was mailed to mothers approximately 8 weeks after the expected date of delivery. Information regarding birth weight and other delivery related items were reported from the mother at this time (ALSAC documentation).

Child-Based (Postnatal) Questionnaires

Beginning at 4 weeks postpartum, a series of "child based questionnaires" were sent to the home of the child and could be completed by the child's mother, father, or other carer. If a pregnancy resulted in multiple births, a separate questionnaire was mailed for each child born. A total of 22 child based questionnaires were mailed out between the ages of 4 weeks and 166 months (13.8 years). The present study is mostly interested in the child-based questionnaires up to age 42 months, where topics including child health, feeding/eating, sleeping habits, temperament and behavior, growth, milestones of physical and cognitive development, child care, and household characteristics are assessed. The mailing schedule for these questionnaires is shown in Table 4.1. It is important to note that questionnaires were made available to participants at the "scheduled" time. However, completed questionnaires were frequently returned several weeks or months later, so the actual age of competition for each questionnaire may vary substantially.

Puberty Questionnaires

Questionnaires intended to assess physical markers of pubertal development, entitled "Growing and Changing" were first mailed when the child reached age 97 months (8.1 years). Five "Growing and Changing" questionnaires followed, through age 175 months

(14.6 years) (Table 4.1). Other maternal, partner and child-completed questionnaires were also sent to the household at these times. The "Growing and Changing" questionnaires collected self-rated data on physical stages of puberty, as well as specific dates for pubertal milestones, such as age at menarche. Separate questionnaires were designed for males and females. Between 97 months (8.1 years) and 157 months (13.1 years), these questionnaires were mailed directly to the "carer," and could be completed by the parent alone, the child alone, or the parent and child together. The 175 month (14.6 years) questionnaire was the first puberty questionnaire sent directly to the teenager. Perhaps as a result, it has a slightly lower response rate than the previous questionnaires. The lower response rate may also have resulted from girls achieving and reporting relevant puberty milestones on questionnaires prior to 175 months. To better characterize these response rates, patterns of response and non-response will be evaluated prior to analysis, particularly with respect to each subject's demographic and developmental characteristics. While most respondents completed the 175 month questionnaire themselves, about 10% reported seeking help from a parent or other individual.

Partner Questionnaires

Select partner information was obtained for the present study in order to gather information regarding paternal demographic, health, and parenting-style information.

b. **Methods: Aim 1**

Aim1: To assess the effects of SBF exposure on early life, sexually dimorphic cognitive and behavioral function.

- Sub Aim 1.1. To assess the effect of SBF on language acquisition within strata of gender

- Sub Aim 1.2. To assess the effect of SBF on gender-role behavior within strata of gender

i. Hypothesis

The hypothesis is that SBF fed males will demonstrate more feminized behaviors, characterized by both a more rapid acquisition of vocabulary, and increased tendency toward feminine play behavior. Given that a rise in testosterone occurs in the postnatal male, it is possible that isoflavone effects on elements of the HPG axis could disrupt this normal trend, resulting in a suppression of the testosterone rise. Such effects on postnatal programming could result in more feminized behaviors in males. It is more difficult to predict how postnatal isoflavone exposure might affect a female's capacity to fully undergo early life sexual differentiation. However, de-feminizing effects of genistein seen in postnatally exposed animal models, such as changes in SDN-POA volume and lordic posturing, suggest that masculinzing effects are possible.

ii. Study Sample

The study sample included subjects of the ALSPAC study, described above, who were

1) term infants (gestational age $\geq$ 37 weeks),
2) singleton births,
3) alive at 1 year
4) had sufficiently complete infant feeding data provided at 6 or 15 months of age, and
5) who completed both a 15 and 24 month vocabulary production assessment (MacArthur Communicative Development Inventory) within a designated time frame (Sub-Aim 1), or who completed a 42 month gender-role evaluation (Preschool Activity Inventory) (Sub-Aim 2).

The study population was limited to term singletons alive at 1 year. Preterm infants (gestational age < 37 weeks) are likely to have both health problems, and some developmental delays, which may influence both SBF use and language acquisition. Likewise, language development and play behavior are likely to be heavily influenced by the presence of a twin sibling. Therefore, excluding preterm and twin births will reduce confounding related to these factors. Of the 14,663 births on which data is available, 390 births were excluded on the basis of a preterm or unknown gestational age. This preterm birth exclusion also included 188 twin births. An additional 202 term twin births were also excluded. Finally, an additional 49 subjects were excluded that were not alive at 1 year, for a total eligible sample size of 12,931.

Subjects were excluded if the data collected on infant feeding methods was not sufficient to characterize infant feeding habits, as described in the *Exposure Assessment and Definition* section below. 4,502 subjects were excluded based on this criterion.

Additional exclusions were made for Sub-Aim 1.1. Subjects that completed the outcome assessment more than 4 months after the intended assessment date (i.e., after 19 months for the 15 month language assessment, or after 28 months for the 24 month assessment) were excluded (n= 43 at 15 months; n = 54 at 24 months). These restrictions were intended to maximize the validity of outcome assessment, since the instrument used for language assessment was designed for children of specific ages. Also, any child with diagnosed complete or partial deafness (n= 27 in eligible population) or later diagnosis of autism (n = 71 in eligible population) was excluded. These characteristics may be related to early life illness (and thus, soy formula use) and also affect language development skills.

Subjects that reported initiating soy formula exposure after 15 months of age were not included in this analysis. This exclusion was required to maintain a proper temporal relationship between the exposure and outcome for Sub-Aim 1.1, since the first language assessment was conducted at 15 months of age. Also, by excluding this population from both the exposed and unexposed classifications, the overall exposure contrast is sharpened to emphasize soy exposures in the first year of life, which are of most interest in this study. Exposures occurring after 15 months of age are probably minimal in comparison to those occurring earlier in infancy. This exclusion applies to 210 subjects out of the 12, 931 otherwise eligible subjects.

iii. Exposure, Outcome and Covariate Assessment

Exposure Assessment and Definition: Mothers completed infant feeding questionnaires at 1, 6, 15 and 24 months postpartum. At 1 month, mothers reported all feeding methods used since birth (breast or bottle), and the type of formula used, if any. At 6, 15, and 24 months, mothers reported current breastfeeding habits, the age at which other milks or formulas were introduced into the child's diet (including formula/baby milk, soy milk, soy formula, goat's milk, hypo-allergenic formula, and cow's milk), and how many feedings per week were given for each of these products at the time of questionnaire completion. Relevant questions regarding formula and milk feeding are shown here,

At 6 months,

Item C1.A. Ordinary Baby Milk
Has your baby ever had ordinary baby milk (formula)? [yes/no]
At what age did your baby start ordinary baby milk (formula)? [age in months]
How often nowadays is your baby fed ordinary baby milk (formula)? [times/week at time of questionnaire completion].

Item C1.C. Soya Milk

Has your baby ever had soya milk? [yes/no]
At what age did your baby start soya milk? [age in months]
How often nowadays is your baby fed soya milk? [times/week at time of questionnaire completion].

Additional feeding options, including follow-on milk (a diet supplement, typically used after 6 months), goat's milk, hypoallergenic formula, and ordinary cow's milk are also assessed in this manner.

At 15 months,

Item D6.A. <u>*Baby Milk (formula)*</u>
Since your child was 6 months old, has he/she had baby milk (formula)? [yes/no]
At what age did your child start baby milk (formula)? [age in months]
How often nowadays is your child fed baby milk (formula)? [times/week at time of questionnaire completion]

Item D6.C. <u>*Soya Formula*</u>
Since your child was 6 months old, has he/she had soya formula? [yes/no]
At what age did your child start soya formula? [age in months]
How often nowadays is your child fed soya formula? [times/week at time of questionnaire completion]

Item D6.F <u>*Soya Milk*</u>
Since your child was 6 months old, has he/she had soya milk? [yes/no]
At what age did your child start soya milk? [age in months]
How often nowadays is your child fed soya milk? [times/week at time of questionnaire completion]

Using this data, along with the reported age that breast feeding ended, subjects were categorized into four mutually exclusive feeding groups: *primarily breastfed, early formula, early soy,* and *late soy* (Figure 4.1). Exposure classification was primarily defined by responses to the questionnaire administered at 6 months postpartum; if these data were missing or incomplete, responses from the 15-month questionnaire were used according to the following criteria. The same exposure definition was used for Sub Aim 1.1 and Sub Aim 1.2.

"Early" exposure to any type of formula or milk (soy or traditional) was defined as the use of a specific formula or milk type occurring ≤ 4 months of age ("*At what age did you start [formula/milk type]?")* through ≥ 6 months of age. Use at 6 months was indicated by any non-zero response to the question, "*How often nowadays is your baby fed [formula/milk type]?"* in the 6 month questionnaire. If the 15 month questionnaire was used instead, "early" exposure to formula was established for any subject that reported introducing the formula or milk ≤ 4 months of age and responded affirmatively to the question "*Since your child was 6 months old, has he/she had [formula/milk type]?"* This definition not only establishes early use of formula, but also a 1 month minimum duration of use.

To note, there is some ambiguity between the terminology of "soya milk" and "soya formula" in these questionnaires. In the 6 month questionnaire, respondents may construe *Item C1.C* as a question regarding either soy-based infant formula, or soy milk. Therefore, exposure definition used in this study will include early life exposure to both soy formula and soy milk. As described in Section II., both types of soy product contain high levels of isoflavones. Exposure to either product, individually or in combination, should result in a large dose of isoflavone compounds. Applying the term birth and singleton birth inclusion criteria, and the age at completion criteria, 796 infants are identified as having *any* soy exposure prior to 15-19 months of age using the questions listed above.

Overall, the exposure groups were classified as follows: *Primarily breastfed infants* were those breast fed until ≥ 6 months of age, with no reported soy use between birth and 24 months and no reported introduction of other milks or formulas before 6 months of age. *Early formula fed infants* were those introduced to any non-soy milk or formula product at or before 4 months of age, who use such products at 6 months of age [*How often nowadays [6-*

10months] is your child fed [non-soy milk/formula]?" **> 0 times/week,** or, *Since your child was 6 months old, has he/she had [non-soy mllk/formula]?"* ***YES]***, and reported no soy use before 24 months. *Early soy fed infants* were those introduced to soy milk or soy formula at or before 4 months of age, and reported sustained use at 6 months of age (Figure 4.2). *Late soy fed infants* were those introduced to soy milk or soy formula any time after 4 months of age through 15 months of age. No restrictions were made in the early formula, early soy, and late soy groups with respect to duration of breast feeding; likewise, there were no restrictions in the early soy or late soy groups with respect to use of non-soy milk or formula.

This study is population-based, and therefore reflects a very "real world" scenario in which infant feeding practices change over time, and often include mixing or supplementation between formula and breast feeding. "Exposed" infants may be exposed to some combination of breast milk, CMF, SBF and soy milk. "Unexposed" infants may be exposed to some combination of breast milk and CMF. To address this, in part, breast feeding will be included as an *a priori* confounder in analysis. The age of soy product introduction is also provided in Appendix I, Table A.1 to provide information on the variability of exposure within this group.

Outcome Assessment:

Language Acquisition: ALSPAC administered modified versions of the MacArthur Communicative Development Inventory (MCDI) at 15, 24 and 38 months. The MCDI is a widely used instrument in clinical and research settings, intended to assess the growth and variability of communicative skills in early life (3). Separate MCDI instruments have been developed for infants (age 8 – 16 months), and toddlers (age 16-30 months). Only the 15 and

24 month assessments were used here; the instrument was not age-appropriate for children as old as 38 months.

In the MCDI, parents indicate which words their child can understand (comprehension) or say (production), according to a list of provided vocabulary words. In the modified version administered here, a list of 124 words was assessed at 15 months, and 135 slightly more complex words were assessed at 24 months. The ALSPAC modified version of the 15 month CDI has been used to gauge cognitive development in studies of prenatal nutrition and environmental factors by Daniels et al., 2004 and Daniels et al., 2007(4, 5). Subjects who completed this questionnaire between 15 and 19 months of age were included in analysis. The 24 month assessment closely resembles the toddler MCDI (3), except in an abbreviated format. This modified version assesses the distinction between "words understood" and "words produced," whereas the original MCDI only assesses word production. Subjects who completed this questionnaire between 24 and 28 months of age will be included in analysis.

Scores for word comprehension and word production were calculated as the sum of understood or spoken words reported by the parents at 15 months for the 15 month scores, and as the sum of the 15 month word score plus the new words spoken at 24 months for the 24 months scores. Twenty-eight words were included on both the 15 and 24 month assessment, but were only counted in scores on the first occasion that they were reported.

Preschool Activities Inventory (PSAI): Childhood play activities are easily observed behavioral characteristics that tend to show clear, gender specific distinctions, even at early ages. Therefore, play behavior is an ideal outcome to use in the assessment of gender roles in children. In this study, gender role behavior was used as a marker to assess effects of early

life soy isoflavone exposure on sexually dimorphic behavioral development. The Preschool Activities Inventory (PSAI) was the instrument used to measure this effect.

The PSAI is an assessment of play behavior among preschool aged children, based on maternal report of child involvement in various sex-typical play behaviors (6). It is composed of three sections: toy preference, activities, and characteristics. In each section, mothers (or caregivers) are instructed to report on how often her child has played with a particular type of toy (7 items), engaged in a particular activity (11 items), and displayed a particular characteristic (6 items), for the past month. Half of these items are "masculine" activities or behaviors, and half are "feminine." Possible responses include "never," "hardly ever," "sometimes," "often," and "very often." Each response is scored on a 5-point scale. The total instrument is scored by summing responses to all masculine items, and subtracting the sum of all feminine items, and applying a transformation (48.25 + (1.1*[Score]), to achieve a "pseudo-T score." Higher scores indicate masculine-typical behavior, and lower scores indicate feminine-typical behavior (6, 7).

This test is a desirable measure of gender role behavior for several reasons. First, the PSAI was developed with the intention for use in research of gender-role behavior, such as studies of the influence of family structure or of gender stereotypes on gender development (7). Therefore, it should be adequately designed for the current research purposes. Additionally, it is specifically designed to assess gender role behavior in children under age 5, which includes the target population for the current study. Finally, in addition to being able to discriminate between play behaviors of males and females, it is also designed to characterize differences in play behavior among children of the same gender, which is of primary interest for the current study.

The instrument has been standardized on over 2,000 children in the UK, US, and Netherlands (6). The population mean score based on these three populations is 51.10. The mean score for boys is 61.66 (SD = 9.4), and the mean score for girls is 38.72 (SD = 9.66), indicating that that this instrument can efficiently detect a pronounced difference between gender populations. Sex-specific age standardizations can be applied to scores, and used when within gender comparisons are of interest.

Certain psychometric properties of this instrument are documented by its authors (7), but are somewhat uninformative due to use of small samples for several of the reliability and validity measures. Based on test-retest and split-half reliability assessments, the PSAI does appear to have moderate reliability. Gender-specific test-retest reliability after 1 year is reported as 0.62 for boys and 0.66 for girls (based on 15 boys and 18 girls), although this value has been criticized for the "re-test" occurring too long after the original test (8). The gender-specific split-half reliability is 0.66 for boys (n = 1260) and 0.80 for girls (n = 1070). PSAI has also been shown to have moderate agreement with teacher ratings of gender behavior, although neither the PSAI nor teacher ratings can be considered a "gold standard." In a population of 45 boys and 57 girls who attended day care (mean age 45.7 months), maternally completed PSAI scores, compared to teacher rankings of a child's "boyish" or "girlish" behavior, had correlations of 0.48 ($p < 0.0002$) for girls and 0.37 ($p < 0.01$) for boys.

The PSAI has been used in several other research studies of gender-role behavior. Increased levels of maternal testosterone during pregnancy, for example, were positively associated with more masculine behavior in 3.5 year old females in the ALSPAC study population (9). Other studies using the PSAI have examined effects of perinatal PCB and dioxin exposure (10), prenatal stress (11) and influences of older siblings (12).

In the ALPSAC study, the PSAI was included in the 30 month, 42 month, and 57 month child-based questionnaires. This study was limited to analysis of the 42 month assessment only, since the age of this population most closely resembles the standardization populations' (mean ages 35.8 -51.4 months). A recent study of the PSAI in the ALSPAC population also suggests that gender-role behavior remains relatively stable over time (13), so there was little need to utilize repeated measures to assess this outcome.

A PSAI score was calculated using a scoring algorithm that accounts for each subject's responses to masculine and feminine items, such that the lower scores represent more feminine behavior and higher scores represent more masculine behavior. Males and females were analyzed separately to look for within gender variations attributable to exposure to SBF.

Covariate Assessment:

A preliminary assessment was performed to determine how certain covariates differed across feeding groups, with particular attention paid to predictors of soy product use. Key variables of interest included various measures of infant health, demographic characteristics, and other characteristics that were associated with the outcomes of interest in relevant literature (Appendix Table A.2). This assessment informed, in part, which covariates were included as potential confounders. Information was collected on key covariates through several questionnaires administered during pregnancy and early childhood. Relevant variables are listed below.

- ***Infant Health:*** Various measures of infant health were acquired during the 6 month questionnaire, including several that might be indications for soy formula or soy milk use. Parental report of "ever/never" for each of the following infant health outcomes was assessed as a dichotomous variable: diarrhea, vomiting, cough, earache, ear discharge, colic, and rash.

- ***Cow's Milk Allergy:*** Child food allergies were assessed in multiple questionnaires. The presence of early allergy to milk was documented in the 6 month questionnaire,

- ***Crying/Fussiness***: Excessive crying or fussiness may be associated with colic, or colic-like symptoms, and thus may be related to soy product use. Excessive crying was assessed at 4 weeks and 6 months of age:

  At 4 weeks:

  ItemD4.a. Do you feel that your child's crying is a problem?[Yes/No]

- ***Maternal Demographics:*** The following maternal characteristics were assessed at various time points in gestation and childhood.

  Maternal Age: *Age of Mother at Birth [years]*

  Race: *At 32 weeks' gestation:*

  > *Item H8. How would you describe the race or ethnic group of yourself, your partner and your parents?[white; black/Caribbean; black/African; black/other (please describe below); Indian; Pakistani; Bangladesh; Chinese; any other ethnic group (please describe)]*

  Education: *At 32 weeks' gestation:*

  > *Mother's Highest Educational Qualifications*
  >
  > *[CSE/None; Vocational; O-Level; A-Level; Degree]*

  Marital Status: *At 8 weeks postpartum:*

  > *F2. a) What is your present marital status?*
  >
  > *[never married; widowed; divorced; separated; married (once); married for 2nd or 3rd time]*

Race was dichotomized to white vs. other. Child's race was further described in the

ALSPAC data as being white if both parents reported being white, and non-white otherwise.

- ***Infant Characteristics/Demographics***

  Gender: from birth records and various questionnaires

  Birth weight: from birth records

Gestational Age: derived from medical records. The variable associated with gestational age is entitled "Best Gestational Age We Could Get."

Additional Potential Confounders for Sub Aim 1.1: Language Acquisition Rate: In addition to the covariates described above, several additional covariates were assessed as potential confounders of the association between feeding group and language acquisition rate. These additional covariates are described below.

- **Ear problems**: Dairy intolerance may lead to a variety of health complications, including chronic ear infections. Accordingly, ear infections may be an impetus to switch from breast or CMF to SBF. Poor ear health may also lead to hearing difficulties that may induce a delay in language skills. Therefore, subjects that have partial or complete deafness were excluded. Furthermore, models were adjusted for having a "suspected hearing problem," as determined by parent report, at 6 months of age.

- **Breast feeding:** Initiation and duration of breastfeeding can be identified using 1, 6 and 15 month questionnaires. Duration of breastfeeding will be determined based on the reported age that breastfeeding stopped, using 1 and 6 month reports in preference of 15 month reports if a there is a discrepancy in reporting. If breastfeeding was never initiated, a duration of 0 was be assigned. Otherwise, a duration in months was assigned. This was be treated as a continuous variable, which can be categorized if needed.

- **Maternal Parenting Score:** Interaction with the child is likely to influence language acquisition rates. It may also be related to infant feeding if one supposes a more involved mother may be more willing to seek out infant feeding methods with perceived health benefit, such as soy. A maternal parenting score, assessed at 18 months, was included as a confounder in language assessment models.

- **Preschool Attendance:** Involvement in a preschool or nursery school environment may advance language acquisition. Nursery school attendance may also impede one's capacity to breastfeed, or may increase a child's rate of infection, which may lead to increased use of either soy formula or traditional formula. Accordingly, regular use of a day nursery was reported at 15 months of age and was included as a potential confounder in this analysis.

- **Presence of older sibling:** Older siblings may advance a child's rate of language acquisition, as well as influence a mothers feeding choice. The presence of older brothers and sisters was assessed at 18 months

- **Pre/Postnatal Smoking:** Exposure to environmental tobacco smoke may influence child health, which may lead to soy product use, and may also affect infant neurologic development. Exposure to tobacco smoke was defined as a dichotomous yes/no variable based on perinatal maternal tobacco use, determined via the question,

  At 8 weeks postnatal,
  B4. Did you smoke regularly in the last 2 months of pregnancy and since having the baby?

- **Maternal Vegan/Vegetarian Status:** At 32 weeks' gestation, and 97 months postpartum:

  Are you, or have you ever been, a vegan (i.e. do/did not eat meat, poultry, fish, eggs, butter, milk or cheese)? [yes, I am now/ yes, in past not now/ no, never]

  Are you, or have you ever been, a vegetarian? [yes, I am now/ yes, in past not now/ no, never]

Vegan/Vegetarian status during pregnancy, and historical reports of vegan/vegetarian status at 97 months postpartum were assessed to infer vegan/vegetarian status at the time of formula feeding. This may inform the motivation for soy formula/soy milk use. Status during pregnancy may also indicate higher exposure to isoflavones during the prenatal period, which is also likely to influence the outcomes of interest.

Additional Potential Confounders for Sub-Aim 1.2: Gender Role Behavior: Several of the variables defined above were assessed as confounders with respect to Sub Aim 1.2. These included breast feeding, preschool attendance, presence (and gender) of older siblings, maternal education, perinatal smoking, and race. In addition, maternal and partner interaction scores at 42 months were assessed as a potential confounders (Appendix Table A.2).

iv. Analysis

Aim 1.1. Language Acquisition: All analyses were conducted separately for boys and girls, so as to highlight within gender differences in language development with respect to

exposure. The population-averaged effect of feeding method on word comprehension and word production over time was assessed using generalized estimating equations (GEE), with a compound symmetric covariance. These models were used to determine the mean effect across exposure groups at 15 and 24 months, and how these means were related across the two time points (14). The change in mean over time was quantified as the interaction between feeding group and age at assessment, expressed as continuous months. The early formula group was used as the referent for all models.

Simplified models were also run that compared the crude mean MCDI scores for word comprehension and word production at each time point separately. Mean differences were compared using linear regression.

GEE models were adjusted for variables associated with feeding method and with language development, as determined by assessment in the data or in relevant literature. Final adjusted models included breast feeding duration, maternal age, presence of an older sibling (yes/no), maternal parenting score at 18 months, daycare attendance at 15 months (yes/no), and suspected hearing problem at 6 months (yes/no), according to the Figure 4.3.

Aim 1.2 Gender-role Behavior: Crude mean PSAI scores were assessed as simple means within exposure groups and within strata of each covariate. Differences in means were assessed using univariable linear regression. Adjusted mean difference in PSAI scores were estimated using multivariable linear regression. Boys and girls were modeled separately to emphasize within gender differences. Unless otherwise noted, the early formula feeding group was used at the referent for all feeding group comparisons.

Confounders included variables thought to be associated with both cultural and environmental influences of gender role behavior, as well as feeding method. Final models

were adjusted for age at PSAI assessment, breast feeding duration (months), presence of an older brother (yes/no), presence of an older sister (yes/no), regular attendance in daycare (yes/no), and maternal factors including age, prenatal smoking status (yes/no), education (5 levels in the United Kingdom, ranked from high to low: University Degree, Advanced Level, Ordinary Level, Vocational, and Certification of Secondary Education (CSE)/None), and interaction score, as illustrated in Figure 4.4. A partner interaction score was also included in models, given the supposed importance of male influence in the home on gender-role behavior, however it did not influence results. Partner and maternal interaction scores were estimated at 42 months of age using a series of questions that assessed the frequency at which each parent participated in a list of 8 activities with the child (score range: 0-36). Partner interaction scores were set to zero if the questionnaire reported that no partner was present. All partners were assumed to be male, given a very low prevalence of mothers in same-sex partnerships in this cohort (<1%) (22).

Missing Data: Eligible subjects that were included in these analyses were compared to other eligible subjects in ALSPAC that were excluded on the basis of incomplete feeding information, outcome information, or other exclusion criteria.

Given the large proportion of data missing in the PSAI analysis, models were analyzed using both complete case analysis (CCA) and multiple imputation (MI) approaches. For the CCA, subjects with missing data for adjustment variables were dropped from models, so only true values for subjects with complete data were modeled. Multiple imputation was implemented to allow the entire study sample to be included in analyses. Values for missing covariates were estimated using PROC MI. Regression models using 5 iterations of imputed

data were run and summarized in PROC MIANALYZE. All analyses were completed using SAS 9.1.3 (SAS Institute Inc., Cary, NC).

v. Sample Size and Power

For Aim 1.1, the sample size is

- Primarily Breast: 683 boys and 684 girls
- Early Formula: 2487 boys and 2329 girls
- Early Soy: 85 boys and 58 girls
- Late Soy: 129 boys and 105 girls

The power calculations presented here show power estimates for the detection of a significant difference in the mean number of words comprehended or produced at a single point in time, at the $\alpha = 0.05$ level, for the comparison between early soy and early formula (Table 4.2). A range of power estimates are given for a range of plausible standard deviations that may be observable in the ALSPAC data. These reflect results from 2-tailed significance tests.

For males, our sample size allows for the detection of a difference of 11 – 15 words with a standard deviation ranging between 30 and 45. For females, a 15 -20 word difference can be detected, given various standard deviations. Comparisons between the early formula and other feeding groups should amply powered to detect more precise results given the increased sample sizes in both the late soy and primarily breast fed groups. Furthermore, when GEE should also allow for a more efficient analysis than a simple, time-specific linear regression model, so these estimates are conservative with respect to the statistical power of the GEE model.

For Aim 1.2, the sample size is

- Primarily Breast: 707 boys and 706 girls

- Early Formula: 2699 boys and 2490 girls
- Early Soy: 89 boys and 68 girls
- Late Soy: 132 boys and 112 girls

The power calculations presented here show power estimates for the detection of a significant difference in PSAI score, at the $\alpha = 0.05$ level, for the comparison between early soy and early formula (Table 4.3). A range of power estimates are given for a range of plausible standard deviations that may be observable in the ALSPAC data. These reflect results from 2-tailed significance tests.

In males, this analysis should have 80% power to detect a 3 point difference in PSAI score between these two exposure groups. In females, a 3-4 point difference is detectible, depending on the standard deviation of mean estimates.

c. Methods: Aim 2

Aim 2: To assess the effect of early life soy product use on time to menarche in adolescent females.

i. Hypothesis

In accordance with a substantial animal literature on genistein, other phytoestrogens, and other exogenous estrogens, it is hypothesized that SBF exposure will advance the age at menarche in this population.

ii. Study Sample

The study sample will include subjects of the ALSPAC study, described above, who were

1) female,
2) term infants (gestational age $\geq$ 37 weeks),

3) singleton births,
4) white,
5) had sufficiently complete infant feeding characterization, and
6) who completed at least one "Growing and Changing" puberty questionnaire (administered between the ages of 8 and 14).

iii. Exposure, Outcome and Covariate Assessment

Exposure Assessment and Definition: The exposure definition used for Aims 1.1 and 1.2 will also be applied here.

Outcome Assessment: Between 1999 and 2007, a series of questionnaires regarding pubertal development, known as the Growing and Changing questionnaires, were administered at approximately 8, 9.5, 10.5, 11.5,13 and 14.5 years of age, as previously described by Rubin et al., 2009 (15). Questionnaires were completed by a care-giving adult or the child of interest. Between 8 and 13 years of age, questionnaires were mailed to parents or guardians. At 14.5 years of age, the questionnaire was mailed directly to the child

In each questionnaire, subjects were asked if the child had had her first period, and if so, what month and year her first period occurred. This date was used in combination with the child's birth date to provide an age in months for menarche. The earliest reported age was used as the age at menarche in the event that multiple questionnaires contained discordant responses for the same individual.

ALSPAC also enrolled some subjects from the main study into a smaller clinic-based cohort study that also assessed age at menarche. For 146 (4.6%) subjects in the present

analysis, , missing questionnaire data on age at menarche was imputed from ALSPAC clinic data.

107 subjects reported a menarche event, but did not report an age. For 67 of these subjects, age at menarche was estimated as the midpoint between the age at which the questionnaire with the first positive menarche response was completed, and the age at which the previous year's questionnaire was completed. If more than one questionnaire was skipped between a negative and positive menarche response, and estimated age was not derived and this subject was not included in any analysis. As an alternative, an imputed age at menarche was estimated for these 67 subjects as part of a larger multiple imputation model for missing data (described below). Ages derived using the midpoint approach were included in the complete case analysis, while imputed values were used in multiple imputation models.

Covariates: Variables examined as potential confounders include child's birth weight, breast feeding duration, milk allergy at 6 months, vegetarian diet in childhood, and maternal perception of infant health, and maternal factors including age and education at delivery, prenatal vegetarian diet, age at menarche, pre-pregnancy body mass index (BMI), race, and pre- and postnatal smoking. Covariates were included as potential confounders if they were associated with an infant feeding method, age at menarche, or censoring in these data or in relevant literature. Duration of breast feeding was not included as confounder, despite having an equivocal association with age at menarche in previous literature (16, 17). Final models were adjusted for pre-pregnancy BMI, smoking in the last 2 months of pregnancy (yes/no), and maternal age at menarche. Since this population is >90% white, analyses were restricted to the white children (defined as having two white parents) to control for any effects due to race. Effect measure modification was examined in relation to childhood

weight status by stratifying above and below the 85th percentile of BMI for age z-scores (18), using measurements taken at any time between 7 and 9.5 years. Associations between feeding group and BMI Z-scores were assessed using linear regression (Appendix Table A.3). The distribution of characteristics among the lost to follow up subjects, including stratum specific median time to menarche values, is shown in Table A.4.

iv. Analysis

All analyses were completed using SAS 9.1.3 (SAS Institute Inc., Cary, NC). Hazard ratios (HR) for time to menarche were estimated using Cox proportional hazards modeling. Relative precision of estimated HRs were compared using confidence limit ratios, calculated as the upper 95% confidence limit divided by the lower 95% confidence limit. The early formula group was used as the referent group in all models, unless otherwise specified. Proportional hazards assumptions (PHA) were assessed using log-log survival density function plots and Cox chi-squared significance tests for time interaction variables ($\alpha = 0.05$), which revealed varying degrees of hazard non-proportionality over time (convergence). This PHA violation can be expected given that, over time, menarche "failure" will inevitably approach 100%, as previously noted by Warner et al.(19). Therefore, models were carried out using categorical ($< / \geq 150$ months (12.5 years)) and continuous-time interactions (exposure*time, pre-pregnancy BMI*time, prenatal smoking*time, and maternal age at menarche*time), to characterize changes in hazards over time. Hazard ratios were estimated at 10,11,12,13 and 14 years for continuous-time Cox models. Cox models were analyzed using both complete case analysis and multiple imputation (MI). For the complete case analysis, only subjects with complete data on necessary covariates were modeled, which resulted in a loss of 18% of subjects. Multiple imputation of missing outcome (n = 67),

adjustment variables and other covariates (BMI) was performed using PROC MI (5 imputations), and simulated results were combined using PROC MIANALYZE. As described previously, final models were adjusted for pre-pregnancy BMI, smoking in the last 2 months of pregnancy (yes/no), and maternal age at menarche (Figure 4.5).

Follow up time was defined for each subject based on the age in months at which she reported a menarche event (*events*), or the age of the last completed questionnaire in which she reported not having reached menarche (*censored).* Censored subjects were distinguished as either lost to follow up, or administratively censored. *Lost to follow up* refers to subjects that dropped out before the end of the study period (14.5 years of age) before reporting an event. *Administratively censored* refers to subjects that completed the study period, but never reported a menarche event. Crude median time-to-menarche and inter-quartile range (IQR: 25^{th}-75^{th} Percentile) estimates were obtained using lifetable analysis (PROC LIFETEST). We obtained counfounding-adjusted median age at menarche by exposure group by calculating inverse probability of exposure weights (20) using polytomous logistic regression, and then performing a weighted lifetable analysis. Crude and adjusted time-to-menarche Kaplan Meier curves were obtained similarly (21).

Sensitivity analyses were performed to assess whether results were affected by informative random censoring (23). In the first analysis, we assumed that all *lost to follow up* subjects were at low risk for reaching menarche in the study period, and their follow-up times were reassigned to resemble the administratively censored (follow-up time: 175 months) ("low risk"). In the second analysis, we assumed that randomly censored subjects were at high risk for reaching menarche in the study period, and all *lost to follow up* follow-up times were modeled as events occurring at the time of drop out ("high risk"). A third sensitivity

analysis was preformed to address early censoring times in the early soy group. Here, we manipulated the distribution of censor times in the early soy group to mimic the censor time distribution seen in other exposure groups (median: ~140 months) ("redistributed"). Categorical-time Cox models were repeated for each of the three hypothetical scenarios to evaluate how these assumptions affected estimates.

Effect measure modification was examined in relation to childhood weight status by stratifying above and below the 85th percentile of BMI, using measurements taken at any time between 7 and 9.5 years. BMI stratified results are presented as time-averaged HRs, since PHA violations were minimal in the stratified samples, and sample size limitations prevented further stratification by time.

v. Sample Size and Power

The sample size for Aim 2 is shown below.

- Primarily Breast: 620
- Early Formula: 2124
- Early Soy: 54
- Late Soy: 86

The power for a "time to event" analysis comparing early soy to early formula was calculated using a simulation in nQuery software. Estimates were based on 100% survival prior to age 10, and 22 and 35% survival rates at the end of study at age 14. The sample size identified above is adequate for detecting these estimates with 81% power at the $\alpha = 0.05$ significance level. It should be strongly noted that survival rates and dropout rates are estimates and approximations, not actual values based on the data. It is therefore possible that power may decrease when actual survival rates and dropout rates are taken into account. The

estimates used for the simulation are shown in Table 4.4 and the results are shown in Table 4.5.

REFERENCES

1. Golding J, Pembrey M, Jones R, ALSPAC Study Team. ALSPAC--the avon longitudinal study of parents and children. I. study methodology. Paediatr Perinat Epidemiol. 2001 Jan;15(1):74-87.

2. Golding J, ALSPAC Study Team. The avon longitudinal study of parents and children (ALSPAC)--study design and collaborative opportunities. Eur J Endocrinol. 2004 Nov;151 Suppl 3:U119-23.

3. Fenson L, Dale PS, Reznick JS, Bates E, Thal DJ, Pethick SJ. Variability in early communicative development. Monographs of the Society for Research in Child Development. 1994;59(5, Serial No. 242):i-185.

4. Daniels JL, Longnecker MP, Rowland AS, Golding J, ALSPAC Study Team. University of Bristol Institute of Child Health. Fish intake during pregnancy and early cognitive development of offspring. Epidemiology. 2004 Jul;15(4):394-402.

5. Daniels JL, Rowland AS, Longnecker MP, Crawford P, Golding J, ALSPAC Study Team. Maternal dental history, child's birth outcome and early cognitive development. Paediatr Perinat Epidemiol. 2007 Sep;21(5):448-57.

6. Golombok S, Rust J. The measurement of gender role behaviour in pre-school children: A research note. J Child Psychol Psychiatry. 1993 Jul;34(5):805-11.

7. Golombok S, Rust J. The pre-school activities inventory: A standardized assessment of gender role in children. Psychological Assessment. 1993;5(2):131-6.

8. Kaufman AS. Critique of vreugdenhil et al.'s study linking PCBs to the play behaviors of dutch girls and boys. Environ Health Perspect. 2003 Jun;111(7):A380; author reply A380-1.

9. Hines M, Golombok S, Rust J, Johnston KJ, Golding J, Avon Longitudinal Study of Parents and Children Study Team. Testosterone during pregnancy and gender role behavior of preschool children: A longitudinal, population study. Child Dev. 2002 Nov-Dec;73(6):1678-87.

10. Vreugdenhil HJ, Slijper FM, Mulder PG, Weisglas-Kuperus N. Effects of perinatal exposure to PCBs and dioxins on play behavior in dutch children at school age. Environ Health Perspect. 2002 Oct;110(10):A593-8.

11. Hines M, Johnston KJ, Golombok S, Rust J, Stevens M, Golding J, et al. Prenatal stress and gender role behavior in girls and boys: A longitudinal, population study. Horm Behav. 2002 Sep;42(2):126-34.

12. Rust J, Golombok S, Hines M, Johnston K, Golding J, ALSPAC Study Team. The role of brothers and sisters in the gender development of preschool children. J Exp Child Psychol. 2000 Dec;77(4):292-303.

13. Golombok S, Rust J, Zervoulis K, Croudace T, Golding J, Hines M. Developmental trajectories of sex-typed behavior in boys and girls: A longitudinal general population study of children aged 2.5-8 years. Child Dev. 2008 Sep-Oct;79(5):1583-93.

14. Applied longitudinal data analysis: Lecture notes ST 732 [homepage on the Internet]. Raleigh: North Carolina State University. 2005. Available from: http://www.stat.ncsu.edu.libproxy.lib.unc.edu/people/davidian/st732/#notes.

15. Rubin C, Maisonet M, Kieszak S, Monteilh C, Holmes A, Flanders D, et al. Timing of maturation and predictors of menarche in girls enrolled in a contemporary british cohort. Paediatr Perinat Epidemiol. 2009;23:492-504.

16. Novotny R, Daida YG, Grove JS, Acharya S, Vogt TM. Formula feeding in infancy in associated with adolescent body fat and earlier menarche. Cell Mol Biol. 2003;49(8):1289-93.

17. Blell M, Pollard TM, Pearce MS. Predictors of age at menarche in the newcastle thousand families study. J Biosoc Sci. 2008;40(4):563-75.

18. A SAS program for the CDC growth charts [homepage on the Internet]. . 2009 November 16, 2009. Available from: http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm.

19. Warner M, Samuels S, Mocarelli P, Gerthoux PM, Needham L, Patterson DG, et al. Serum dioxin concentrations and age at menarche. Environ Health Perspect. 2004;112(13):1289-92.

20. Hernan MA, Robins JM. Estimating causal effects from epidemiological data. J Epidemiol Community Health. 2006 Jul;60(7):578-86.

21. Cole SR, Hernan MA. Adjusted survival curves with inverse probability weights. Comput Meth Prog Bio. 2004;75:45-9.

22. Remsberg KE, Demerath EW, Schubert CM, Chumlea WC, Sun SS, Siervogel RM. Early menarche and the development of cardiovascular disease risk factors in adolescent girls: The fels longitudinal study. J Clin Endocrinol Metab. 2005 May;90(5):2718-24.

23. Allison PD. Survival analysis using the SAS system: A practical guide. Cary, NC: SAS Press; 1995.

Figure 4.1. Exposure Characterization

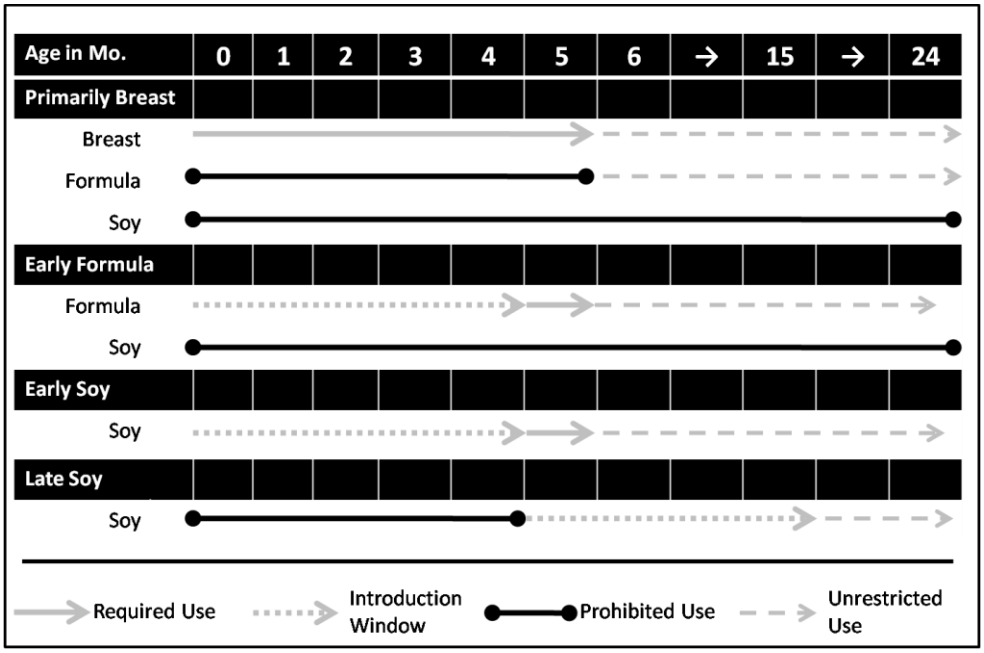


Figure 4.2. 'Early Soy' Exposure Criteria: 4 to 6 Month Exposures

1) *Item C1.C:*

"*Has your baby ever had soya milk?"* ***YES***

"*At what age did you start soya milk?"* ***≤ 4 months***

"How often nowadays [6-10months] is your child fed soya milk?"

> 0 times/week,

Or,

2) *Item D6.C:*

"*Since your child was 6 months old, has he/she had soya formula?"* ***YES***

"At what age did your child start soya formula?" ***≤ 4 months***

Or,

3) *Item D6. F:*

"*Since your child was 6 months old, has he/she had soya milk?"* ***YES***

"At what age did your child start soya milk?" ***≤ 4 months***.

Figure 4.3. Conceptual diagram of association between infant feeding and language acquisition, noting confounding relationships.

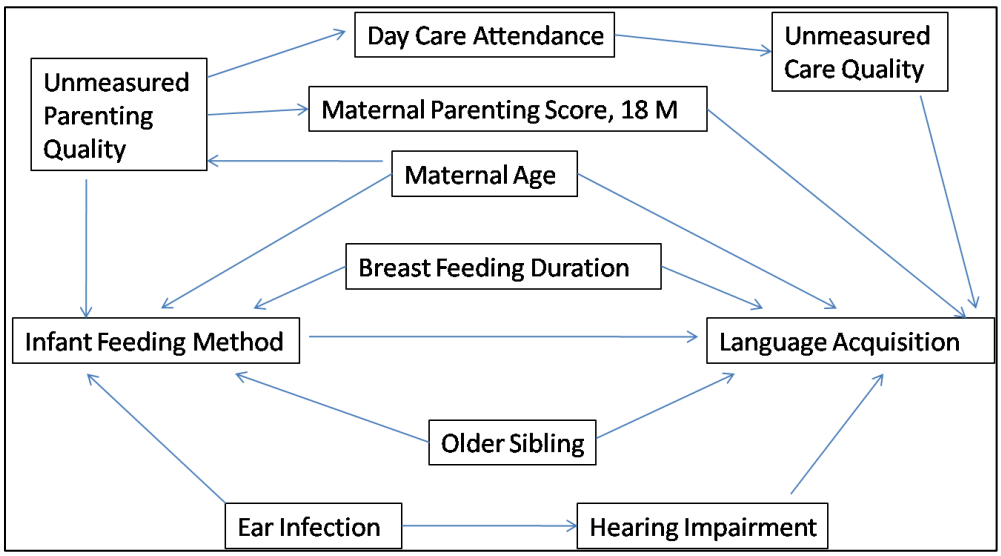


Figure 4.4. Conceptual diagram of association between infant feeding and gender-role play behavior, noting confounding relationships.

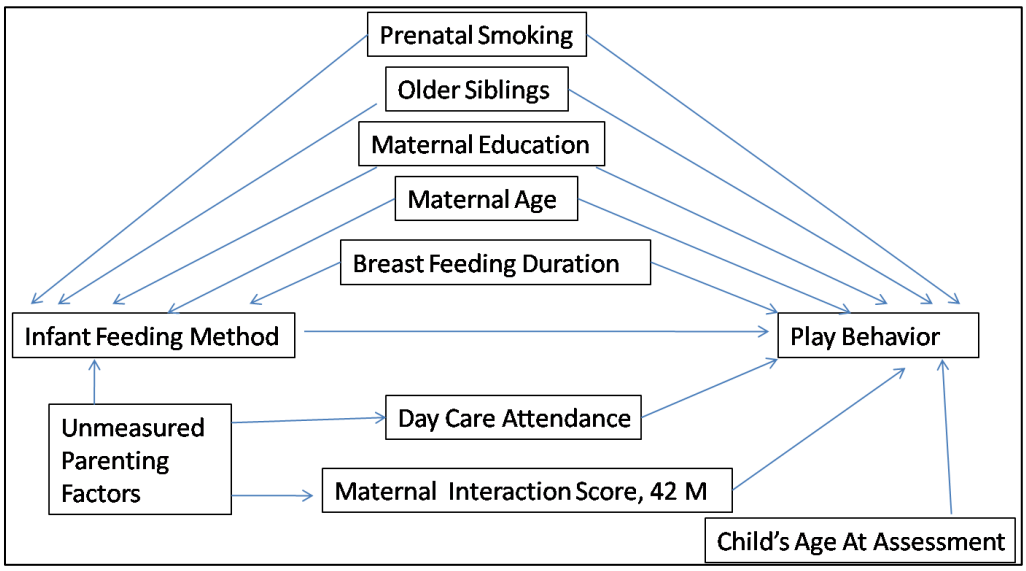


Figure 4.5. Conceptual diagram of association between infant feeding and time-to-menarche, noting confounding relationships.

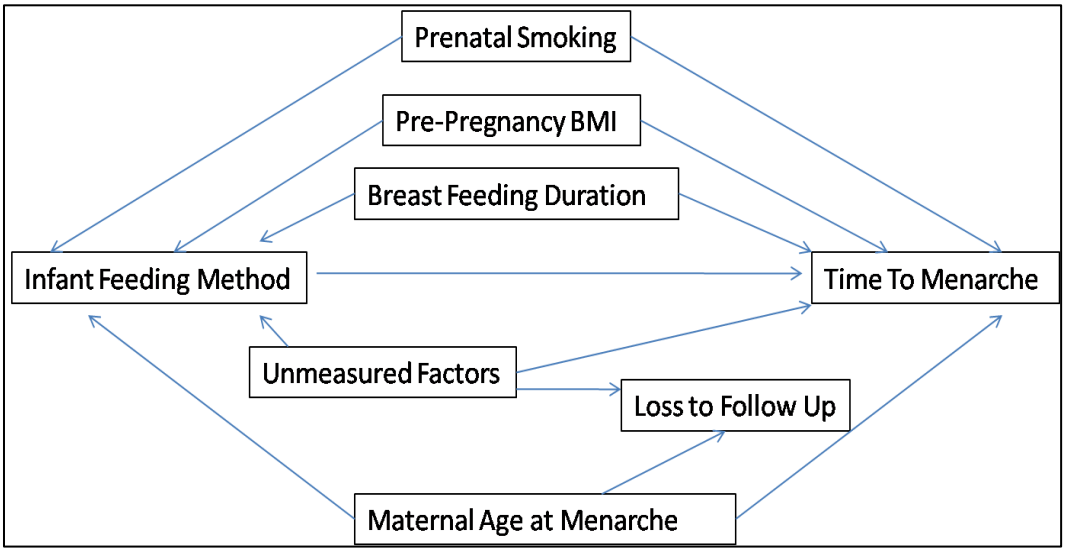


Table 4.1. Number of Respondents Per Questionnaire, According to ALSPAC Documentation

Questionnaire	Goal Age for Administration	# of Subjects	% of Enrolled Pregnancies*	% of Births*	% of 4 week Postnatal Respondents*
Maternal					
Your Environment	< 28 wks G	12,571	84.4		
Having a Baby	< 23 wks G	12,213	82.0		
Your Home and Lifestyle[1]	24-41 wks G	977	6.6		
Your Pregnancy	32-40 wks G	12,084	81.1		
Filling the Gaps[2]	12 m	334	2.2		
About Yourself	14 wks - PD	12,471	83.7		
Child Based					
My Young Baby Boy/Girl	4 wks	12,353	82.9	88.4	
My Son/Daughter	6 m	11,485	77.1	82.2	92.97
My Infant Son/Daughter	15 m	11,073	74.4	79.2	89.64
Boy/Girl Toddler	18 m	10,750	72.2	76.9	87.02
My Little Girl/Boy	24 m	10,431	70.0	74.6	84.44
My Study Son/Daughter	30 m	10,359	69.6	74.1	83.86
My 3-year old Girl/Boy	38 m	10,145	68.1	72.6	82.13
My S/D's Health and Behavior	42 m	10,083	67.7	72.1	81.62
My Young 4 Year Old Boy/Girl	54 m	9722	65.3	69.6	78.70
Puberty					
Growing and Changing 1	97 m	6255	42.0	44.7	50.64
Growing and Changing 2	115 m	7017	47.1	50.2	56.80
Growing and Changing 3	128 m	6629	44.5	47.4	53.66
Growing and Changing 4	140 m	6293	42.3	45.0	50.94
Growing and Changing 5	157 m	6075	40.8	43.5	49.18
Growing and Changing 6	175 m	5163	34.7	36.9	41.80
Partner					
You and Your Environment	12 wks G	8645	58.0	61.8	69.98

[1] Responses pooled with Your Environment and Having a Baby

[2] Responses pooled with Your Pregnancy

*pregnancy denominator = 14,893, birth denominator = 13,978, 4 week denominator = 12,353

G = Gestation PD = Post-Delivery

Table 4.2. Power estimations for mean word comprehension and production differences

			Range of Standard Deviations			
	Early Soy/ Early Formula	Difference in mean words	30	35	40	45
Males	85/2487	9	77	64	53	44
		11	91	81	70	60
		13	97	92	83	47
		15	99	97	92	85
Females	58/2329	9	61	48	39	32
		11	78	65	54	45
		13	90	79	68	58
		15	95	89	80	70
		20	99	99	96	91

Table 4.3. Power estimations for mean PSAI score differences

			Range of Standard Deviations		
	Early Soy/ Early Formula	Difference in mean PSAI Score	8	9	10
Males	89/2699	2	64	54	45
		3	93	87	79
		4	99	98	96
Girls	68/2490	2	52	43	36
		3	86	77	68
		4	98	95	90

Table 4.4. Time-dependent hazard info for log-rank survival test

	1	2	3	4	5	6
End of period, time t	0.00	10.000	11.000	12.000	13.000	14.000
Accrual (% of total)	100.00	0.000	0.000	0.000	0.000	0.000
Group 1 exponential hazard rate	0.000	0.0020	0.1054	0.3052	0.3242	0.2800
Group 2 exponential hazard rate	0.000	0.0033	0.2960	0.4265	0.2948	0.4643
Group 1 expected % surviving time t	100.00	98.000	88.200	65.000	47.000	35.530
Group 2 expected % surviving time t	100.00	96.800	72.000	47.000	35.000	22.000
Common exponential dropout rate,d	0.00	0.0000	0.0000	0.0050	0.0250	0.0100

Table 4.5. Log-rank test of survival in two groups, simulation with specified rates (unequal n's)

Test significance level, alpha	0.050
1 or 2 sided test?	2
Number of periods	5
n_1	2124
n_2	54
Power (%)	81

V. AIM 1.1 SUMMARY

Background:

In a typical child, word comprehension begins around 9 months of age, and word production begins around 12 months of age. Following 12 months, a "fan effect" occurs in which rapid word producers and slow word producers begin to differentiate. At around16 to 18 months, a significant difference can be observed between males and females in both the number of words understood, and the number of words produced (1), with females consistently demonstrating higher word production than males (2-7).

Language acquisition rates can be affected by a number of environmental influences in addition to age and sex. These may include social factors, such as parental tendency to engage verbally with the child ("language input"), or presence of an older sibling. Physical factors such as oral development, hearing, and general health may also influence language development.

In addition to the social and physical environment, sex-specific biologic development may also play a large role in influencing language acquisition. Specifically, the biological basis for observed sex differences in language development may be related to sexual differentiation of the brain in early life. Sex hormone concentrations can influence sexually dimorphic brain development, and consequently, the manifestation of sex-specific learning and behavior characteristics (8). Under this mechanism, typical sex-specific developmental patterns may, in theory, be altered by exposure to endocrine disrupting compounds in early life. This analysis will assess whether exposure to soy-based infant formula, common source

of infant nutrition that contains high levels of estrogen-like plant based compounds, affects the rate of language development in boys and girls between the ages of 15 and 24 months.

Methods:

Study Sample. The Avon Longitudinal Study of Parents and Children (ALSPAC) is an ongoing, prospective, longitudinal study which enrolled pregnant women residing in the Avon region of the United Kingdom, who were expected to deliver between April 1, 1991 and December 31, 1992 . Women were informed of the study by community clinicians or local media campaigns, and were recruited into the study after expressing interest in participation. 14,062 live births were recruited into the study during pregnancy. Of these, 13,978 (7220 boys and 6756 girls) were twins or singletons alive at one year. Mothers provided consent for participation. Ethical approval for the study was obtained from the ALSAPC Law and Ethics Committee and the Local Research Ethics Committees.

The present investigation was restricted to term singletons who were alive at one year of age (n = 8,492), for whom comprehensive infant feeding data was available and for whom a language assessment was completed at approximately 15 and 24 months of age. Subjects were excluded if they completed either questionnaire more than 4 months past the intended administration time or had a missing completion age, if they had total or partial deafness, or were later diagnosed with autism spectrum disorder. The final sample size for this analysis is $n_{boys} = 3384$, $n_{girls} = 3176$. The present analysis was approved by the Institutional Review Board of the University of North Carolina at Chapel Hill.

Exposure: Exposure was assessed using methods identical to those described elsewhere in this document. Subjects were categorized into four mutually exclusive feeding groups:

primarily breastfed, early formula, early soy, and late soy, based on previously described definitions.

Outcome: The outcomes assessed here included word comprehension and word production at approximately 15 and 24 months, as assessed by a modified version of the MacArthur Communicative Development Inventory (MCDI). The MCDI is a widely used instrument in clinical and research settings, intended to assess the growth and variability of communicative skills in early life (1). In the MCDI, parents indicate which words their child can understand (comprehension) or say (production), according to a list of provided vocabulary words. In the modified version administered here, a list of 134 words was assessed at 15 months, and 123 slightly more complex words were assessed at 24 months. Scores for word comprehension and word production were calculated as the sum of understood or spoken words reported by the parents at 15 months for the 15 month scores, and as the sum of the 15 month word score plus the new words spoken at 24 months for the 24 months scores. Twenty-eight words were included on both the 15 and 24 month assessment, but were only counted in scores on the first occasion that they were reported.

Analysis: All analyses were conducted separately for boys and girls, so as to highlight within sex differences in language development with respect to exposure. Crude mean MCDI scores for word comprehension and word production were calculated as simple averages for each of the feeding groups. Mean differences were compared using linear regression. Each time point was assessed separately in preliminary analyses.

The population-averaged effect of feeding method on word comprehension and word production over time was assessed using generalized estimating equations (GEE) with a compound symmetric covariance assumption. These models were used to determine the

mean effect across exposure groups at 15 and 24 months, and how these means are related across the two time points (9). The change in mean over time was quantified as the interaction between feeding group and age at assessment, expressed as continuous months. The early formula group was used as the referent for all models.

GEE models were adjusted for variables associated with feeding method and with language development, as determined by assessment in this data, or in relevant literature. Final adjusted models included breast feeding duration, maternal age, presence of an older sibling (yes/no), maternal parenting score at 18 months, daycare attendance at 15 months (yes/no), and suspected hearing problem at 6 months (yes/no).

Results:

The characteristics of the study sample, as compared to the total eligible ALSPAC cohort, are shown in Table 5.1. Generally, the study sample is similar to the source cohort. Minimal differences between these groups included slightly lower word production scores at 15 months (14.5 vs. 15.2), longer breast feeding duration (4.1 months vs. 3.8 months) and maternal age (28.8 years vs. 28.0 years), and a higher proportion of subjects with no older siblings (45.7% vs. 44.2%) and of mothers with a university degree (15.2% vs. 13.8%) in the study sample versus the eligible ALSPAC cohort (not accounting for distributions of missing data).

Mean word comprehension and word production values at 15 and 24 months, and the crude difference in these means, are shown in Tables 5.2 and 5.3. The crude change in mean scores is also shown graphically in Figure 5.1. In boys, word comprehension at 15 months and 24 months is higher in those that were breast fed, compared to the early formula group. However, the overall difference in mean words comprehended (i.e., words acquired between

15 and 24 months) is higher in all feeding groups compared to the early formula group. A similar effect is seen in girls, with the exception that there is no difference in girls' comprehension scores at 15 months across feeding groups. Word production was elevated at 24 months among breast fed boys and among early soy fed girls.

GEE model results are shown in Tables 5.4 and 5.5. The mean changes over time, as expressed by β, demonstrate similar trends as were observed in the mean difference comparisons discussed above. Adjustment did not substantially affect results. Early soy was associated with a small increase in mean change in word comprehension in boys. Word production among early and late soy fed boys was also elevated, but with poor precision. Both outcomes modestly increased in primarily breast fed boys, as compared to the early formula feeders. In girls, the effect of early soy exposure on word comprehension was larger than in boys, but imprecise in comparison to the effect observed in the primarily breast fed. Word production in early soy exposed girls increased by 1.46 words per month compared to early formula fed girls (95% CI: 0.56, 2.29); no effect was observed in the other feeding groups with respect to this outcome.

Conclusions:

Here, we demonstrate a small and imprecise increase in word comprehension and word production associated with soy product exposure in early life, adjusting for multiple social and environmental factors, in both boys and girls.

Since girls typically acquire language more rapidly than boys, an increase in rate of vocabulary development between 15 and 24 months may be interpreted as a "feminized" effect. Accordingly, the small increase in word comprehension, as well as the small and imprecise increase in word production, among early and late soy fed boys is consistent with

animal literature suggesting that male mammals exposed to the estrogenic components of soy demonstrate feminine characteristics (10-14). However, these results should be interpreted cautiously with respect to broad conclusions regarding endocrine disruption. It is not clear whether such modest changes are truly indicative of biological changes in brain development, particularly since similar effects were observed in the primarily breast fed boys as well. Furthermore, this study is not sufficiently powered to detect subtle differences between either group of soy fed boys and early formula fed boys; effect estimates among the soy groups are imprecise and subject to Type II error.

The effect of early soy exposure on girls' word production was the most substantial of all effects observed. Following the endocrine disruptor hypothesis, it could be argued that early life exposure to exogenous estrogen may be enhancing the development of female-typical structures and functions, thus enhancing soy exposed girls to be "hyper feminine." However, as in boys, the effects seen in girls are small and imprecise. It is not appropriate to make broad conclusions on the effects of soy and female brain development. This finding, however, does highlight an interesting hypothesis that should be explored in future studies.

No effect was observed in the late soy exposed girls. This may be due mostly to low statistical power. However, in the absence of a Type II error, this finding also may suggest that if biologically active components of soy are inducing changes in sexually dimorphic development, it is likely happening in the early, as opposed to later stages of infancy in girls. This is in contrast to boys, where a similar magnitude of effect was observed in both the early and late soy feeding groups, highlighting potentially variable stages of susceptibility between the sexes.

Unmeasured confounding is important to acknowledge here. Language development is a complex, multi-factorial process. While sex, age, and many of the social environmental factors used in this study are important in determining the rate at which language develops, other important factors play a role as well. These factors may include the quality of "language input" from parents, or the quality of childcare that the child received. It is possible, for instance, that early soy fed girls were fed soy for reasons that also coincided with increased adult attention (illness, for example), which subsequently enhanced language development. Although our models attempted to control for such factors, by including terms such as the maternal parenting score, it is still important to note that factors other than biologically based endocrine disruption may have influenced findings.

Since this study only had two time points to assess the rate of language acquisition, we are limited in the conclusions that can be drawn, particularly with respect to change over time. Future studies should integrate multiple language assessments to facilitate the development of language development trajectories. Additionally, there were few subjects in both the early soy and late soy groups. This led to largely imprecise estimates for both groups. It is possible that some of the null findings reported here may be a result of low statistical power. Since the prevalence of soy use is low in this population, future studies should be attempted in areas with higher proportions of soy use during infancy, such as the United States.

REFERENCES

1. Fenson L, Dale PS, Reznick JS, Bates E, Thal DJ, Pethick SJ. Variability in early communicative development. Monographs of the Society for Research in Child Development. 1994;59(5, Serial No. 242):i-185.

2. Bauer DJ, Goldfield ba, Reznick JS. Alternative approaches to analyzing individual differences in the rate of early vocabulary development. 2002;23:313-35.

3. Westerlund M, Lagerberg D. Expressive vocabulary in 18-month-old children in relation to demographic factors, mother and child characteristics, communication style and shared reading. Child Care Health Dev. 2008 Mar;34(2):257-66.

4. Stokes SF, Klee T. Factors that influence vocabulary development in two-year-old children. J Child Psychol Psychiatry. 2009 Apr;50(4):498-505.

5. Berglund E, Eriksson M, Westerlund M. Communicative skills in relation to gender, birth order, childcare and socioeconomic status in 18-month-old children. Scand J Psychol. 2005 Dec;46(6):485-91.

6. Bornstein MH, Haynes MO, Painter KM. Sources of child vocabulary competence: A multivariate model. J Child Lang. 1998 Jun;25(2):367-93.

7. Reilly S, Wake M, Bavin EL, Prior M, Williams J, Bretherton L, et al. Predicting language at 2 years of age: A prospective community study. Pediatrics. 2007;120(6):e1441-9.

8. Scallet AC, Divine RL, Newbold RR, Delclos KB. Increased volume of the calbindin D28k-labeled sexually dimorphic hypothalamus in genistein and nonylphenol-treated male rats. Toxicol Sci. 2004 Dec;82(2):570-6.

9. Applied longitudinal data analysis: Lecture notes ST 732 [homepage on the Internet]. Raleigh: North Carolina State University. 2005. Available from: http://www.stat.ncsu.edu.libproxy.lib.unc.edu/people/davidian/st732/#notes.

10. Slikker W,Jr, Scallet AC, Doerge DR, Ferguson SA. Gender-based differences in rats after chronic dietary exposure to genistein. Int J Toxicol. 2001 May-Jun;20(3):175-9.

11. Lund TD, Rhees RW, Setchell KD, Lephart ED. Altered sexually dimorphic nucleus of the preoptic area (SDN-POA) volume in adult long-evans rats by dietary soy phytoestrogens. Brain Res. 2001 Sep 28;914(1-2):92-9.

12. Lephart ED, Setchell KD, Handa RJ, Lund TD. Behavioral effects of endocrine-disrupting substances: Phytoestrogens. ILAR J. 2004;45(4):443-54.

13. Lephart ED, West TW, Weber KS, Rhees RW, Setchell KD, Adlercreutz H, et al. Neurobehavioral effects of dietary soy phytoestrogens. Neurotoxicol Teratol. 2002 Jan-Feb;24(1):5-16.

14. Wisniewski AB, Klein SL, Lakshmanan Y, Gearhart JP. Exposure to genistein during gestation and lactation demasculinizes the reproductive system in rats. J Urol. 2003 Apr;169(4):1582-6.

Figure 5.1. Change in mean word comprehension and production over time, by feeding group

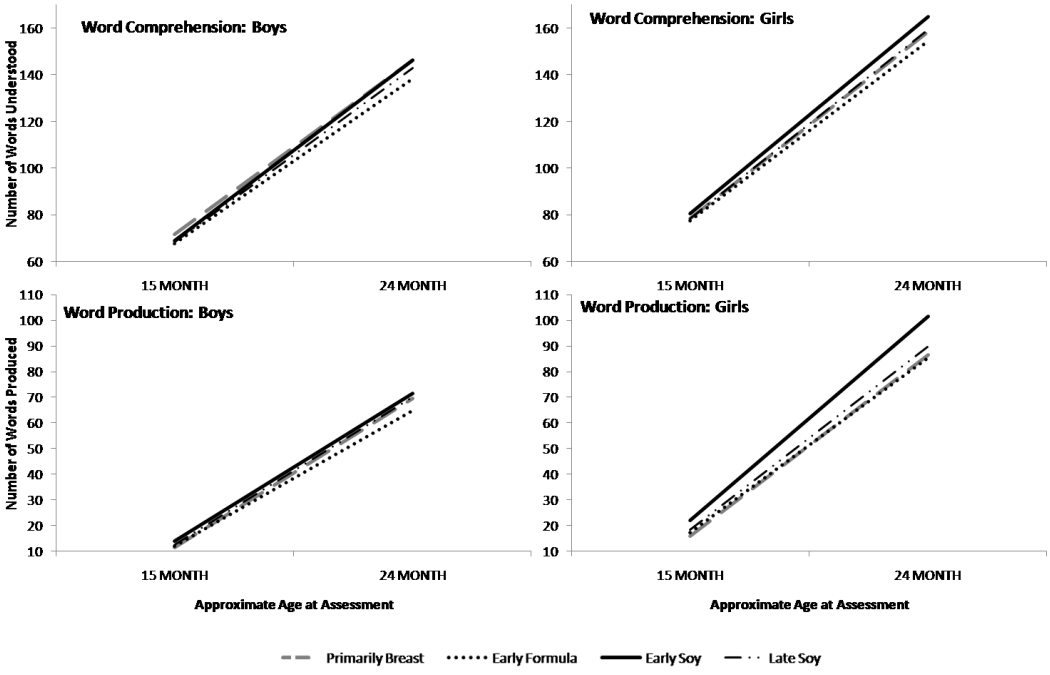


Table 5.1. Study Sample Characteristics

Characteristics	Study Sample			Eligible ALSPAC Subjects[1] (n = 12,931)
	Boys (n = 3384)	Girls (n = 3176)	Total (n = 6560)	
Word Comp Score, 15 M, mean (SD)	68.6 (31.5)	77.8 (31.0)	73.0 (31.6)	73.3 (32.4)
Missing, n	0	0	0	2679
Word Prod Score, 15 M, mean (SD)	12.0 (16.1)	17.1 (19.8)	14.5 (18.2)	15.2 (19.1)
Missing, n	0	0	0	2679
Word Comp Score, 24 M, mean (SD)	140.2 (43.7)	155.8 (41.8)	147.8 (43.5)	147.4 (44.2)
Missing, n	0	0	0	3789
Word Prod Score, 24 M, mean (SD)	66.2 (43.0)	86.0 (43.7)	75.8 (44.4)	17.9 (44.5)
Missing, n	0	0	0	3789
Breast Feeding Duration, mean (SD)	4.0 (4.5)	4.2 (4.6)	4.1 (4.6)	3.8 (4.5)
Missing, n	0	0	0	1011
Maternal Age, mean (SD)	28.9 (4.7)	28.7 (4.6)	28.8 (4.6)	28.0 (5.0)
Missing, n	0	0	0	0
Maternal Parenting Score, mean (SD)	40.4 (4.6)	41.0 (4.3)	40.7 (4.4)	40.7 (4.5)
Missing, n	93	70	163	2626
Infant Feeding Method, n (%)				
Early Formula	2487 (73.5)	2329 (73.3)	4816 (73.4)	6294 (74.7)
Early Soy	85 (2.5)	58 (1.8)	143 (2.2)	182 (2.6)
Late Soy	129 (3.8)	105 (3.3)	234 (3.6)	286 (3.4)
Primarily Breastfed	683 (20.2)	684 (21.5)	1367 (20.8)	1667 (19.8)
Missing, n	-	-	-	4502
Presence of Older Sibling, n (%)				
No	1465 (44.7)	1445 (46.8)	2910 (45.7)	4537 (44.2)
Yes (≥ 1)	1811 (55.3)	1646 (53.3)	3457 (54.3)	5717 (55.8)
Missing, n	108	85	193	2677
Regular Daycare Attendance, n (%)				
No	3151 (93.9)	2955 (94.0)	6106 (93.9)	9574 (94.1)
Yes	206 (6.1)	190 (6.0)	396 (6.1)	605 (5.9)
Missing, n	27	31	58	2752
Suspected Hearing Problem, n (%)				
No	3163 (96.8)	2981 (97.0)	6144 (96.9)	10,213 (96.4)
Yes	106 (3.2)	91 (3.0)	197 (3.1)	380 (3.6)
Missing, n	115	104	219	2338
Maternal Education, n (%)				
University Degree	467 (14.6)	471 (15.8)	938 (15.2)	1497 (13.8)
Advanced Level	814 (25.5)	767 (25.8)	1581 (25.6)	2607 (24.1)
Ordinary Level	1229 (38.5)	1112 (37.4)	2341 (38.0)	4002 (36.9)
Vocational	316 (9.9)	277 (9.3)	593 (9.6)	1132 (10.5)
CSE[2]/None	366 (11.5)	348 (11.7)	714 (11.6)	1596 (14.7)
Missing, n	192	201	393	2097

[1] Eligible subjects include term, singleton births alive at 1 year of age

[2] Certificate of Secondary Education

Table 5.2. Mean Word Comprehension MCDI Scores

Feeding Group[a]				
		Boys (n = 3384)		
	n	**15 M**	**24 M**	**Mean Difference**
ES	85	69 (34)	146 (47)	77 (23)*
LS	129	68 (31)	143 (45)	74 (21)*
PB	683	72 (31)*	146(41)*	74 (19)*
EF (ref)	2487	68(32)	138 (44)	70 (22)
		Girls (n = 3176)		
	n	**15 M**	**24 M**	**Mean Difference**
ES	58	81 (31)	165 (37)	84 (15)*
LS	105	78 (30)	159 (40)	81 (20)*
PB	684	78 (31)	158 (42)*	80 (19)*
EF (ref)	2329	78 (31)	154 (42)	77 (20)

[a] ES: Early Soy; LS: Late Soy; PB: Primarily Breast; EF: Early Formula

*$P < 0.05$

Table 5.3. Mean Word Production MCDI Scores

Feeding Group[a]				
		Boys (n = 3384)		
	n	**15 M**	**24 M**	**Mean Difference**
ES	85	14 (18)	72 (45)	58 (34)
LS	129	13 (16)	70 (42)	57 (32)
PB	683	12 (15)	70 (42)*	58 (34)*
EF (ref)	2487	12 (16)	65 (43)	53 (34)
		Girls (n = 3176)		
	n	**15 M**	**24 M**	**Mean Difference**
ES	58	22 (25)	101 (44)*	79 (28)*
LS	105	19 (23)	90 (46)	71 (31)
PB	684	16 (19)	86 (44)	70 (32)
EF (ref)	2329	17 (20)	85 (43)	68 (32)

[a] ES: Early Soy; LS: Late Soy; PB: Primarily Breast; EF: Early Formula

*$P < 0.05$

Table 5.4. Mean response estimates for MCDI word comprehension from 15 to 24 months

Feeding Group[a]			
		Boys	
		Crude	Adjusted[b]
	n	β (95% CI)	β (95% CI)
ES	84	0.66 (0.13, 1.18)	0.66 (0.13, 1.18)
LS	120	0.41 (-0.06, 0.88)	0.41 (-0.05, 0.88)
PB	639	0.37 (0.18, 0.57)	0.37 (0.18, 0.57)
EF	2303	0.	0.
		Girls	
		Crude	Adjusted[b]
	n	β (95% CI)	β (95% CI)
ES	52	0.87 (0.36, 1.38)	0.87 (0.36, 1.38)
LS	96	-0.00 (-0.48, 0.48)	-0.01(-0.49, 0.43)
PB	637	0.19 (0.10, 0.48)	029 (0.10, 0.48)
EF	2187	0.	0.

[a] ES: Early Soy; LS: Late Soy; PB: Primarily Breast; EF: Early Formula

[b]Adjusted for breast feeding duration, maternal age, presence of an older sibling (yes/no), maternal parenting score at 18 months, daycare attendance at 15 months (yes/no), and suspected hearing problem at 6 months (yes/no)

Table 5.5. Mean response estimates for MCDI word production from 15 to 24 months

Feeding Group[a]			
		Boys	
	n	Crude β (95% CI)	Adjusted[b] β (95% CI)
ES	84	0.51 (-0.31, 1.32)	0.50 (-0.32, 1.31)
LS	120	0.50 (-0.19, 1.19)	0.51 (-0.18, 1.19)
PB	639	0.46 (0.14, 0791)	0.47 (0.14, 0.79)
EF	2303	0.	0.
		Girls	
	n	Crude β (95% CI)	Adjusted[b] β (95% CI)
ES	52	1.43 (0.55, 2.30)	1.46 (0.56, 2.29)
LS	96	-0.05 (-0.72, 0.63)	-0.06 (-0.73, 0.61)
PB	637	0.27 (-0.03, 0.59)	0.27 (-0.04, 0.59)
EF	2187	0.	0.

[a] ES: Early Soy; LS: Late Soy; PB: Primarily Breast; EF: Early Formula

[b]Adjusted for breast feeding duration, maternal age, presence of an older sibling (yes/no), maternal parenting score at 18 months, daycare attendance at 15 months (yes/no), and suspected hearing problem at 6 months (yes/no)

VI. AIM 1.2 MANUSCRIPT

Introduction

Soy-based infant formula (SBF) is a commonly used alternative to cow's milk based infant formula, particularly in instances of milk intolerance or preference for a vegan diet. It accounts for approximately 20% of the infant formula sold in the United States, and 7% in the United Kingdom (1, 2). While it is thought to be nutritionally adequate for term infants (2), SBF also contains high levels of phytoestrogens, plant compounds with structural and functional similarity to the steroid hormone,17β-estradiol. These phytoestrogen compounds, specifically the isoflavones genistein and daidzein, can bind to estrogen receptors (ER) and can act as either estrogen agonists or antagonists (1, 3-5). Since steroid hormones play an important role in sexually dimorphic brain and reproductive development during the pre- and early postnatal periods (6, 7), it is important to explore whether early life exposure to these hormonally active soy-derived compounds affect hormonally driven developmental characteristics.

Animal models have demonstrated that sexually dimorphic behaviors are sensitive to early life hormonal exposures, including exposure to the isoflavones in soy. For instance, visual-spatial memory, a characteristic on which male rats test better, is enhanced in females and decreased in males exposed to a phytoestrogen rich diet (8, 9). Low doses of genistein decreased offensive behaviors in male rats (10), while neonatal genistein exposures decreased reproductive posturing behaviors among female rats (a "defeminized" effect consistent with genistein functioning as a weaker agonist than the estradiol it displaced from

receptors) (11). However, no studies have assessed the effects of soy or its constituents on sexually dimorphic behaviors in children.

Since breast milk and formula are the only sources of nutrition for infants in the first months of life, an infant diet of soy-based formula can result in extremely high exposure to isoflavones. In fact, soy formula fed infants experience isoflavone exposures that are 500 times that of those fed cow's milk formula (12). Exposure of this magnitude may be of concern since brain development in early infancy may be particularly susceptible to endocrine disrupting effects (13-15) .

In this study, we assessed gender-role play behavior in boys and girls exposed to soy products in early infancy. Gender-role play behaviors are characterized by a child's preference for certain masculine- or feminine-typical toys, activities, and attitudes. Gender dimorphism in play is detectible at 12 months of age, and becomes pronounced by 36 months (16-19). These play behaviors are influenced largely by the social environment. However, the biological influence of early life hormone concentrations on play behavior has been clearly documented. Girls with high levels of androgens in the prenatal and very early postnatal periods due to congenital adrenal hyperplasia have consistently shown preference for male-typical toys and interests in multiple studies (20-22). Subsequent studies of environmental endocrine disruptors have yielded less clear results. Boys have exhibited less masculine play behavior following exposure to polychlorinated biphenyl, dioxins, and phthalates; girls have exhibited play behaviors that were more masculine, more feminine and not associated with the same exposures, respectively (23-25). Here, we assess the association between gender-role play behavior and early life soy exposure to further explore the influence of endocrine disruption on sexually dimorphic characteristics during development.

Methods

Study Sample. Women who were pregnant, residing in the Avon region of the United Kingdom, and expected to deliver between April 1, 1991 and December 31, 1992 were eligible for the Avon Longitudinal Study of Parents and Children (ALSPAC). Women were informed of the study by community clinicians or local media campaigns, and were recruited into the study after expressing interest in participation. 14,062 pregnancies were recruited into the study that resulted in live births. Of these, 13,978 (7,220 boys and 6,756 girls) were twins or singletons alive at one year. The present investigation was restricted to term singletons (n = 12,931) for whom complete infant feeding data were available (n = 8,492) and for whom a play behavior outcome assessment was completed at approximately 42 months of age, yielding a total study sample of 7,003 subjects (3,627 boys and 3,376 girls). Mothers provided consent for participation. Ethical approval for the study was obtained from the ALSAPC Law and Ethics Committee and the Local Research Ethics Committees. The present analysis was approved by the Institutional Review Board of the University of North Carolina at Chapel Hill.

Exposure Assessment. Mothers completed infant feeding questionnaires at 1, 6, 15 and 24 months postpartum. At 1 month, mothers reported all feeding methods used since birth (breast or bottle), and the type of formula used, if any. At 6, 15, and 24 months, mothers reported current breastfeeding habits, the age at which other milks or formulas were introduced into the child's diet (including formula/baby milk, , soy milk, soy formula, goat's milk, hypo-allergenic formula, and cow's milk), and how many feedings per week were given for each of these products at the time of questionnaire completion.

Exposure classification was defined by responses to the questionnaire administered at 6 months postpartum; if these data were missing or incomplete, responses from the 15-month questionnaire were used. "Early" exposure to any type of formula was defined as the use of a specific formula or milk type occurring ≤ 4 months of age ("*At what age did you start [formula/milk type]?")* through ≥ 6 months of age. Use at 6 months was indicated by any non-zero response to the question, "*How often nowadays is your baby fed [formula/milk type]?"* in the 6 month questionnaire. If the 15 month questionnaire was used instead, "early" exposure to formula was established for any subject that reported introducing the formula or milk ≤ 4 months of age and responded affirmatively to the question "*Since your child was 6 months old, has he/she had [formula/milk type]?"* This definition not only establishes early use of formula, but also a 1 month minimum duration of use.

Subjects were categorized into four mutually exclusive feeding groups: primarily breastfed, early formula, early soy, and late soy (Figure 6.1). *Primarily breastfed infants* were those who were breast fed until ≥ 6 months of age, who had no reported introduction of other milks or formulas before 6 months of age and no reported soy milk/formula use before 24 months of age. *Early formula fed infants* were introduced to any non-soy milk or formula product at or before 4 months of age, sustained use of such products at 6 months of age, and reported no soy use before 24 months of age. *Early soy fed infants* were introduced to soy milk or soy formula at or before 4 months of age, and sustained use at 6 months of age. *Late soy fed* infants were introduced to soy milk or soy formula any time after 4 months of age through 15 months of age. No restrictions were made in the early formula, early soy and late soy groups with respect to duration of breast feeding; likewise there were no restrictions in the early soy or late soy groups with respect to use of non-soy formula.

Subjects were excluded if feeding profiles were not sufficiently complete to estimate duration of a particular feeding method. Subjects who only reported soy use between 15 and 24 months were also excluded because it was assumed that exposure would be low compared to the earlier time periods when milk or formula comprised most of the diet. Responses from the 1 month questionnaire were used to verify that no soy was used in early infancy among primarily breast fed, early formula and late soy subjects. Exposure definitions do not take into account exposure to solid foods or their corresponding soy content, if any.

Outcome assessment. Gender-role play behavior was assessed at approximately 42 months of age using the Preschool Activities Inventory (PSAI), a psychometric test designed to assess within and between gender differences in early life play (26, 27). To complete the PSAI, mothers or other primary caregivers reported how often her child had played with certain toys (7 items), engaged in certain activities (11 items), and displayed certain characteristics (6 items) for the past month. Half of these items were "masculine," and half were "feminine." Each response was scored on a 5-point scale ("never," "hardly ever," "sometimes," "often," and "very often"). The total instrument was scored by summing responses to all masculine items, subtracting the sum of all feminine items, and applying a transformation (48.25 + 1.1*[Score]) to achieve a "pseudo-T score." Higher scores indicated masculine-typical behavior, and lower scores indicated feminine-typical behavior (27).

Covariates: Demographic factors, family composition, and lifestyle factors were assessed through parent report on various self-completed questionnaires. Given the potential importance of parental influences on both feeding practices and children's development, the mother's and partner's interaction with the child was estimated when the child was42 months of age using a series of questions assessing the frequency at which each parent participated in

a list of 8 activities with the child (score range: 0-36). Partner interaction scores were set to zero if the questionnaire reported that no partner was present. All partners were assumed to be male, given a very low prevalence of mothers in same-sex partnerships in this cohort (<1%) (28).

Analysis. Crude mean PSAI scores were assessed as simple means within exposure groups and within strata of each covariate. Differences in means were assessed using linear regression. Adjusted mean difference in PSAI scores were estimated using multivariable linear regression. Boys and girls were modeled separately to distinguish within gender differences. Unless otherwise noted, the early formula feeding group was used at the referent for all feeding group comparisons.

We assessed confounding by variables thought to be associated with both cultural and environmental influences of gender role behavior, as well as feeding method in the literature and in univariate investigations of these data. Final models were adjusted for age at PSAI assessment, breast feeding duration (continuous months), presence of an older brother (yes/no), presence of an older sister (yes/no), regular attendance in daycare (yes/no), maternal and partner interaction scores, and other maternal factors including age at delivery, smoking in the third trimester of pregnancy (yes/no), and education (5 levels in the United Kingdom, ranked from high to low: University Degree, Advanced Level, Ordinary Level, Vocational, and Certification of Secondary Education (CSE)/None).

Data were analyzed using both complete case analysis (CCA) and multiple imputation (MI) approaches. For the CCA, subjects with missing data for adjustment variables were dropped from models, so only true values for subjects with complete data were modeled. Approximately 17% of the study sample was excluded due to missing covariate data.

Multiple imputation was implemented to allow the entire study sample to be included in analyses. Values for missing covariates were estimated using PROC MI. Regression models of imputed data were run and summarized in PROC MIANALYZE. All analyses were completed using SAS 9.1.3 (SAS Institute Inc., Cary, NC).

Results

The subjects included in this analysis were similar to other term, singleton ALSPAC births with respect to the distribution of PSAI scores (among non-missing), gender distribution (not shown), feeding method distribution, and other select covariates (Table 6.1). The study sample did have slightly lower proportions of subjects with older brothers (32.1% vs. 33.6%), older sisters (30.9% vs. 31.8%), and prenatal smokers (16.9% vs. 19.5%), and slightly higher mean maternal age (28.7 years vs. 28.0 years), mean partner interaction scores (20.3 vs. 19.8), and proportions of mothers with a university degree (14.9% vs. 13.8%), than the larger ALSPAC cohort (not accounting for distribution of missing data within each covariate). Participating boys and girls were demographically similar, although girls were slightly more likely to have an older brother or sister (Table 6.1). Approximately 35% of the study sample attended day care regularly. Most mothers were non-smokers during the prenatal period and had mid- to high-level education. Approximately 6% of households reported the absence of a partner in the home.

Boys' and girls' PSAI scores had distinct normal distributions (Figure 2; mean (SD): boys, 62.3(8.6), range 20.8-95.6; girls, 36.9(9.3), range 4.3-85.7). Boys' scores were slightly higher and girls' scores were slightly lower in this study sample compared to the scores previously documented by the instrument's developers (mean (SD), boys: 61.7 (9.4); girls

38.7 (9.7)) (26). The PSAI assessment was completed between 41-53 months for boys, and 41-54 months for girls. Early soy fed infants accounted for 2.5% of boys and 2.0% of girls.

Among girls, early soy exposure was associated with a higher crude PSAI score, as compared to early formula feeding (Table 6.2). The other feeding regimes were not associated with PSAI scores. Among boys, lower (more feminine) scores were observed among the primarily breast fed as compared to early formula fed boys. No substantial difference was observed between early soy and early formula fed boys. Scores for both genders were higher in the presence of an older brother and prenatal tobacco smoke exposure, and decreased in the presence of an older sister. Women with more education reported lower scores for their boys and higher scores for their girls. Scores were lower for boys and higher for girls with higher levels of maternal education. Boys' scores also lowered with increasing duration of breast feeding, while girls' scores elevated with increasing maternal age and lowered with increasing age at PSAI assessment. Daycare attendance and parental interaction scores were not associated with any difference in PSAI score for either gender.

In the adjusted complete case analysis, early soy feeding was also associated with an elevation in PSAI score among girls ($\beta = 2.68$; 95% CI: 0.20, 5.15) compared to early formula feeding, but no association was observed in boys for any feeding group (Table 6.3). Estimates for early soy exposure derived using multiple imputation were similar, but more precise ($\beta_{girls} = 2.87$; 95% CI: 0.67, 5.06;$\beta_{boys} = 0.96$;95% CI: -0.82, 2.74). No association was observed in the late soy or primarily breast fed exposure groups.

Several control variables were also associated with change in PSAI score in the adjusted model. For both genders, adjusted mean change in PSAI scores were associated with the

presence of an older brother, presence of an older sister, prenatal smoking, maternal education, and maternal age. Breast feeding duration was associated with lower scores among males, while age at assessment was associated with lower scores among females.

Discussion

Early life soy exposure was associated with a slightly higher (more "masculine") PSAI score among girls, while no association between soy and PSAI score was observed in boys. The association among girls was robust to adjustment for multiple social and environmental factors. Although the average PSAI score was nearly 3 points higher among early soy fed girls, they remained within the normal range for feminine behavior. As a point of comparison, the association between early soy exposed girls and gender-role behavior is similar to, but slightly less than, that observed among girls with older brothers in this study, though much less precise. Our findings suggests that early soy exposure may be associated with girls exhibiting slightly *less feminine* behavior than their non-soy fed counterparts, but not overtly *masculine* or "boy-like" behavior.

A subtle de-feminizing effect on behavior in females is consistent with previous investigations of early life exposure to endocrine disrupting compounds. The binding of estrogen receptors to a weak, exogenous estrogen such as genistein may induce an antagonistic effect on normal, highly regulated estradiol-dependant mechanisms, resulting in impaired development of normal female characteristics (3). Animal models have suggested such effects occur structurally in sexually dimorphic regions of the female brain following exposures to testosterone, excess estrogen, or exogenous estrogen-like compounds (6). For example, the volume of the sexually dimorphic nucleus of the medial pre-optic area (SDN-POA) of the hypothalamus, which plays a role in male sexual behavior and is normally larger

in males than females, has been shown to increase in female rats following postnatal exposure to high levels of genistein (29-31). Neonatal genistein exposure has also been shown to affect estrus cycling and alter reproductive posturing behaviors in exposed female rats (11). These studies emphasize the early postnatal periods, which operates in agreement with our findings that early soy exposure, but not late soy exposure, had an effect on play behavior.

Epidemiologic studies of other hormonal and endocrine disrupting exposures also support our findings in girls. The association between of early life hormone concentrations and gender-role play behavior has been observed among girls with congenital adrenal hyperplasia, who clearly and consistently show preference for male-typical toys when compared to unexposed female relatives (20, 21). This hypothesis that early life endocrine disruption affects sexually dimorphic behavior in girls has been further supported in studies showing increased masculine typical behaviors in girls following prenatal exposures to polychlorinated biphenyls and bisphenol A (23, 32).

Most animal studies also suggest a feminizing effect among males following soy isoflavone exposure, which was not observed in this study. Our null findings may be explained, in part, by biased reporting of feminine behaviors among parents of boys. If, for instance, social stigma prevented parents from accurately reporting feminine behaviors among boys in a manner that was non-differential by exposure status, our results would then be biased towards the null. Alternatively, it is possible that the biologic effect of soy in boys is too subtle, if present at all, to be detected by a parent-report instrument such as the PSAI, or that the susceptible developmental period for these effects may be in the prenatal rather than postnatal period.

We used longitudinal exposure assessment data to characterize soy product use in early and late infancy. However, early and late soy users were not necessarily *exclusively* fed soy products at any time point and we were not able characterize the relative dose of soy; thus we were unable to assess a dose-response relationship between soy feeding and PSAI score. The exposure classification was also unable to account for other dietary or prenatal exposures to soy isoflavones. Improved characterization of early life soy exposure, either through more detailed questionnaire or isoflavone biomarker measures, should be implemented in future studies.

While the ALSPAC cohort is a large study that is widely generalizable to the United Kingdom, the number of soy users in this population is quite small, resulting in imprecise estimates of association. This study characteristic should be considered when determining whether to generalize these findings to other regions where soy use is more prevalent, such as the United States.

Our results suggest that early life exposure to soy products may subtly increase masculine type play behaviors in girls, but are not associated with gender-role play behaviors in boys. Given the low prevalence of soy use in this study sample, associations between soy exposure and PSAI score were imprecise and results should be interpreted cautiously. If replicated, these findings may be more important in populations with higher prevalence of soy use than observed here. Accordingly, additional studies should explore the relationship between early life soy exposure and sexually dimorphic development, particularly in populations with high soy use prevalence. Additional attention should also be paid to the effects that soy isoflavones may be having in the prenatal period, so as to better characterize effects across multiple developmental stages.

REFERENCES

1. Essex C. Phytoestrogens and soy based infant formula. BMJ. 1996 Aug 31;313(7056):507-8.

2. Bhatia J, Greer F, American Academy of Pediatrics Committee on Nutrition. Use of soy protein-based formulas in infant feeding. Pediatrics. 2008 May;121(5):1062-8.

3. Hwang CS, Kwak HS, Lim HJ, Lee SH, Kang YS, Choe TB, et al. Isoflavone metabolites and their in vitro dual functions: They can act as an estrogenic agonist or antagonist depending on the estrogen concentration. J Steroid Biochem Mol Biol. 2006 Nov;101(4-5):246-53.

4. Dixon RA, Ferreira D. Genistein. Phytochemistry. 2002 Jun;60(3):205-11.

5. Patisaul HB. Phytoestrogen action in the adult and developing brain. J Neuroendocrinol. 2005 Jan;17(1):57-64.

6. Scallet AC, Divine RL, Newbold RR, Delclos KB. Increased volume of the calbindin D28k-labeled sexually dimorphic hypothalamus in genistein and nonylphenol-treated male rats. Toxicol Sci. 2004 Dec;82(2):570-6.

7. Davis EC, Popper P, Gorski RA. The role of apoptosis in sexual differentiation of the rat sexually dimorphic nucleus of the preoptic area. Brain Res. 1996;734(1-2):10-8.

8. Lephart ED, West TW, Weber KS, Rhees RW, Setchell KD, Adlercreutz H, et al. Neurobehavioral effects of dietary soy phytoestrogens. Neurotoxicol Teratol. 2002 Jan-Feb;24(1):5-16.

9. Lephart ED, Setchell KD, Handa RJ, Lund TD. Behavioral effects of endocrine-disrupting substances: Phytoestrogens. ILAR J. 2004;45(4):443-54.

10. Wisniewski AB, Klein SL, Lakshmanan Y, Gearhart JP. Exposure to genistein during gestation and lactation demasculinizes the reproductive system in rats. J Urol. 2003 Apr;169(4):1582-6.

11. Kouki T, Kishitake M, Okamoto M, Oosuka I, Takebe M, Yamanouchi K. Effects of neonatal treatment with phytoestrogens, genistein and daidzein, on sex difference in female rat brain function: Estrous cycle and lordosis. Horm Behav. 2003 Aug;44(2):140-5.

12. Cao Y, Calafat AM, Doerge DR, Umbach DM, Bernbaum JC, Twaddle NC, et al. Isoflavones in urine, saliva, and blood of infants: Data from a pilot study on the estrogenic activity of soy formula. J Expo Sci Environ Epidemiol. 2009 Feb;19(2):223-34.

13. Cohen-Bendahan CC, van de Beek C, Berenbaum SA. Prenatal sex hormone effects on child and adult sex-typed behavior: Methods and findings. Neurosci Biobehav Rev. 2005 Apr;29(2):353-84.

14. Chellakooty M, Schmidt IM, Haavisto AM, Boisen KA, Damgaard IN, Mau C, et al. Inhibin A, inhibin B, follicle-stimulating hormone, luteinizing hormone, estradiol, and sex hormone-binding globulin levels in 473 healthy infant girls. J Clin Endocrinol Metab. 2003 Aug;88(8):3515-20.

15. Sehested A, Juul AA, Andersson AM, Petersen JH, Jensen TK, Muller J, et al. Serum inhibin A and inhibin B in healthy prepubertal, pubertal, and adolescent girls and adult women: Relation to age, stage of puberty, menstrual cycle, follicle-stimulating hormone, luteinizing hormone, and estradiol levels. J Clin Endocrinol Metab. 2000 Apr;85(4):1634-40.

16. Golombok S, Rust J, Zervoulis K, Croudace T, Golding J, Hines M. Developmental trajectories of sex-typed behavior in boys and girls: A longitudinal general population study of children aged 2.5-8 years. Child Dev. 2008 Sep-Oct;79(5):1583-93.

17. Snow ME, Jacklin CN, Maccoby EE. Sex-of-child differences in father-child interaction at one year of age. Child Dev. 1983;54(1):227-32.

18. O'Brien M, Huston AC. Activity level and sex-sterotyped toy choice in toddler boys and girls. J Genet Psychol. 1985;146(4):527-33.

19. Servin A, Bohlin G, Berlin L. Sex differences in 1-, 3-, and 5-year-olds' toy-choice in a structured play-session. Scand J Psychol. 1999 Mar;40(1):43-8.

20. Berenbaum SA, Duck SC, Bryk K. Behavioral effects of prenatal versus postnatal androgen excess in children with 21-hydroxylase-deficient congenital adrenal hyperplasia. J Clin Endocrinol Metab. 2000 Feb;85(2):727-33.

21. Berenbaum SA, Hines M. Early androgens are related to childhood sex-typed toy preferences. Psychol Sci. 1992;3(3):203-6.

22. Iijima M, Arisaka O, Minamoto F, Arai Y. Sex differences in children's free drawings: A study on girls with congenital adrenal hyperplasia. Horm Behav. 2001 Sep;40(2):99-104.

23. Vreugdenhil HJ, Slijper FM, Mulder PG, Weisglas-Kuperus N. Effects of perinatal exposure to PCBs and dioxins on play behavior in dutch children at school age. Environ Health Perspect. 2002 Oct;110(10):A593-8.

24. Swan SH, Liu F, Hines M, Kruse RL, Wang C, Redmon JB, et al. Prenatal phthalate exposure and reduced masculine play in boys. Int J Androl. 2010 Apr;33(2):259-69.

25. Hines M, Johnston KJ, Golombok S, Rust J, Stevens M, Golding J, et al. Prenatal stress and gender role behavior in girls and boys: A longitudinal, population study. Horm Behav. 2002 Sep;42(2):126-34.

26. Golombok S, Rust J. The pre-school activities inventory: A standardized assessment of gender role in children. Psychological Assessment. 1993;5(2):131-6.

27. Golombok S, Rust J. The measurement of gender role behaviour in pre-school children: A research note. J Child Psychol Psychiatry. 1993 Jul;34(5):805-11.

28. Golombok S, Perry B, Burston A, Murray C, Mooney-Somers J, Stevens M, et al. Children with lesbian parents: A community study. Dev Psychol. 2003;39(1):20-33.

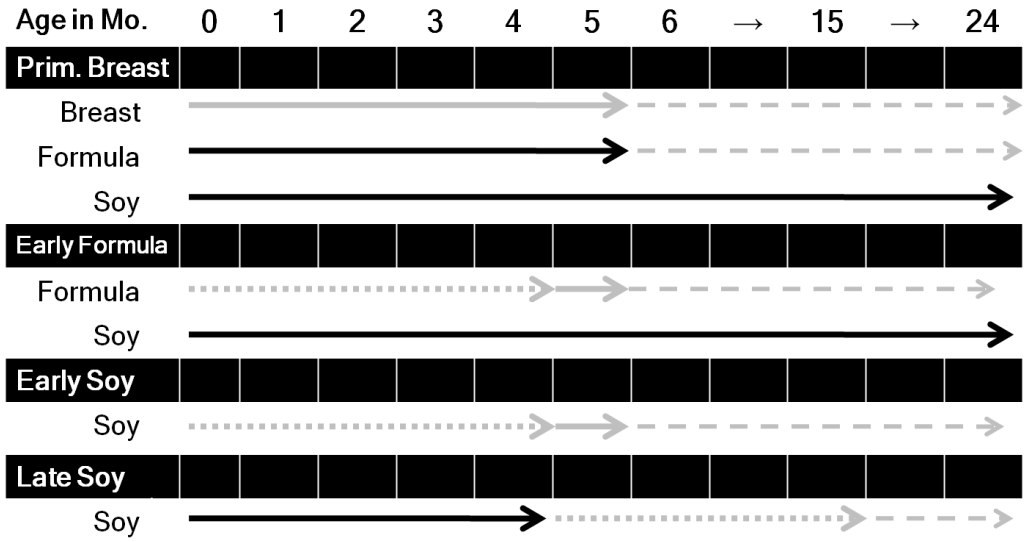
29. Faber KA, Hughes CL,Jr. Dose-response characteristics of neonatal exposure to genistein on pituitary responsiveness to gonadotropin releasing hormone and volume of the sexually dimorphic nucleus of the preoptic area (SDN-POA) in postpubertal castrated female rats. Reprod Toxicol. 1993;7(1):35-9.

30. Faber KA, Hughes CL,Jr. The effect of neonatal exposure to diethylstilbestrol, genistein, and zearalenone on pituitary responsiveness and sexually dimorphic nucleus volume in the castrated adult rat. Biol Reprod. 1991 Oct;45(4):649-53.

31. Lewis RW, Brooks N, Milburn GM, Soames A, Stone S, Hall M, et al. The effects of the phytoestrogen genistein on the postnatal development of the rat. Toxicol Sci. 2003 Jan;71(1):74-83.

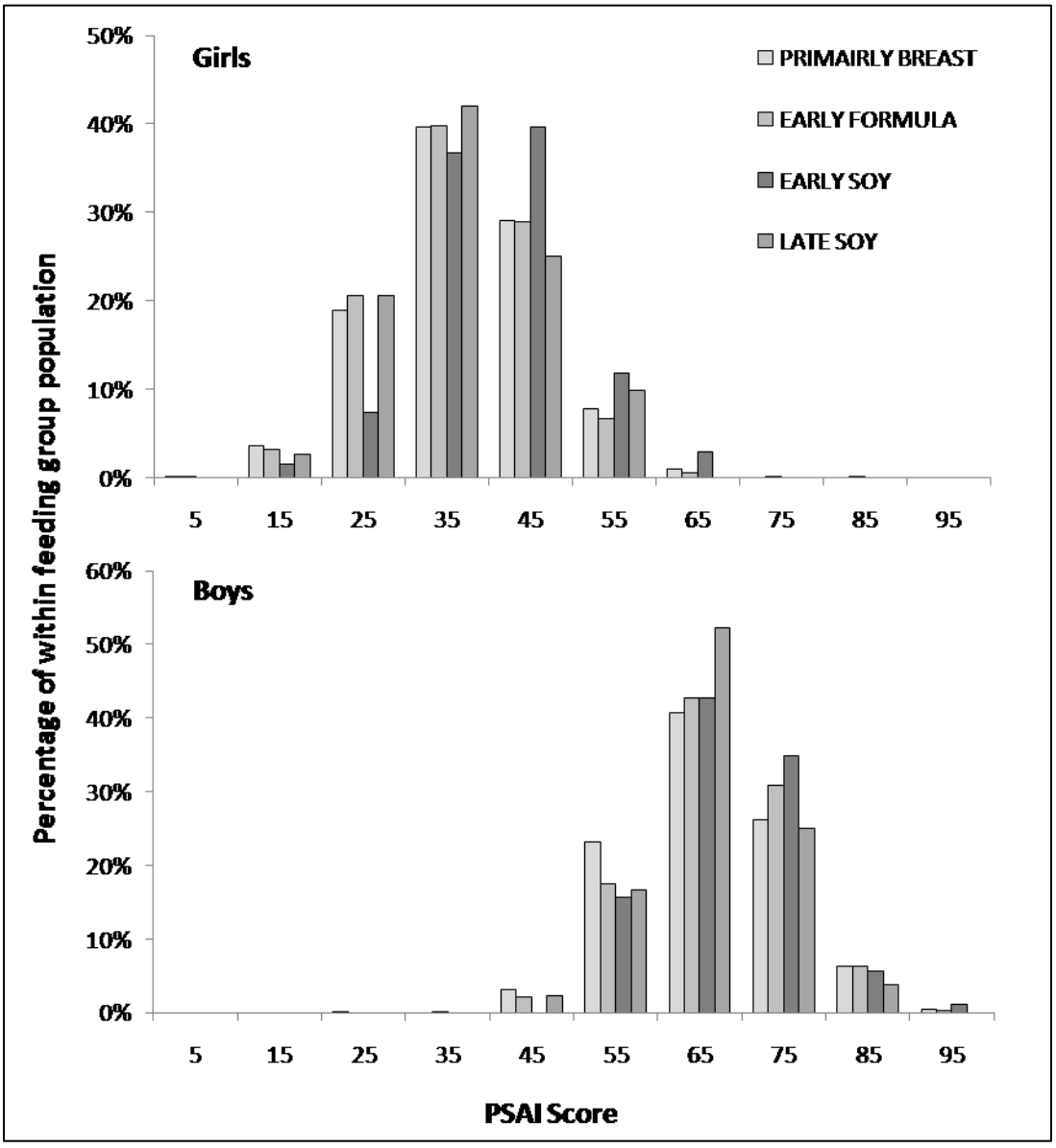
32. Braun JM, Yolton K, Dietrich KN, Hornung R, Ye X, Calafat AM, et al. Prenatal bisphenol A exposure and early childhood behavior. Environ Health Perspect. 2009 Dec;117(12):1945-52.

Figure 6.1. Exposure Characterization for Infant Feeding Groups



Exposure Characterization for Infant Feeding Groups. Solid arrows (→) indicate the *required* time period (age in months) of use for each particular feeding method. Short dotted arrows (⋯▸) indicate the time period (age in months) during which "early use" of a product (either soy-based or traditional milk or formula) could have been introduced. Solid bars(•—•) indicate time periods for which use of a particular milk or formula product was prohibited for a particular feeding group. Long dotted arrows (- - ▸) indicate periods of time for which use of a particular feeding method was not restricted or defined according to the exposure definition.

Figure 6.2. Distribution of PSAI scores, by gender and feeding group



Distribution of PSAI Scores By Gender and Feeding Group. Histograms representing the distribution of PSAI scores for girls (top) and boys (bottom). Colored bars (see key) correspond to each of the four infant feeding group categories.

Table 6.1. Characteristics of study sample (n = 7,003) and eligible subjects (n = 12,931)

	Study Sample			
Characteristic	**Boys**	**Girls**	**Total**	**Eligible[a]**
PSAI Score, mean (SD)	62.3 (8.6)	36.9 (9.3)	50.1 (15.5)	50.0 (15.6)
Missing, n	--	--	--	3567
Age at PSAI completion, mean (SD)	42.3 (0.8)	42.3 (0.8)	42.3 (0.8)	42.3 (0.9)
Missing, n	81	85	166	4872
Breast Feeding Duration, mean (SD)	3.9 (4.5)	4.1 (4.6)	4.0 (4.5)	3.8 (4.5)
Missing, n	0	0	0	1011
Maternal Age, mean (SD)	28.8 (4.7)	28.5 (4.6)	28.7 (4.7)	28.0 (5.0)
Missing, n	0	0	0	0
Mother Interaction Score, mean (SD)	28.5 (4.9)	28.8 (4.7)	28.6 (4.8)	28.6 (4.8)
Missing, n	12	6	18	3597
Partner Interaction Score, mean (SD)	20.4 (7.9)	20.1(8.1)	20.3 (8.0)	19.8 (8.6)
Missing, n	19	25	44	3567
Infant Feeding Method, n (%)				
Early Formula	2699 (74.4)	2490 (73.8)	5189 (74.1)	6294(74.7)
Early Soy	89 (2.5)	68 (2.0)	157 (2.2)	182 (2.2)
Late Soy	132 (3.6)	112 (3.3)	244 (3.5)	286 (3.4)
Primarily Breastfed	707 (19.5)	706 (20.9)	1413 (20.2)	1667(19.8)
Missing, n	--	--	--	4502
Presence of Older Brother, n (%)				
No	2324 (66.7)	2235 (69.1)	4559 (67.9)	6816(66.4)
Yes (≥ 1)	1160 (32.0)	999 (29.6)	2159 (32.1)	3443(33.6)
Missing, n	143	142	285	2672
Presence of Older Sister, n (%)				
No	2370(68.1)	2268(70.1)	4638(69.1)	6991(68.2)
Yes (≥ 1)	1112 (31.9)	965 (29.9)	2077 (30.9)	3264(31.8)
Missing, n	145	143	288	2676
Regular Daycare Attendance, n (%)				
No	2212 (64.5)	1998 (62.4)	4210 (63.5)	5943(63.0)
Yes	1220 (35.6)	1205 (37.6)	2425 (36.6)	3490(37.0)
Missing, n	195	173	368	3498
Prenatal Smoking, n (%)				
No	2895(82.7)	2734 (83.6)	5629 (83.1)	8898(80.5)
Yes	604 (17.3)	538 (16.4)	1142 (16.9)	2162(19.5)
Missing, n	128	104	232	1871
Maternal Education, n (%)				
University Degree	489 (14.4)	488 (15.4)	977 (14.9)	1407(13.8)
Advanced Level	869 (25.4)	814 (25.8)	1674 (25.6)	2607(24.1)
Ordinary Level	1306 (38.5)	1186 (37.5)	2492 (38.0)	4002(36.9)
Vocational	335 (9.9)	297 (9.4)	632 (9.64)	1132(10.5)
CSE[b]/None	403 (11.9)	375 (11.9)	778 (11.9)	1596(14.7)
Missing, n	234	216	450	2097

[a] Eligible ALSPAC subjects are limited to term, singleton infants alive at 1 year
[b]CSE: Certificate of Secondary Education

Table 6.2. Crude mean Preschool Activities Inventory (PSAI) (mean (SD)) scores and regression estimates (β (SE)) for exposure groups and select categorical and continuous covariates

Covariate	**Boys**	**Girls**
Infant Feeding Method, Mean (SD)		
Early Formula[a]	62.6 (8.5)	36.7 (9.2)
Early Soy	63.0 (8.1)	40.8 (9.1)*
Late Soy	61.7 (7.8)	36.9 (9.8)
Primarily Breastfed	61.3 (9.1)*	37.1 (9.4)
Presence of Older Brother, Mean (SD)		
No[a]	61.5 (8.5)	35.9 (9.0)
Yes (≥ 1)	63.7 (8.7)*	39.3 (9.5)*
Missing	62.8 (8.4)	36.3 (9.3)
Presence of Older Sister, Mean (SD)		
No[a]	63.2 (8.5)	37.6 (9.4)
Yes (≥ 1)	60.2 (8.6)*	35.5 (8.9)*
Missing	62.7 (8.4)	36.3 (9.3)
Regular Daycare Attendance, Mean (SD)		
No[a]	62.2 (8.6)	36.7 (9.2)
Yes	62.4 (8.5)	37.3 (9.5)
Missing	62.7 (9.3)	37.2 (9.3)
Prenatal Smoking, Mean (SD)		
No[a]	62.0 (8.5)	36.7 (9.2)
Yes	63.7 (8.8)*	38.2 (9.8)*
Missing	62.1 (9.0)	36.2 (8.7)
Maternal Education, Mean (SD)		
University Degree[a]	60.6 (8.5)	38.6 (9.3)
Advanced Level	61.9 (9.0)*	37.4 (9.4)*
Ordinary Level	62.7 (8.3)*	36.1 (9.1)*
Vocational	62.4 (8.1)*	36.0 (9.0)*
CSE[b]/None	63.2 (8.7)*	36.9 (9.7)*
Missing	63.2 (9.0)*	36.7 (9.5)*
Age at PSAI Assessment, β (SE)	0.25 (0.18)	-0.44 (0.19)*
Breast Feeding Duration, β (SE)	-0.16 (0.03)*	0.05 (0.04)
Maternal Age at Delivery, β (SE)	-0.01 (0.03)	0.17 (0.03)*
Maternal Interaction Score, β (SE)	0.02 (0.03)	0.02 (0.03)
Partner Interaction Score, β (SE)	0.00 (0.02)	0.02 (0.02)

[a] Referent Category

[b] CSE: Certificate of Secondary Education

*p <0.05 for mean difference estimates, compared to referent category

Table 6.3. Adjusted [a] change in mean Preschool Activities Inventory (PSAI) scores for boys and girls

	β (95% CI)	
Covariate	**Boys (n = 2979)**	**Girls (n = 2788)**
Infant Feeding Method		
Early Formula	0.	0.
Early Soy	1.24 (-0.67, 3.14)	2.68 (0.20, 5.15)
Late Soy	-0.40 (-2.05, 1.24)	-0.56 (-2.59, 1.47)
Primarily Breastfed	0.29 (-0.82, 1.39)	0.00 (-1.22, 1.23)
Presence of Older Brother		
No	0.	0.
Yes (≥ 1)	2.06 (1.39, 2.73)	3.26 (2.50, 4.03)
Presence of Older Sister		
No	0.	0.
Yes (≥ 1)	-2.76 (-3.44, -2.09)	-1.88 (-2.65, -1.11)
Regular Daycare Attendance		
No	0.	0.
Yes	0.37 (-0.26, 1.01)	0.46 (-0.25, 1.16)
Prenatal Smoking		
No	0.	0.
Yes	1.35 (0.50, 2.21)	2.07 (1.10, 3.04)
Maternal Education		
University Degree	0.	0.
Advanced Level	1.04 (0.04, 2.03)	-1.32 (-2.40, -0.25)
Ordinary Level	1.68 (0.69, 2.66)	-2.64 (-3.73, -1.55)
Vocational	1.34 (0.02, 2.65)	-2.75 (-4.21, -1.28)
CSE[b]/None	2.19 (0.91, 3.48)	-2.29 (-3.72, -0.85)
Age at PSAI Assessment	0.15 (-0.25, 0.55)	-0.61 (-1.07, -0.15)
Breast Feeding Duration	-0.14 (-0.24, -0.03)	-0.03 (-0.14, 0.09)
Maternal Age	0.08 (0.01, 0.15)	0.10 (0.01, 0.18)
Maternal Interaction Score	0.03 (-0.03, 0.05)	0.02 (-0.06, 0.09)
Partner Interaction Score	0.01 (-0.03, 0.05)	0.02 (-0.02, 0.07)

[a] Adjustment variables include age at PSAI assessment, breast feeding duration (months), presence of an older brother (yes/no), presence of an older sister (yes/no), regular attendance in daycare (yes/no), partner interaction score, and maternal factors including age, prenatal smoking status (yes/no), education, and interaction score

[b] CSE: Certificate of Secondary Education

VII. AIM 2 MANUSCRIPT

Introduction

The soy isoflavones, genistein and daidzein, are weak estrogenic compounds contained in soy protein and various products derived from soybeans (1). Demonstrating structural and functional similarity to 17β-estradiol, soy isoflavones can bind to estrogen receptors and can act as either estrogen agonists or antagonists (1-4). The biological activity of soy isoflavones has been demonstrated widely *in vitro* and in animal models(5-9), as well as in adult humans(10, 11). Few studies, however, have addressed the effects of early life soy protein exposure on long-term outcomes such as reproductive development and function.

Endogenous sex hormones play an important role in brain and reproductive development in the pre- and neonatal periods (12-14). Their biologic activity is also important in the timing of pubertal onset and in reproductive function during adolescence (15). Consequently, exposure to soy isoflavones in early infancy may have lasting effects on later reproductive development and function. Animal studies have shown that pubertal markers, such as the age at vaginal opening, occur earlier in female rodents fed genistein during various time points in early development (16-19). Early pubertal onset has also been observed in girls exposed to other endocrine disrupting compounds pre- and postnatally (20). Infants exposed to soy-based products, such as soy-based infant formula (SBF), have not been prospectively followed to evaluate the potential for early life soy exposure to disrupt endocrine system activity. SBF is commonly used in the United States, accounting for approximately 20% of the infant formula sold (21). Isoflavone content in SBF is

approximately 4 orders of magnitude higher than other common sources of infant nutrition (isoflavone range in SBF: 32 to 47 mg/L; breast milk: ~ 5.6 μg/L; cow's milk: 0.1-2.0 μg/L).(22-24) The urinary concentration of total isoflavones among infants exclusively fed SBF is approximately 500 times the concentration of those fed cow's milk formula (25), and plasma isoflavone concentrations per bodyweight were an order of magnitude higher in SBF fed infants than in adults consuming diets containing soy protein (26). No other postnatal exposure to an endocrine disruptor approaches this exposure level (27). We have investigated the association between soy product use during early infancy and age at menarche among girls enrolled in the Avon Longitudinal Study of Parents and Children (ALSPAC). In this study, age at menarche may serve as an easily observable marker of possible early life endocrine disruption.

Methods

Study Sample. ALSPAC is an ongoing, prospective, longitudinal study which enrolled pregnant women residing in the Avon region of the United Kingdom, who were expected to deliver between April 1, 1991 and December 31, 1992. 14,062 live births were recruited into the study during pregnancy. Of these, 13,978 (7,220 boys and 6,756 girls) were twins or singletons alive at one year. Mothers provided consent for participation.

The present investigation was restricted to singleton, term, white females who were alive at one year of age (n = 5,230), then further restricted to those for whom comprehensive infant feeding data was available and for whom at least one puberty questionnaire was completed between the ages of approximately 8 and 14 years of age (n = 2,884). We restricted the analysis to white race was because over 90% of the study sample were white, which limited the power to control potential confounding by race. Ethical approval for the study was

obtained from the ALSAPC Law and Ethics Committee and the Local Research Ethics Committees and the present analysis was approved by the Institutional Review Board of the University of North Carolina at Chapel Hill.

Exposure Assessment. Mothers completed infant feeding questionnaires at 1, 6, 15 and 24 months postpartum. At 1 month, mothers reported all feeding methods used since birth. At 6, 15, and 24 months, mothers reported current breastfeeding habits, the age at which other milks or formulas were introduced into the child's diet (including formula/baby milk, soy milk, soy formula, goat's milk, hypo-allergenic formula, and cow's milk), and how many feedings per week were given for each product at the time of questionnaire completion.

Exposure classification was defined by responses to the questionnaire administered at 6 months postpartum; if these data were missing or incomplete, responses from the 15-month questionnaire were used. Subjects were categorized into four mutually exclusive feeding groups: primarily breastfed, early formula, early soy, and late soy (Figure 7.1). *Primarily breastfed infants* were breast fed until ≥ 6 months of age, had no reported soy use between birth and 24 months and no reported introduction of other milks or formulas before 6 months of age. *Early formula fed infants* were introduced to any non-soy milk or formula product at or before 4 months of age, sustained use of such products at 6 months of age, and reported no soy use before 24 months. *Early soy fed infants* were introduced to soy milk or soy formula at or before 4 months of age, and reported sustained use at 6 months of age. *Late soy fed infants* were introduced to soy milk or soy formula any time after 4 months of age through 15 months of age. No restrictions were made in the early formula, early soy, and late soy groups with respect to duration of breast feeding; likewise, there were no restrictions in the early soy

or late soy groups with respect to use of non-soy milk or formula. Exposure definitions do not take into account exposure to solid foods or their corresponding soy content, if any.

Outcome assessment. Between 1999 and 2007, a series of questionnaires regarding pubertal development (the "Growing and Changing" questionnaires), were administered at approximately 8, 9.5, 10.5, 11.5,13 and 14.5 years of age.(28) Questionnaires were completed by a care-giving adult or the child of interest. In each questionnaire, subjects were asked if the child had had her first period, and if so, what month and year her first period occurred. Her age at menarche was defined as her age in months at this time. The earliest reported age was used as the age at menarche in the event that multiple questionnaires contained discordant responses for the same individual.

ALSPAC also enrolled some subjects from the main study into a smaller clinic-based cohort study that also assessed age at menarche. For 146 (4.6%) subjects in the present analysis, missing questionnaire data on age at menarche was obtained from ALSPAC clinic data.

Some subjects reported a menarche event, but did not report an age (n = 107). For 67 of these subjects, age at menarche was estimated as the midpoint between the age at which the questionnaire with the first positive menarche response was completed, and the age at which the previous year's questionnaire was completed. If more than one questionnaire was skipped between a negative and positive menarche response, and estimated age was not derived and this subject was not included in any analysis (n = 40). As an alternative, an imputed age at menarche was estimated for these 67 subjects as part of a larger multiple imputation model for missing data (described below). Ages derived using the midpoint approach were included

in the complete case analysis, while imputed values were used in multiple imputation models.

Analysis. All analyses were completed using SAS 9.1.3 (SAS Institute Inc., Cary, NC). Hazard ratios (HR) for time-to-menarche were estimated using Cox proportional hazards modeling. The early formula group was used as the referent group in all models. Relative precision of estimated HRs were compared using confidence limit ratios (CLR), calculated as the upper 95% confidence limit divided by the lower 95% confidence limit. Proportional hazards assumptions (PHA) were assessed using log-log survival density function plots and Cox chi-squared significance tests for time interaction variables ($\alpha = 0.05$). These diagnostics revealed varying degrees of hazard convergence over time, including a notable convergence at approximately 150 months of age. Therefore, models were carried out using categorical ($< / \geq 150$ months (12.5 years)) and continuous-time interactions (exposure*time, pre-pregnancy BMI*time, prenatal smoking*time, and maternal age at menarche*time), to characterize changes in hazards over time. Hazard ratios were estimated at 10,11,12,13 and 14 years for continuous-time Cox models. Cox models were analyzed using both complete case analysis and multiple imputation (MI). For the complete case analysis, only subjects with complete data on necessary covariates were modeled, which resulted in a loss of 18% of subjects. Multiple imputation of missing outcome (n = 67), adjustment variables and other covariates (BMI) was performed using PROC MI (5 imputations), and simulated results were combined using PROC MIANALYZE.

Follow up time was defined for each subject based on the age in months at which she reported a menarche event (*events*), or the age of the last completed questionnaire in which she reported not having reached menarche (*censored).* Censored subjects (those that did not

report a menarche event) were distinguished as either *lost to follow up* (dropped out before the end of the study), or *administratively censored* (completed the study). Crude Kaplan Meier survival curves, median time-to-menarche and inter-quartile range (IQR: 25^{th}-75^{th} Percentile) estimates were obtained using lifetable analysis (PROC LIFETEST). We obtained confounding-adjusted Kaplan Meier curves and median time-to-menarche by exposure group by calculating inverse probability of exposure weights (29, 30) using polytomous logistic regression, and then performing a weighted lifetable analysis.

Sensitivity analyses were performed to assess whether results were affected by informative censoring (31). In the first analysis, we assumed that all *lost to follow up* subjects were at low risk for reaching menarche in the study period, and their follow-up times were reassigned to resemble the administratively censored (follow-up time: 175 months) ("low risk"). In the second analysis, we assumed that randomly censored subjects were at high risk for reaching menarche in the study period, and all *lost to follow up* follow-up times were modeled as events occurring at the time of drop out ("high risk"). A third sensitivity analysis was preformed to address early censoring times in the early soy group. Here, we manipulated the distribution of censor times in the early soy group to mimic the censor time distribution seen in other exposure groups (median: ~140 months) ("redistributed"). Categorical-time Cox models were repeated for each of the three hypothetical scenarios to evaluate how these assumptions affected estimates.

Covariates were included as potential confounders if they were associated with an infant feeding method, age at menarche, or censoring in these data (via univariate association) or in relevant literature. Variables examined included child's birth weight, breast feeding duration, milk allergy at 6 months, vegetarian diet in childhood, and maternal

perception of infant health, and maternal factors including age and education at delivery, prenatal vegetarian diet, age at menarche, pre-pregnancy body mass index (BMI), and pre- and postnatal smoking. The final models were adjusted for pre-pregnancy BMI, smoking in the last 2 months of pregnancy (yes/no), and maternal age at menarche (continuous years). Postnatal smoking was not included to avoid collinearity with the prenatal smoking.

Effect measure modification was examined in relation to childhood weight status by stratifying above and below the 85th percentile of BMI, using measurements taken at any time between 7 and 9.5 years. BMI stratified results are presented as time-averaged HRs, since PHA violations were minimal in the stratified samples, and sample size limitations prevented further stratification by time. Associations between feeding group and BMI age adjusted Z-scores were assessed using linear regression.(32)

Results

Among 5,230 eligible ALSPAC subjects, 2,884 had sufficient infant feeding and puberty data to be included in this analysis. Of the 2,346 excluded girls, most had insufficient feeding data to characterize exposure (n=1,480), or had soy exposure only after 15 months of age when milk/formula is a less prominent component of the diet (n=86). Other excluded girls did not complete any "Growing and Changing" questionnaire (n = 780). On average, excluded girls had lower birth weight, shorter breastfeeding duration, were more likely to be ill as infants, exposed to prenatal tobacco smoke, and born to younger mothers than those that remained in the final study sample (Table 7.1). Excluded girls were also more likely to have missing data on key covariates. There was no difference between the included and excluded subjects with respect to soy product use. However, excluded girls were less likely

to be classified in the primarily breast fed group, and more likely to be classified in the early formula group, as compared to the included study sample.

The median time-to-menarche was 153 months (12.8 years) [IQR, 144-163]. The median follow-up time contributed by those lost to follow up (n = 658 [22.8%]) was 140 months (11.7 years) [IQR, 118-157]. Three percent of girls were administratively censored. Two percent of girls were fed soy at or prior to 4 months of age (*early soy*). Over 20% of subjects in this group initiated soy use in the first month, while approximately 37% initiated soy use at age 4 months.

In crude bivariate analyses, there was no effect of birth weight, breast feeding duration, infant health, or maternal age at delivery on median time-to-menarche. Time-to-menarche was earlier among girls with high BMI, and among girls whose mothers had a high pre-pregnancy BMI or a young age at menarche. However, early censoring may have contributed to early time-to-menarche among girls with high BMI, and among those whose mothers had high pre-pregnancy BMI (respective median [IQR] follow-up times among lost to follow up subjects: 128 [115-141] and 132 [124 -141]). Girls with mothers who had an older age at menarche had a later age at menarche (Table 7.1).

Across infant feeding groups, the crude median time-to-menarche was earliest for girls receiving an early soy diet (149 months (12.4 years) [IQR, 140-159]), and latest among those who were primarily breast fed (154 months (12.8 years) [IQR, 145-165]). There was considerable overlap of the inter-quartile ranges for all comparison groups. Crude and adjusted Kaplan Meier curves (Figure 7.2) suggest a more rapid time-to-menarche in the early soy group, as compared to all other feeding groups, prior to approximately 150 months. However, these curves, along with adjusted median time-to-menarche estimates (Table 7.2)

further suggest that that this difference converges across all feeding groups at approximately 150 months of age (12.5 years).

Adjusted hazard ratios for the complete case analysis are presented in Table 7.2. The hazard of reaching menarche by age 12.5 (150 months) among girls in the early soy feeding group is approximately 1.42 times [95% CI, 0.92, 2.20] higher than among girls in the early formula feeding group, adjusted for maternal pre-pregnancy BMI, prenatal smoking, and maternal age at menarche. This association was similar when implementing multiple imputation to account for missing covariates (MI HR: 1.42 [95% CI, 0.94, 2.14]). Accordingly, MI results will not be presented except where estimates differed. No association was observed in the other feeding groups. After age 12.5, a small decrease in hazard is associated with the primarily breast fed, however, this association is largely attenuated when multiple imputation is applied (MI HR: 0.93 [95% CI, 0.77, 1.11])

A clear trend towards decreasing hazard ratios over time among early soy users was observed when continuous time interaction variables were modeled (Figure 7.3). The most precise hazard ratio estimates occur at age 12 (144 months) (HR: 1.33 [95% CI, 0.96, 1.85]). Hazard ratios for other feeding groups, as compared to the early formula group, remained close to the null over time.

Stratifying on childhood weight status, feeding group was not associated with hazard of menarche among girls of normal weight over time. Among those with a BMI for age above the 85th percentile, the hazard of menarche in the early soy group was elevated (HR: 1.60 [95% CI, 0.98, 2.61]; MI HR:1.66 [95% CI, 1.04, 2.64]). Estimates among those with high BMI for age are particularly imprecise given the low number of subjects in this group. Early

soy feeding was not associated with a change in BMI z-score at any age point between 7 and 9, as compared to either early formula feeding or primarily breast feeding (data not shown).

The distribution of censoring (lost to follow up and administrative) by main exposure status for the total study sample is shown in Table 7.3. Loss to follow up in the early soy group tended to occur earlier in the study than in the other feeding groups, but the overall proportion of censoring in this group is less than other feeding groups (14.8%). The effect of these censored observations on reported effect estimates was assessed in sensitivity analyses (Table 7.4). Hazard ratios prior to age 12.5, while consistently elevated for the early soy group, are slightly attenuated in the "high risk" and "redistributed" models. Estimates after age 12.5 increased in the "low risk" model for both early and late soy feeding, but are otherwise similar to the original analysis.

Discussion

In this study, early life exposure to soy products was associated with a small, but imprecise, increase in the hazard for reaching menarche in early adolescence. Over the full course of the study period the median time-to-menarche did not differ substantially between each of the feeding groups. However time interaction Cox models and Kaplan Meier survival curves support that early soy exposure may be associated with earlier time-to-menarche before age 12.5, while soy does not influence menarche after this time.

The association between early soy exposure and early time-to-menarche is supported biologically. The physiological processes that regulate the onset of puberty are complex, and may have origins in the fetal or neonatal periods. For example, the hypothalamic-pituitary-gonadal (HPG) axis regulates hormone signaling that eventually leads to ovulation and menarche during puberty. This system, which integrates the central nervous system and the

reproductive tract, is also active during the mid-fetal and infant stages of human development (15). Estrogen receptors are present in the hypothalamus (33, 34), suggesting that this region in particular may be susceptible to isoflavone binding. Accordingly, rodent models have demonstrated that neonatal exposure to estrogenic compounds has altered hypothalamic characteristics and function (16).

Our findings also suggest that the effect of an early soy diet was modified by childhood BMI. Early soy product use was associated with earlier time-to-menarche among girls whose BMI was above the 85th percentile for age between ages 7 and 9, but not below the 85th percentile. Early soy exposure was not associated with increased BMI in this study sample, as compared to *early formula* exposure. However, soy isoflavones and other endocrine disruptors have been hypothesized to affect adipogenesis and obesity (6, 35). While based on few observations, the observed association among overweight girls deserves further investigation and may present leads to evaluate potential biological mechanisms relating isoflavone exposure, adiposity, and menarche.

The finding that late soy feeding was not associated with time-to-menarche suggests that dose and timing of soy isoflavone exposure may be important for the induction of developmental effects. The diet becomes increasingly diverse with increasing age, so exposure to soy isoflavones is presumably higher during early infancy when formula provides a greater proportion of the child's nutrition compared to later. Furthermore, the endocrine system is more active in the first several months of infancy, compared to later (36-38), implying that this early stage may be more susceptible to endocrine disrupting effects than later in infancy.

Strom et al.(39) observed various endocrine-sensitive outcomes among adults who had been involved in cow's milk formula and soy formula clinical trials in infancy, and reported no difference in age at menarche between the two feeding groups (adjusted mean difference in years = -0.03 [95% CI, -0.32,0.26]. We also noted little difference in the median age at menarche over the full course of this study. However, using Cox models that isolated risks of menarche during distinct time periods in adolescence, we observed differential effects of soy exposure across time. During earlier, but not later ages, the magnitude of association in our study was notable.

Previous reports by Strom et al. relied largely on age at menarche recall in adulthood for outcome assessment. Validation studies have suggested that accuracy in reporting age at menarche decreases with age (40-42), so non-differential inaccuracies in recall may have biased these results towards the null. To improve upon this, our study used a series of puberty questionnaires that were administered approximately every 1 to 1.5 years between ages 8 and 14.5. This approach maximized the proximity of data collection to the actual menarche event, and was thus a potentially more reliable and useful method of outcome assessment.

The longitudinal nature of our study, however, may have resulted in potential selection bias with respect to inclusion and exclusion of potential study subjects. To address this, important covariates were assessed among eligible ALSPAC subjects who were included and excluded from the final study sample. Of note, excluded subjects had a higher proportion of prenatal smoking, a characteristic that was associated both with earlier time to menarche and also with any type of formula feeding (as compared to breast feeding). The median time-to-menarche in all of the formula feeding groups, then, may be underestimated in our study

population given the artificially low representation of prenatal smoke exposures. This may have biased our results towards the null. However, the relative effect of prenatal smoking on age at menarche is probably small compared to other major predictors, and adjustment for prenatal smoking in all models likely negates any bias that this omission might have created. An additional source of bias may be present if there is an unmeasured characteristic that is associated with both infant feeding and loss to follow up. Certain factors associated with loss to follow up, such as age at menarche and pre-pregnancy BMI, and breast feeding duration were included in models to address this issue, but it is possible that other factors exist that allow for some selection bias to persist.

Sensitivity analysis demonstrated that our results have the potential to be biased by censoring. For early soy exposures, the models that assumed early values for menarche and the redistributed models moved early adolescent hazard ratio estimates towards the null.. The model assigning late ages of menarche to missing data moved the later adolescent hazard ratio estimates away from the null. Based on these findings, if the true distribution of event times among the censored more closely resembles either hypothetical distribution, rather than our assumption that they resemble the event time distribution seen in our population, then our results may be biased. However, no characteristic of the lost to follow up subjects (Table 7.1) support that these subjects are particularly high or low risk. Since these model assumptions are less likely than the assumption that *lost to follow up* event times are distributed similarly to the observed event times in our population, the high and low risk model results do not change our overall interpretation of results. However, the redistributed model allows for more plausible treatment of *lost to follow up* subjects in the early soy group.

The attenuation of association in this model (HR: 1.31 [95% CI, 0.85, 2.02]) does indicate that our results are slightly overestimated due to early censoring in the early soy group.

This analysis was conducted using a large, longitudinal cohort study that is generalizable to the United Kingdom (43). This cohort is unique in that it allows investigators to relate early life exposures to adolescent outcomes, without the hindrance of inaccurate, long term recall. However, as stated previously, the proportion of subjects using soy products in early infancy in this population is small (2%) and our analysis was limited to a white race sample. Thus our power to detect associations was limited and the generalizability of these results may be limited in regions where soy use is more prevalent or populations are not primarily white.

In summary, our study suggests that early life soy product exposure may be contributing to a small decrease in time-to-menarche in white populations, particularly in early adolescence. This study contributes to a growing literature on potential endocrine disruptors and pubertal onset, uniquely making use of a wide exposure contrast between those who were and were not fed soy products in infancy. In contrast to many studies of other environmental endocrine disruptors, where exposures typically occur passively, ubiquitously, and at low levels, the use of soy was actively and continuously fed to infants at regular doses during the defined exposure period. This improves our confidence in the quality of exposure classification and the noted associations.

Since early puberty can put children at risk for obesity, psychosocial abuse, behavior disorders, potential sexual abuse and certain cancers (44, 45), assessing nutritional and environmental factors that may modify the onset of puberty is of great public health significance. Given the growing prevalence of soy in infant diets, further investigation into

these associations, specifically focusing on the role of BMI, deserve attention. Supportive findings from future studies in this area may allow for revised use and manufacturing of soy based products marketed towards infants, as well as a greater understanding of the role of endocrine disruption on development in general.

REFERENCES

1. Santell RC, Chang YC, Nair MG, Helferich WG. Dietary genistein exerts estrogenic effects upon the uterus, mammary gland and the hypothalamic/pituitary axis in rats. J Nutr. 1997 Feb;127(2):263-9.

2. Hwang CS, Kwak HS, Lim HJ, Lee SH, Kang YS, Choe TB, et al. Isoflavone metabolites and their in vitro dual functions: They can act as an estrogenic agonist or antagonist depending on the estrogen concentration. J Steroid Biochem Mol Biol. 2006 Nov;101(4-5):246-53.

3. Dixon RA, Ferreira D. Genistein. Phytochemistry. 2002 Jun;60(3):205-11.

4. Patisaul HB. Phytoestrogen action in the adult and developing brain. J Neuroendocrinol. 2005 Jan;17(1):57-64.

5. Akiyama T, Ishida J, Nakagawa S, Ogawara H, Watanabe S, Itoh N, et al. Genistein, a specific inhibitor or tyrosine-specific protein kinases. J Biol Chem. 1987;262(12):5592-5.

6. Dang ZC. Dose-dependent effects of soy phyto-oestrogen genistein on adipocytes: Mechanisms of action. Obes Rev. 2009 Jan 15.

7. Dang ZC, Audinot V, Papapoulos SE, Boutin JA, Lowik CW. Peroxisome proliferator-activated receptor gamma (PPARgamma) as a molecular target for the soy phytoestrogen genistein. J Biol Chem. 2003 Jan 10;278(2):962-7.

8. Wagner JD, Zhang L, Shadoan MK, Kavanagh K, Chen H, Tresnasari K, et al. Effects of soy protein and isoflavones on insulin resistance and adiponectin in male monkeys. Metabolism. 2008 Jul;57(7 Suppl 1):S24-31.

9. Kavanagh K, Jones KL, Zhang L, Flynn DM, Shadoan MK, Wagner JD. High isoflavone soy diet increases insulin secretion without decreasing insulin sensitivity in premenopausal nonhuman primates. Nutr Res. 2008 Jun;28(6):368-76.

10. Cassidy A, Bingham S, Setchell KD. Biological effects of a diet of soy protein rich in isoflavones on the menstrual cycle of premenopausal women. Am J Clin Nutr. 1994 Sep;60(3):333-40.

11. Hooper L, Ryder JJ, Kurzer MS, Lampe JW, Messina MJ, Phipps WR, et al. Effects of soy protein and isoflavones on circulating hormone concentrations in pre- and post-menopausal women: A systematic review and meta-analysis. Hum Reprod Update. 2009 Jul-Aug;15(4):423-40.

12. Scallet AC, Divine RL, Newbold RR, Delclos KB. Increased volume of the calbindin D28k-labeled sexually dimorphic hypothalamus in genistein and nonylphenol-treated male rats. Toxicol Sci. 2004 Dec;82(2):570-6.

13. Davis EC, Popper P, Gorski RA. The role of apoptosis in sexual differentiation of the rat sexually dimorphic nucleus of the preoptic area. Brain Res. 1996;734(1-2):10-8.

14. Davies MJ, Norman RJ. Programming and reproductive function. Trends Endocrinol Metab. 2002;13(9):386-92.

15. Buck Louis GM, Gray LE,Jr, Marcus M, Ojeda SR, Pescovitz OH, Witchel SF, et al. Environmental factors and puberty timing: Expert panel research needs. Pediatrics. 2008 Feb;121 Suppl 3:S192-207.

16. Bateman HL, Patisaul HB. Disrupted female reproductive physiology following neonatal exposure to phytoestrogens or estrogen specific ligands is associated with decreased GnRH activation and kisspeptin fiber density in the hypothalamus. Neurotoxicology. 2008 Nov;29(6):988-97.

17. Lewis RW, Brooks N, Milburn GM, Soames A, Stone S, Hall M, et al. The effects of the phytoestrogen genistein on the postnatal development of the rat. Toxicol Sci. 2003 Jan;71(1):74-83.

18. Gallo D, Cantelmo F, Distefano M, Ferlini C, Zannoni GF, Riva A, et al. Reproductive effects of dietary soy in female wistar rats. Food Chem Toxicol. 1999 May;37(5):493-502.

19. Casanova M, You L, Gaido KW, Archibeque-Engle S, Janszen DB, Heck HA. Developmental effects of dietary phytoestrogens in sprague-dawley rats and interactions of genistein and daidzein with rat estrogen receptors alpha and beta in vitro. Toxicol Sci. 1999 Oct;51(2):236-44.

20. Blanck HM, Marcus M, Tolbert PE, Rubin C, Henderson AK, Hertzberg VS, et al. Age at menarche and tanner stage in girls exposed in utero and postnatally to polybrominated biphenyl. Epidemiology. 2000 Nov;11(6):641-7.

21. Bhatia J, Greer F, American Academy of Pediatrics Committee on Nutrition. Use of soy protein-based formulas in infant feeding. Pediatrics. 2008 May;121(5):1062-8.

22. Setchell KD, Zimmer-Nechemias L, Cai J, Heubi JE. Isoflavone content of infant formulas and the metabolic fate of these phytoestrogens in early life. Am J Clin Nutr. 1998 Dec;68(6 Suppl):1453S-61S.

23. Knight DC, Eden JA, Huang JL, Waring MA. Isoflavone content of infant foods and formulas. J Paediatr Child Health. 1998 Apr;34(2):135-8.

24. Antignac J, Carou R, Le Bizec B, Cravedi J, Andre F. Identification of phytoestrogens in bovine milk using liquid chromatography/electrospray tandem mass spectrometry. Rapid Commun Mass Spectrom. 2003;17:1256-64.

25. Cao Y, Calafat AM, Doerge DR, Umbach DM, Bernbaum JC, Twaddle NC, et al. Isoflavones in urine, saliva, and blood of infants: Data from a pilot study on the estrogenic activity of soy formula. J Expo Sci Environ Epidemiol. 2009 Feb;19(2):223-34.

26. Setchell KD, Zimmer-Nechemias L, Cai J, Heubi JE. Exposure of infants to phyto-oestrogens from soy-based infant formula. Lancet. 1997 Jul 5;350(9070):23-7.

27. Committee on Hormonally Active Agents in the Environment, National Research Council. Hormonally active agents in the environment. Washington, DC: National Academy Press; 1999.

28. Rubin C, Maisonet M, Kieszak S, Monteilh C, Holmes A, Flanders D, et al. Timing of maturation and predictors of menarche in girls enrolled in a contemporary british cohort. Paediatr Perinat Epidemiol. 2009;23:492-504.

29. Hernan MA, Robins JM. Estimating causal effects from epidemiological data. J Epidemiol Community Health. 2006 Jul;60(7):578-86.

30. Cole SR, Hernan MA. Adjusted survival curves with inverse probability weights. Comput Meth Prog Bio. 2004;75:45-9.

31. Allison PD. Survival analysis using the SAS system: A practical guide. Cary, NC: SAS Press; 1995.

32. A SAS program for the CDC growth charts [homepage on the Internet]. . 2009 November 16, 2009. Available from: http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm.

33. Perez SE, Chen EY, Mufson EJ. Distribution of estrogen receptor alpha and beta immunoreactive profiles in the postnatal rat brain. Brain Res Dev Brain Res. 2003 Oct 10;145(1):117-39.

34. Shughrue PJ, Lane MV, Merchenthaler I. Comparative distribution of estrogen receptor-alpha and -beta mRNA in the rat central nervous system. J Comp Neurol. 1997 Dec 1;388(4):507-25.

35. Newbold RR, Padilla-Banks E, Jefferson WN, Heindel JJ. Effects of endocrine disruptors on obesity. Int J Androl. 2008 Apr;31(2):201-8.

36. Cohen-Bendahan CC, van de Beek C, Berenbaum SA. Prenatal sex hormone effects on child and adult sex-typed behavior: Methods and findings. Neurosci Biobehav Rev. 2005 Apr;29(2):353-84.

37. Chellakooty M, Schmidt IM, Haavisto AM, Boisen KA, Damgaard IN, Mau C, et al. Inhibin A, inhibin B, follicle-stimulating hormone, luteinizing hormone, estradiol, and sex

hormone-binding globulin levels in 473 healthy infant girls. J Clin Endocrinol Metab. 2003 Aug;88(8):3515-20.

38. Sehested A, Juul AA, Andersson AM, Petersen JH, Jensen TK, Muller J, et al. Serum inhibin A and inhibin B in healthy prepubertal, pubertal, and adolescent girls and adult women: Relation to age, stage of puberty, menstrual cycle, follicle-stimulating hormone, luteinizing hormone, and estradiol levels. J Clin Endocrinol Metab. 2000 Apr;85(4):1634-40.

39. Strom BL, Schinnar R, Ziegler EE, Barnhart KT, Sammel MD, Macones GA, et al. Exposure to soy-based formula in infancy and endocrinological and reproductive outcomes in young adulthood. JAMA. 2001 Aug 15;286(7):807-14.

40. Cooper R, Blell M, Hardy R, Black S, Pollard TM, Wadsworth ME, et al. Validity of age at menarche self-reported in adulthood. J Epidemiol Community Health. 2006 Nov;60(11):993-7.

41. Casey VA, Dwyer JT, Coleman KA, Krall EA, Gardner J, Valadian I. Accuracy of recall by middle-aged participants in a longitudinal study of their body size and indices of maturation earlier in life. Ann Hum Biol. 1991 Mar-Apr;18(2):155-66.

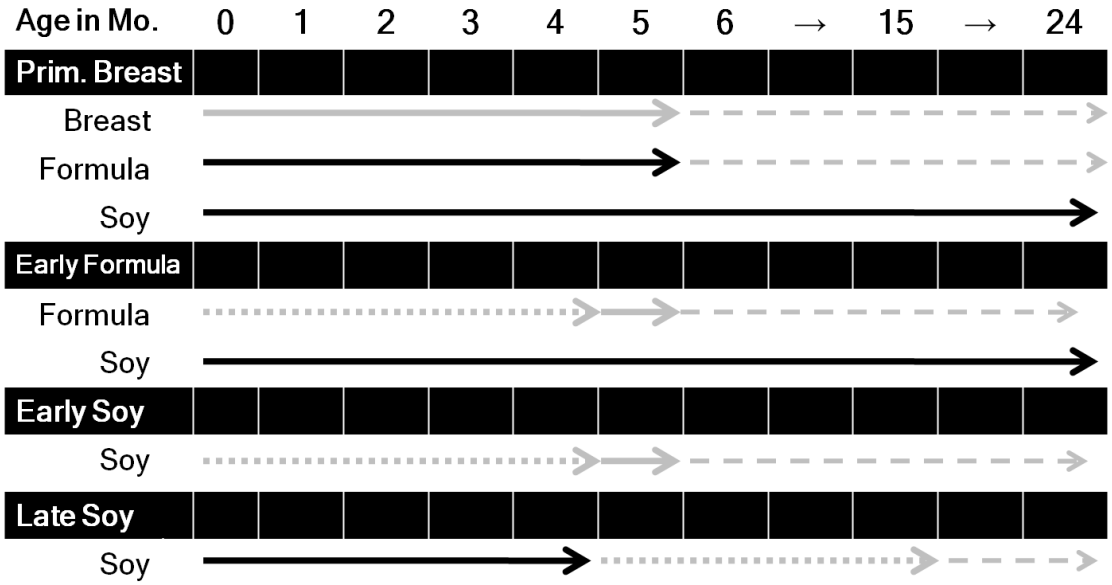
42. Must A, Phillips SM, Naumova EN, Blum M, Harris S, Dawson-Hughes B, et al. Recall of early menstrual history and menarcheal body size: After 30 years, how well do women remember? Am J Epidemiol. 2002 Apr 1;155(7):672-9.

43. Golding J, Pembrey M, Jones R, ALSPAC Study Team. ALSPAC--the avon longitudinal study of parents and children. I. study methodology. Paediatr Perinat Epidemiol. 2001 Jan;15(1):74-87.

44. Jacobson-Dickman E, Lee MM. The influence of endocrine disruptors on pubertal timing. Curr Opin Endocrinol Diabetes Obes. 2009 Feb;16(1):25-30.

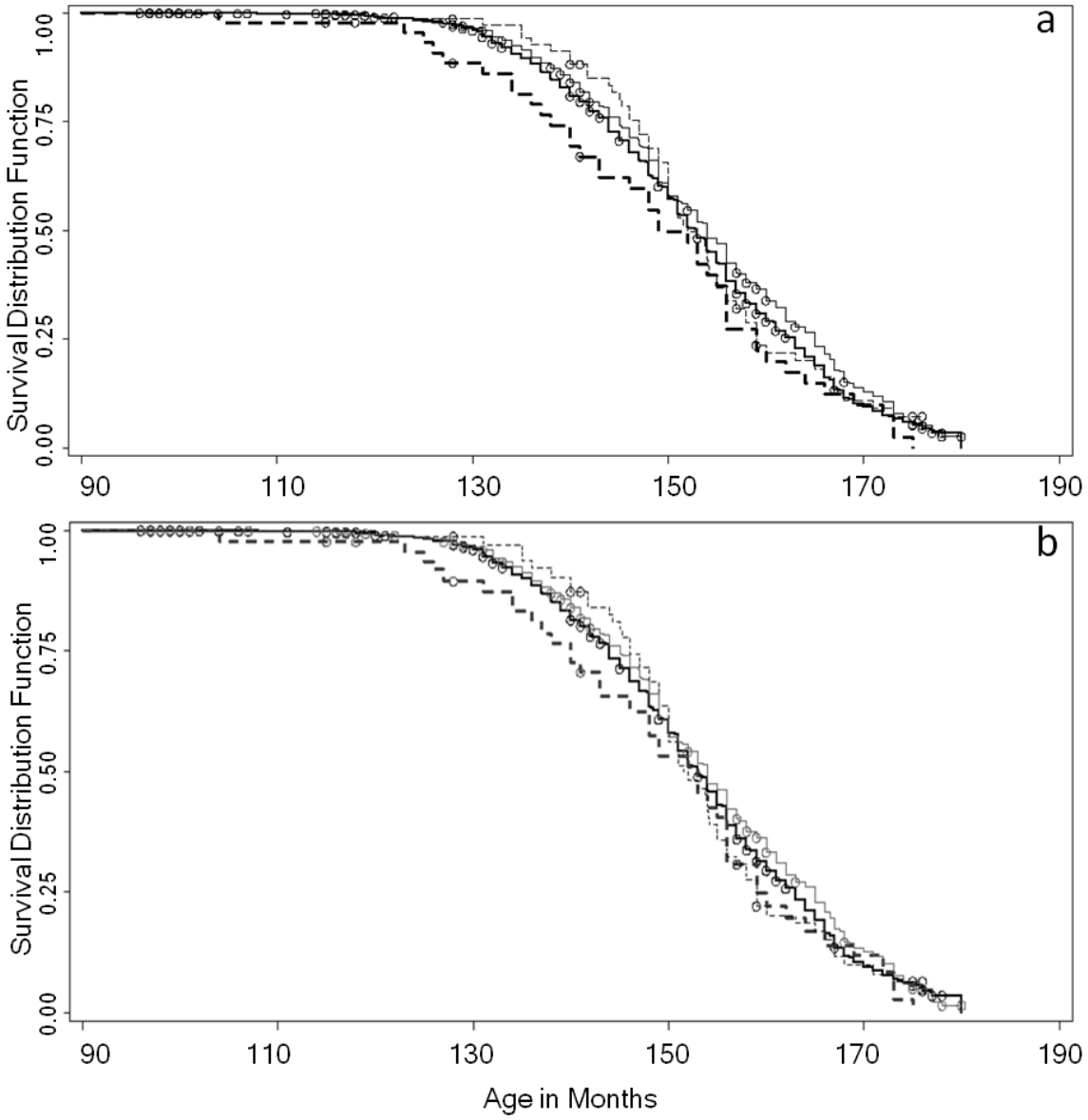
45. Feigelson HS, Henderson BE. Estrogens and breast cancer. Carcinogenesis. 1996 Nov;17(11):2279-84.

Figure 7.1. Exposure Characterization



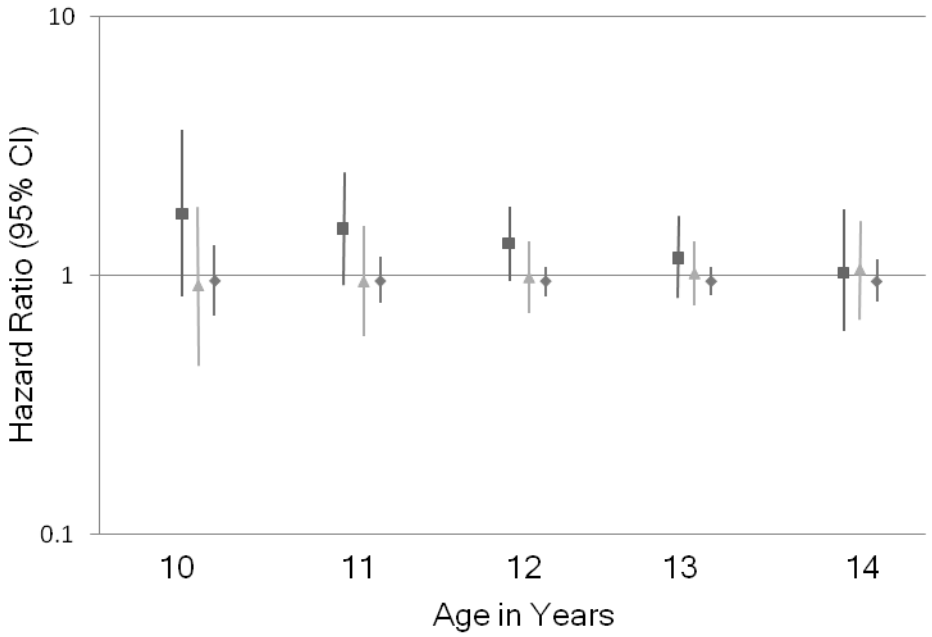
Exposure Characterization. Infant feeding characteristics for the four mutually exclusive exposure categories are illustrated above. Solid arrows (⟶) indicate the *required* time period (age in months) of use for each particular feeding method. Short dotted arrows (⋯▸) indicate the time period (age in months) during which "early use" of a product (either soy-based or traditional milk or formula) *could have been* introduced. Solid bars (•⟶•) indicate time periods for which use of a particular milk or formula product was prohibited. Long dotted arrows (- - ▸) indicate periods of time for which use of a particular feeding method was not restricted or defined according to the exposure definition.

Figure 7.2. Crude (a) and Adjusted (b) Kaplan Meier survival curves, by feeding group



Crude (a) and counfounding-adjusted (b) Kaplan Meier survival curves for each feeding group are shown above. Censored observations are indicated by ○. Survival for each feeding group is incidated as follows: ——— *early formula;* – – – *early soy;* - - - - *late soy;* ——— *primarily breast.* Confounding-adjusted curves are adjusted for pre-pregnancy BMI, maternal age at menarche, and prenatal smoking.

Figure 7.3. Period specific hazard ratios by feeding group (*early formula* = referent)



Period specific hazard ratios for ages 10, 11, 12, 13 and 14, are shown above. Estimates were derived from Cox proportional hazards models with continuous exposure*time and covariate*time interaction terms.

The *early formula* is the referent exposure category. ■ indicate hazard ratios for the *early soy* group; ▲ indicate hazard ratios for the *late soy* group; ◆ indicate hazard ratios for the *primarily breast fed* group. Vertical lines correspond to 95% confidence intervals.

Table 7.1. Characteristics of eligible[a] ALSPAC study sample (n = 5,230), distinguished as those included in the present analysis (n = 2,884) and those excluded for missing exposure or outcome data (n = 2,346)

	Study Sample Included in Present Analysis			Excluded
	n (%)	Median Time-to-Menarche (mo) [IQR[b]]	Lost to Follow Up[c] n (%)	n (%)
Study Sample	2884	153 [144 – 163]	658	2346
Feeding Group				
Early Formula	2124 (73.7)	153 [144-163]	513(78.0)	607 (77.8)*
Early Soy	54 (1.9)	149 [140-159]	8 (1.2)	12 (1.5)
Late Soy	86 (3.0)	151 [146-159]	13 (2.0)	24 (3.1)
Primarily Breast	620 (21.5)	154 [145-165]	124 (18.8)	137 (17.6)*
Missing	--	--	--	1566
Birth Weight				
≤ 2500g	35 (1.2)	151 [146-160]	10 (1.5)	53 (2.3)*
>2500g	2808 (98.8)	153 [144-163]	638 (98.5)	2266 (97.7)
Missing	41	153 [144-161]	10	27
Mean (SD) (g)	3429 (436)	-	-	3408 (463)*
Breast Feeding Duration				
Mean (SD) (mo)	4.3 (4.6)	-	-	3.8 (4.4)*
Missing	0	-	0	142
Pre-pregnancy BMI				
≥ 25	535 (20.1)	150 [140-160]	128 (21.4)	436 (21.6)
< 25	2124 (79.9)	154 [145-164]	470 (78.6)	1584 (78.4)
Missing	225	153 [143-162]	60	326
Mean (SD)	22.8 (3.7)	-	-	23.1 (4.0)
Maternal Age at Menarche				
8-11	490 (19.2)	146 [138-155]	97 (16.9)	379 (19.5)
12-14	1748 (68.5)	154 [145-163]	384 (67.0)	1299(66.8)
15+	313 (12.3)	161 [151-171]	92 (16.1)	267 (13.7)
Missing	333	154 [146-165]	85	401
Mean (SD)(y)	12.8 (1.5)	-	-	12.9 (1.6)
Prenatal Smoking				
Yes	423 (15.0)	152 [142-163]	128 (20.5)	419 (20.8)*
No	2387 (85.0)	154 [144-163]	498 (79.6)	1595 (79.2)
Missing	64	153 [147-160]	32	332
BMI for age, age 7-9				
>85th Percentile	754 (28.1)	148 [137-155]	142 (25.0)	340 (27.9)
≤ 85th Percentile	1926 (71.9)	156 [147-165]	425 (75.0)	880 (72.1)
Missing	204	154 [142-162]	91	1126
Infant Health at 6 months				
Healthy/ Minor Problems	2745 (98.0)	153 [144-163]	609 (96.8)	1813 (96.5)
Sometimes ill/unwell	56 (2.0)	153 [148-170]	20 (3.2)	66 (3.5)*
Missing	83	154 [144-160]	29	467
Maternal Age at Delivery				
≤ 30	1924 (66.7)	154 [144-163]	471 (71.6)	1692 (72.1)*
>30	960 (33.3)	153 [144-164]	187 (28.4)	654 (27.9)*
Missing	0	-	0	0
Mean (y)	28.7 (4.5)	-	-	27.6 (5.0)*

*$p < 0.05$, comparing proportion of excluded to proportion included.
[a] Term, singleton, white females born in the ALSPAC cohort.
[b] Interquartile Range: 25th – 75th Percentile
[c] Subjects who completed at least one "Growing and Changing" questionnaire, but did not report menarche and did not complete the study; a subset of study sample (n = 2,884).

Table 7.2. Adjusted[a] hazard ratios and median time-to-menarche estimates

Feeding Group	**N**	**$HR_{< 150 \text{ months}}$ [95% CI]**	**CLR[b]**	**$HR_{\geq 150 \text{ months}}$ [95% CI]**	**CLR**	**Adjusted Median Time-to-Menarche[a]** [IQR]
Early Formula	1718	1.0		1.0		153 [144-163]
Early Soy	46	1.42 [0.92, 2.20]	2.39	1.11 [0.71, 1.74]	2.45	153 [140-159]
Late Soy	72	0.84 [0.55, 1.29]	2.35	1.13 [0.80, 1.59]	1.99	151.5 [146-159]
Primarily Breast	518	1.03 [0.87, 1.22]	1.40	0.89 [0.76, 1.04]	1.37	154 [145-165]
Total Observations			2354			
Events			1774			
Censored (%)			580 (24.6)			

[a] Adjusted for pre-pregnancy BMI, maternal age at menarche, prenatal smoking (yes, no)
[b] Confidence Limit Ratio

Table 7.3. Proportion of censoring in study sample: lost to follow up (*LTF*) and administrative censored subjects [n (%)], by feeding group

Feeding Group	Lost To Follow Up: Approximate Age at Last Completed Questionnaire						Administratively Censored[a]	*Total Censored*[b]
	8[a]	9 ½[a]	10 ½[a]	11 ½[a]	13[a]	*Total LTF*[b]		
Early Formula	63 (10.9)	84 (14.5)	97 (16.8)	123 (21.2)	146 (24.8)	*513 (24.1)*	66 (11.4)	*579 (27.3)*
Early Soy	1 (12.5)	3 (37.5)	3 (37.5)	1 (12.5)	0 (0)	*8 (14.8)*	0 (0)	*8 (14.8)*
Late Soy	2 (11.1)	0 (0)	4 (22.2)	5 (27.8)	2 (11.1)	*13 (15.1)*	5 (27.8)	*18 (20.9)*
Primarily Breast	20 (13.9)	17 (11.8)	16 (11.1)	34 (23.6)	37 (25.7)	*124 (20.0)*	20 (13.9)	*144 (23.2)*
Total	86 (11.5)	104 (13.9)	120 (16.0)	163 (21.8)	185 (24.7)	*658 (22.8)*	91 (12.2)	*749 (26.0)*

[a] Percent expressed as percentage of total censored within feeding group

[b] Percentage expressed as percentage of within feeding group totals for the study sample

Table 7.4. Sensitivity analysis hazard ratios and median survival times

Feeding Group	HR $_{<150}$ [95% CI] [a]	CLR[b]	HR $_{\geq 150}$ [95% CI][a]	CLR
Low Risk Follow Up Time Assumption				
Early Formula	1.0		1.0	
Early Soy	1.43 [0.93, 2.21]	2.37	1.51 [0.97, 2.35]	2.42
Late Soy	0.84 [0.55, 1.29]	2.35	1.37 [0.97, 1.93]	1.99
Primarily Breast	1.02 [0.85, 1.16]	1.36	0.99 [0.85, 1.16]	1.36
High Risk Follow Up Time Assumption				
Early Formula	1.0		1.0	
Early Soy	1.19 [0.81, 1.76]	2.17	0.95 [0.61, 1.49]	2.44
Late Soy	0.82 [0.57, 1.17]	2.05	1.03 [0.74, 1.43]	1.93
Primarily Breast	0.98 [0.76, 1.01]	1.33	0.88 [0.76, 1.01]	1.33
Redistribution of Early Soy Censor Times				
Early Formula	1.0		1.0	
Early Soy	1.31 [0.85, 2.02]	1.87	1.07 [0.68, 1.66]	1.87
Late Soy	0.84 [0.54, 1.29]	1.70	1.13 [0.80, 1.59]	1.74
Primarily Breast	1.02 [0.87, 1.21]	1.25	0.89 [0.76, 1.03]	1.38

[a]Adjusted for pre-pregnancy BMI, maternal age at menarche, and prenatal smoking
[b]Confidence Limit Ratio

VIII. CONCLUSIONS

Soy-based infant formula is a commonly used alternative to cow's milk based infant formula, particularly in instances of milk intolerance or preference for a vegan diet. However, soy products contain high levels of phytoestrogens, plant compounds with structural and functional similarity to 17β-estradiol, which may have long term effects on developing infants. Animal models have shown that the phytoestrogens found in soy are capable of inducing disruptive effects on the endocrine system. However, few longitudinal studies of early life soy exposure have been conducted in human populations. Therefore, the overall long term safety of soy exposure in infancy is unclear.

This epidemiologic investigation aimed to assess the association between early life soy product exposure and hormonally sensitive outcomes related to neurological and reproductive development. This investigation was carried out in the Avon Longitudinal Study of Parents and Children (ALSPAC), a pregnancy cohort of over 14,000 live births recruited in 1991-1992 in the United Kingdom, and followed to the present day.

Three major analyses were carried out to address two specific aims,

1) to assess the effects of soy product exposure on early life sexually dimorphic cognition and behavior, and

2) to assess the effect of early life soy product use on time-to-menarche in adolescent females.

A common exposure definition was used for all analyses, and is fully described in *Section IV. Methods* of this document. A summary of this study's findings are provided below, along

with a discussion of the strengths, limitations, implications, and future directions of this research.

a. Summary, Aim 1

The objective of Aim 1 was to assess the effects of soy product exposure on early life sexually dimorphic cognition and behavior. This aim was investigated using two approaches. First, in Aim 1.1, we assessed word comprehension and word production at 15 and 24 months across the four feeding groups, with the intention of emphasizing the contrast between soy fed children and formula fed children, within each gender. Next, in Aim 1.2, we similarly assessed the within gender differences for gender-role play behaviors across feeding groups. Both of these outcomes have been shown to have clear, sexually dimorphic patterns in normal development, and so were used to assess the overall objective of Aim 1.

The hypotheses behind Aims 1.1 and 1.2 were that soy exposed boys would exhibit "feminized" behavior, as exhibited in accelerated language acquisition and lower scores on the gender-role play behavior assessment. This hypothesis was based on ample animal literature, as cited throughout this document. A masculinizing effect (decreased language acquisition and higher gender-role play behavior scores) was hypothesized for soy exposed girls, although the body of previous literature supporting this effect was not as consistent as it was for boys.

The results for Aim 1.1 are summarized in Section V of this document. Among boys, a minimal increase in word comprehension and word production was observed between 15 and 24 months among both early and late soy exposed boys. However, the differences between the soy fed boys and the early formula fed boys were small and imprecise. Furthermore, results in the early and late soy group were not substantially different from the

results observed in the primarily breast fed group, with the exception that the estimates in the primarily breast fed group had increased precision due to larger sample size.

While these findings do agree with the hypothesized direction of effect, sample size limitations, combined with modest effect estimates, prevent valid conclusions from being drawn in this analysis with respect to soy use. Recalling preliminary power calculations (Section IV. Methods), a mean difference of approximately 11 words was the minimum detectible effect, at 80% power, given the sample size of the early soy group. Since the crude mean difference between early soy word comprehension and word production did not exceed 8 words (Tables 5.2 and 5.3), we can conclude that this analysis was not sufficiently powered to detect the subtle effect that was observed. While our results can *suggest* a very small increase in language acquisition over time, our effect estimates were imprecise. This study sample was not appropriate for answering this research question in boys without acknowledging a high likelihood for Type II error.

In girls, the effect observed in the early soy fed group was greater than was observed in the boys. Therefore, the limitations with regard to power did not apply to the analysis in girls. An elevated effect was observed in word comprehension and word production among both early soy fed girls and primarily breast fed girls. In both comparisons, the effect in the early soy fed girls was substantially larger than what was observed in the primarily breast fed girls. No effect was observed in the late soy fed group. Interestingly, the direction of effect observed in the early soy fed girls is opposite of what was hypothesized. Rather than a "masculinizing effect," or decrease in acquisition rate, soy appeared to induce a "hyper-feminizing" effect. Given that our confidence in the direction of effect in our hypothesis was somewhat weak, this effect is not entirely surprising. Alternatively,this result may suggest

that soy was inducing a biological effect that was not related to sexual-dimorphism, per se, but rather an alternative biological mechanism that enhanced language development. In addition, while all analyses were adjusted for breast feeding duration, maternal age, the presence of an older sibling, maternal parenting, day care attendance, and potential hearing problems, it is possible that unmeasured confounding was biasing estimates away from the null.

For Aim 1.2, gender-role play behavior scores were slightly elevated in early soy exposed boys and girls. As in Aim 1.1, the effect in boys was negligible. However, our study was likely not sufficiently powered to allow for strong conclusions to be drawn, given the high likelihood for Type II error. The effect in girls was more substantial, although still imprecise.

In contrast to Aim 1.1, the finding among early soy fed girls was in the masculine, not feminine, direction. Because the observed effect was small (an increase of 2.7 points on a 'pseudo T-score' scale of approximately 100), and did not place soy fed girls outside the normal range of female scores, it would not be appropriate assert that soy fed girls have been "masculinized." Rather, they may be considered slightly *less feminine* than their unexposed counterparts, and experience an effect similar to having an older brother in the home.

It is difficult to interpret Aims 1.1 and 1.2 in summary, because the effects in girls suggest that early soy exposure is inducing opposite effects in each analysis. Several possible explanations may account for this discrepancy. First, it is possible that soy isoflavones are acting as both estrogen agonists and antagonists in various regions of the brain, resulting in the somewhat inconsistent direction of effect. In addition, it is also possible that language acquisition and gender-role play behavior are not adequate indicators

of sexually-dimorphic development, as they were intended to be. Simply assessing an increase or decrease in language or play behavior score may be too crude a measure to truly assert any sort of "masculinzing" or "feminizing" effect, particularly when effect estimates are small, and not outside the range of normal within-gender development.

Next, it is possible that other components of soy products, apart from isoflavones, influence early development in ways that are not necessarily related to sex hormone concentrations and sexually dimorphic development. For example, compared to breast milk and cow's milk formula, soy formula has high concentrations of phosphatidylcholine (1), an essential precursor to phospholipids and an important part of early development (2). The results of our language acquisition analysis could be attributed to this property of soy formula, as opposed to isoflavone content, in which case a sexually-dimorphic effect would not necessarily be expected.

Lastly, unmeasured confounding is also important here. Both outcomes are multi-factorial in nature, in that many social, environmental, and biological factors contribute to their development. These analyses attempted to control for many of these factors, but it is plausible that unmeasured confounding persists. For example, a common characteristic may be shared among parents that chose soy formula over traditional formula. This common characteristic may have also influenced parenting styles, such that soy fed girls were conditioned to be more physical, outgoing, assertive, and verbal than other girls.

b. Summary, Aim 2

The objective of Aim 2 was to assess the effect of early life soy product use on time-to-menarche in adolescent females. The hypothesis behind this aim was that, in accordance with a substantial animal literature on genistein, other phytoestrogens, and other exogenous

estrogens, early life soy exposure would be associated with earlier age at menarche. For this aim, the outcome was assessed using a series of puberty questionnaires administered approximately annually between the ages of 8 and 14.

In this study, early life exposure to soy products was associated with a small, but imprecise, increase in the hazard for reaching menarche in early adolescence. When averaged over the course of the study period, time-to-menarche did not differ substantially between each of the feeding groups. However, categorical and continuous time interaction models and adjusted Kaplan Meier curves support that early soy exposure may be associated with reaching menarche before age 12.5, and that the influence of soy on menarche after this time is attenuated.

Our findings also suggested that the effect of an early soy diet was modified by childhood BMI. Early soy product use was associated with earlier time-to-menarche among girls whose BMI was above the 85th percentile for age between ages 7 and 9, but not below the 85th percentile. Early soy exposure was not associated with increased BMI in this study sample.

Sensitivity analysis demonstrated that our results also have the potential to be biased by censoring. In hypothetical models where *lost to follow up* follow up times were modeled as event times at the time of dropout, hazard ratio estimates before age 12.5 moved towards the null. Since this model assumed the earliest possible even time for lost to follow up subjects, it can be viewed as a lower bound of the amount of bias that is introduced by lost to follow up subjects. Based on these findings, we acknowledge that, if the true distribution of event times among the censored more closely resembles this low-end hypothetical distribution, rather than our assumption that they resemble the event time distribution seen in our population, then our results may be slightly overestimating effects. Additionally, if the

follow up times for the *lost to follow up* early soy subjects were redistributed to resemble follow up times of other exposure groups, estimates before age 12.5 were also slightly attenuated. These "redistributed" results suggest that early censoring in the early soy group was affecting results, and are acknowledged.

Overall, the aim of assessing the effects of early soy exposure on time-to-menarche was successfully met in this analysis. Although potentially biased by loss to follow up, both proportional hazards models supported that, before age 12.5, early soy fed girls were achieving menarche earlier in this study sample. As with all analyses described here, the number of soy exposed individuals was small, and our results were imprecise. As will be discussed in greater detail below, these findings do not support any sort of causal inference with respect to early life soy exposure and early age at menarche. Rather, this research should be viewed as an early contribution to a currently sparse literature on early life endocrine disruption and pubertal onset, and should be used to generate hypotheses for future studies rather than promote changes in public health practice.

c. Strengths

This study addresses novel, biologically based research questions that investigate the role of early life soy product exposure in sexually dimorphic and reproductive development. Previous studies of soy use have largely focused on gross developmental outcomes, such as growth and immune response in first few years of life. No studies have addressed the question of gender-specific effects of soy on sexually differential development, and few have explored the role of early life soy exposure on timing of pubertal onset. The outcomes studied here were specifically selected to explore changes in hormonally-sensitive development, so as to address the role of soy isoflavones as endocrine disruptors.

Particularly for Aim 1, the outcomes are subtle, and may not convey significant public health impact. However, all outcomes explored here do contribute to an understanding of soy exposure as an endocrine disruptor, which may have great public health impact given the prevalence of soy use in the American diet.

This study utilizes a large, population based longitudinal cohort. Applying the ALSPAC study data to this research question is efficient in many ways. Because of the large sample size, this study was able to identify a suitable number of soy exposed children, despite a low prevalence of use in the United Kingdom (UK) during the 1990s. The ALSPAC study was also expertly designed to collect longitudinal and repeated measures data that were utilized here without the time or financial expense that is associated with other longitudinal studies.

This study population has been shown to be generally representative of the UK on a range of demographic factors. Thus, results from this study will be generalizable to the larger UK population, with the exception of Aim 2, where generalizablity is limited only to the white population. The ALPSC population is also likely representative of "real world" soy use prevalence and patterns. The patterns associated with infant feeding often involve a degree of mixing and switching between breast milk, and various types of infant formula. This behavior is reflected in this study. Most previous studies have been performed in randomized clinical trial settings. While clinical trials are advantageous in many respects, it is also informative to be able to assess the effects of soy as it is actually used, provided careful controlling for confounding is possible.

Outcome assessment for all study aims was a major strength of this study. Word comprehension and production (MCDI, Aim 1.1) and gender role play behavior (PSAI, Aim 1.2) were assessed efficiently through parent report. Both instruments have been used

widely in other studies to explore developmental variability with respect to nutritional and environmental factors (3-8). These instruments have also been shown to be reliable and valid (9, 10). Any misclassification in outcome assessment is likely to be non-differential with respect to exposure status, meaning that any bias introduced by inaccuracies in outcome assessment should lead to overly conservative results. In addition, the assessment of age at menarche in Aim 2 was well designed to improve upon previous studies of time-to-menarche. Whereas many previous studies have relied largely on adult recall of age at menarche, our study used a series of puberty questionnaires that were administered approximately annually between ages 8 and 14 to assess the age in months that menarche occurred. This approach maximized the proximity of data collection to the actual menarche event, as well as the precision of the age estimate (months instead of years), and was thus a more reliable and useful method of outcome assessment than has been used in many previous studies.

The prospective nature of the ALSPAC study also allowed for quality exposure assessment. Infant feeding habits were assessed at several time points in infancy. The 6 month assessment, which was predominantly used for our exposure assessment, occurred rather proximally to the time period of interest, and thus can be viewed as an accurate method of recall of early life infant feeding habits. Our exposure definition also took into account the age at which children were introduced to the various formulas, as well as an implied duration of use (use at or beyond 6 months of age). This definition is advantageous because it isolates subjects who used formula for at least one month from those who may have reported only "ever" using the formula.

A final strength of this study, and of the ALSPAC study in general, is the availability of extensive covariate data. While the potential for bias due to unmeasured confounding still persists, these analyses were able to adjust for most reasonably identified confounders.

d. Limitations

Despite being a very large prospective cohort, the proportion of infants using soy products in this sample is low compared to what might have been observed in other regions or time periods. The power to detect small changes in effect is therefore limited, and thus this study is only suited to detect moderate to large effects of exposure. This limitation was particularly clear in the analysis of boys in Aim 1, but was also evident in consistently imprecise effect estimates, as indicated by wide confidence intervals.

In addition, this study relied on questionnaire-based maternal report of both exposure and outcome. Some misclassification may result. With respect to outcome misclassification, it is not likely that it occurred differentially across exposure groups, so any existing bias should be towards the null. Regarding exposure, reliance on questionnaire data is admittedly less than ideal. Improvements in future studies should attempt to assess total isoflavone exposure through the use of biomarkers, as opposed to maternal report of feeding method. Even within exposure categories, as they were defined in this study, there is potential for considerable heterogeneity of exposure to exist. For example, the age of soy product initiation was shown to vary substantially across the first 4 months of life (Appendix Table A.1). However, subjects that started soy formula at birth and those that started at age 4 months are treated as though they had identical exposures. Our effect estimates, then, may be underestimating the effect of exclusive soy feeding from birth, or over estimating effects of moderate feeding that began several months later.

As mentioned previously, there is potential that our findings may be affected by some degree of unmeasured confounding. Since only approximately 2% of our study sample included parents who chose to feed soy formula, it is possible that some unmeasured factor motivated parents to choose a somewhat rare feeding method. It is also possible that this same unmeasured factor may have influenced cognitive and behavioral development, as well as other environmental and behavioral factors that influenced age at menarche. While this study attempted to identify predictors of soy use (Appendix Table A.2), the examined covariates may not fully characterize personal characteristics of infants or parents, or parenting styles that may also have influenced development in soy fed children.

Finally, given the longitudinal nature of this study, selection bias is important to consider as a potential limitation to this study. For both aims, analyses were limited to subjects that had data on both the exposure and outcome. This requirement limits the overall generalizablity of our study, in that our findings can only truly be generalized to subjects who *would have* stayed in this study long enough to complete all necessary questionnaires. With respect to exposure status, this may have occurred differentially if a large proportion of soy users dropped out of the study before reporting soy use because they, for example, had a sick baby. We therefore may only be capturing a select portion of actual soy users in this study. It is unlikely, however, among those with complete feeding data, that loss to follow up by outcome was differential by exposure group, since the proportion of subjects in each feeding group was consistent across the 3 analyses. In Aim 2, loss to follow up after age 8 (subjects that contributed to the puberty assessment, but dropped out before reaching menarche) may be particularly problematic, as shown in the sensitivity analyses which indicate that these loss

to follow up individuals could possibly have biased our results. However, such a bias would only exist if extreme assumptions applied to these individuals, which is unlikely.

e. Implications

Collectively, the research presented here provides a much needed and worthwhile contribution to the literature in the fields of endocrine disruptor research and pediatric epidemiology. Using three distinct outcomes, we have demonstrated small associations between early life soy exposure and developmental effects among exposed girls. This study was not sufficiently powered to detect small effects in the male study subjects. Since this is the first study to look at several of these outcomes, it is premature to assert that early life soy exposure has a true effect on developmental outcomes. However, our findings do provide groundwork for future studies of this kind.

The findings presented here are novel, particularly with respect to language development and play behavior. Again, it is not appropriate at this stage to make any sort of causal inference with respect to early life soy use and developmental outcomes, nor is it appropriate to make firm conclusions regarding the potential for soy isoflavones to act as endocrine disruptors based on this work. Instead, this study should be regarded as a preliminary, hypothesis generating body of work. Future studies of infant feeding should integrate the novel research questions explored in these analyses into study protocols. This is feasible, since the outcomes assessed here were low-cost and easily administered questionnaires. A greater body of knowledge on how early life feeding of soy based products affect subtle, sex-specific developmental outcomes is needed to clarify the findings of this work, as well as to more fully understand the impact that early life soy exposure may have on the population level.

f. Future Directions

This work suggests a need for future research on the effects of early life soy exposure. Future studies should identify a larger population of soy formula users from which to sample the study participants. Ideally, biomarkers of isoflavone concentrations could be used to improved exposure classification. Detailed, month to month feeding data would also be an improvement on the methods used here.

In addition, to gain understanding of the overall biologic activity of soy products, it may also be interesting and useful in future studies to assess the effects of soy exposure in the prenatal period, as well as later in childhood. Such studies would meaningfully contribute to the understanding of developmental periods of susceptibility for the outcomes assessed here.

REFERENCES

1. Holmes-McNary MQ, Cheng WL, Mar MH, Fussell S, Zeisel SH. Choline and choline esters in human and rat milk and in infant formulas. Am J Clin Nutr. 1996 Oct;64(4):572-6.

2. Wurtman RJ. Synapse formation and cognitive brain development: Effect of docosahexaenoic acid and other dietary constituents. Metabolism. 2008 Oct;57 Suppl 2:S6-10.

3. Daniels JL, Rowland AS, Longnecker MP, Crawford P, Golding J, ALSPAC Study Team. Maternal dental history, child's birth outcome and early cognitive development. Paediatr Perinat Epidemiol. 2007 Sep;21(5):448-57.

4. Daniels JL, Longnecker MP, Rowland AS, Golding J, ALSPAC Study Team. University of Bristol Institute of Child Health. Fish intake during pregnancy and early cognitive development of offspring. Epidemiology. 2004 Jul;15(4):394-402.

5. Hines M, Golombok S, Rust J, Johnston KJ, Golding J, Avon Longitudinal Study of Parents and Children Study Team. Testosterone during pregnancy and gender role behavior of preschool children: A longitudinal, population study. Child Dev. 2002 Nov-Dec;73(6):1678-87.

6. Hines M, Johnston KJ, Golombok S, Rust J, Stevens M, Golding J, et al. Prenatal stress and gender role behavior in girls and boys: A longitudinal, population study. Horm Behav. 2002 Sep;42(2):126-34.

7. Rust J, Golombok S, Hines M, Johnston K, Golding J, ALSPAC Study Team. The role of brothers and sisters in the gender development of preschool children. J Exp Child Psychol. 2000 Dec;77(4):292-303.

8. Vreugdenhil HJ, Slijper FM, Mulder PG, Weisglas-Kuperus N. Effects of perinatal exposure to PCBs and dioxins on play behavior in dutch children at school age. Environ Health Perspect. 2002 Oct;110(10):A593-8.

9. Fenson L, Dale PS, Reznick JS, Bates E, Thal DJ, Pethick SJ. Variability in early communicative development. Monographs of the Society for Research in Child Development. 1994;59(5, Serial No. 242):i-185.

10. Golombok S, Rust J. The pre-school activities inventory: A standardized assessment of gender role in children. Psychological Assessment. 1993;5(2):131-6.

APPENDIX I: ADDITIONAL TABLES

Tables:

Table A.1. Age of introduction of soy product in early soy exposure group (n = 182)

Age (months) of Soy Introduction	N (%)
0	35 (19%)
1	20 (11%)
2	30 (16%)
3	41 (23%)
4	56 (31%)

Table A.2. Covariate assessment by feeding group

	Breast	Early Formula	Early Soy	Late Soy	Total
N*	1667	6298	183	288	8436
Male Gender, n (%)	827 (49.6)	3263 (51.8)	**109 (59.6)**	158 (54.9)	4357 (51.7)
Gestational Age, mean (SD)	39.8 (1.25)	39.8 (1.32)	39.9 (1.31)	39.8 (1.26)	38.36 (5.5)
Birthweight, n(%)					
≥ 2500 g	1636 (99.3)	6115 (98.2)	173 (96.1)	278 (96.5)	13089 (98.4)
<2500 g	12 (0.7)	**112 (1.8)^{B}**	**7 (3.9)^{B,F}**	**6(2.1)^{B}**	812 (1.6)
Birthweight, mean (sd) g	3537 (448)	**3458 (472)^{B,L}**	3495 (531)	3544 (497)	3477 (470)
Child Race, n (%)					
White	1516 (95.0)	5638 (96.0)	165 (96.5)	256 (91.8)	7575 (95.7)
Non-White	79 (5.0)	234 (4.0)	6 (3.5)	**23 (8.2)^{B,F,E}**	342 (4.3)
Breast Feeding Duration					
Never	0	1494 (23.7)	42 (23.0)	26 (9.0)	1562 (19.5)
<3 months	0	2456 (39.0)	58 (31.7)	55 (19.1)	2569 (32.1)
3-<6 months	0	1231 (19.6)	35 (19.1)	34 (11.8)	1300 (16.3)
6+ months	**1667 (100)**	**695 (11.0)^{B}**	**38 (20.8)^{B,F}**	**165 (57.3)^{B,E,F}**	2565 (32.1)
Mean (sd) months of BF	**10.4 (3.2)**	**2.2 (2.8)**	**3.3 (3.9)**	**6.6 (4.8)**	4.2 (4.5)
BMI at age 8 (TOTAL)					
Mean (SD)	16.9 (2.4)	**17.2 (2.5)^{B,L}**	17.4 (2.5)	16.6 (2.0)	17.1 (2.5)
Missing	*	*	*	*	*
>85th percentile					
BMI at age 8 (Girls)					
Mean (sd)	16.9 (2.5)	**17.4 (2.6)^{L}**	17.6 (2.6)	16.7 (2.0)	17.3 (2.6)
>85th percentile	139 (23.9)	**533 (28.7)^{B,L}**	**15 (34.9)^{L}**	16 (18.8)	703 (27.4)
BMI at age 8 (Boys)					
Mean (sd)	16.9 (2.3)	17.0 (2.4)	17.3 (2.4)	16.6 (2.1)	16.9 (2.3)
>85th percentile	123 (23.0)	461 (24.0)	17 (25.0)	21 (19.1)	622 (23.6)
Infant Health Condition Reported by 6 Months, n(%)					
Colic	**551 (34.9)**	**2411 (39.7)^{B}**	**103 (59.2)^{B,F}**	**131 (47.3)^{B,F,E}**	3196 (39.5)
Earache	**153 (9.7)**	**599 (9.9)**	**28 (16.1)^{B,F,L}**	**26 (9.4)**	806 (10.0)
Vomiting	**391 (24.8)**	**1970 (32.5)^{B}**	**80 (46.0)^{B,F}**	**109 (39.4)^{B,F}**	2550 (31.5)
Diarrhea	**288 (18.2)**	**2200 (36.2)^{B}**	**80 (46.0)^{B,F,L}**	**101 (36.5)^{B}**	2669 (33.0)
Rash	**603 (38.2)**	**2266 (37.3)**	**83 (47.7)^{B,F}**	**151 (54.5)^{B,F}**	3103 (38.3)
Cough	**970 (61.4)**	**3987 (65.7)^{B}**	**125 (71.8)^{B}**	**194 (70.0)^{B}**	5276 (65.1)
Wheezing	281 (17.8)	**1289 (21.2)^{B}**	**59 (33.9)^{B,F}**	**80 (28.9)^{B,F}**	1709 (21.1)
Suspected Hearing Problem	62 (4.0)	195 (3.2)	**11 (6.4)^{F}**	**13 (4.8)**	281 (3.5)
Milk Allergy	19 (1.2)	69 (1.1)	**91 (52.3)^{B,F,l}**	**58 (20.9)^{B,F}**	237 (2.9)
LATER Milk Allergy (81M) - Yes	16 (1.4)	55 (1.4)	**17 (14.9)^{B,F}**	**29 (15.8)^{B,F}**	117 (2.1)

Table A2 continued

4 weeks: Mom thinks crying is a problem					
Yes/Sometimes	65	308	28	20	377
No	1544 (96.0)	5696 (94.9)	**147 (84.0)[B,F,L]**	256 (92.8)	7643 (94.8)
Maternal Perception of Infant Health					
Very Healthy	**1034 (65.7)**	**3677 (60.8)[B]**	**56 (32.6)[B,F,L]**	**132 (48.2)[B,F]**	**4899 (60.7)**
Minor Problems	**501 (31.9)**	**2195 (36.3)**	**93 (54.1)**	**130 (47.5)**	**2919 (36.2)**
Sometimes quite ill	**34 (2.2)**	**133 (2.2)**	**16 (9.3)**	**9 (3.3)**	**192 (2.4)**
Mostly unwell	**4 (0.3)**	**43 (0.7)**	**7 (4.1)**	**3 (1.1)**	**57 (0.7)**
Childhood Soy Consumption					
Yes (Veget/Veg/High Soy Meat or Milk)	**52 (4.4)**	**115 (2.8)[B]**	**12 (10.3)[B,F]**	**25 (12.3)[B,F]**	5353 (96.3)
No	**1130 (95.6)**	**3939 (97.2)**	**105 (89.7)**	**179 (87.8)**	204 (3.7)
Pregnancy Vegetarian Status					
Yes (at 32-40 weeks)	122 (7.6)	255 (4.3)	19 (11.1)	39 (14.0)	435 (5.4)
No (does not excl. previous)	1474 (92.4)	**5733 (95.7)[B]**	**153 (88.9)[F]**	**239 (86.0)[B,F]**	239 (94.6)
Maternal Smoking (3rd Trimester) n(%)	162 (10.1)	**1227 (20.6)[B]**	**40 (22.6)[B]**	**51 (18.6)[B]**	1480 (18.5)
Maternal Smoking (Postnatal) n(%)	190 (11.9)	**1463 (24.6)[B]**	**48 (27.1)[B]**	**60 (21.8)[B]**	1761 (22.0)
Maternal Race					
White	1571 (97.6)	5888(98.2)	170 (98.8)	267 (95.0)	7896 (98.0)
Non-white	39 (2.4)	106 (1.8)	2 (1.2)	**14 (5.0)[B,E,F]**	161 (2.0)
Pre-pregnancy BMI, mean (sd)	22.4 (3.2)	**23.1(3.9)[B,L]**	22.8 (3.6)	22.3 (3.6)	22.9 (3.8)
Maternal Age					
Maternal Age, mean (sd)	30.1 (4.5)	**27.9 (4.8)[B,E,L]**	29.4 (4.7)	29.8 (4.7)	28.0 (5.0)
Maternal Education					
O-Level, Voc, CSE, None	726 (46.3)	3799 (66.6)	80 (47.6)	103 (37.5)	4708 (61.0)
A-Level, Degree	841 (53.7)	**1904 (33.4)[B]**	**88 (52.4)[F]**	**172 (62.6)[B,F,E]**	3005 (39.0)
Marital Status					
Single	232 (14.9)	1203(20.6)	44 (26.2)	55 (20.2)	1534 (19.6)
Married	1327 (85.1)	**4629 (79.4)[B]**	**124 (73.8)[B]**	**217 (79.8)[B]**	6297 (80.4)
Maternal Age at Menarche (DERIVED)					
<12	246 (16.7)	**1084 (20.1)**	**38 (23.9)**	36 (14.0)	1404 (19.3)
12-<15	1032 (70.0)	**3614 (67.2)**	**101 (63.5)**	181 (70.2)	4928 (67.8)
15+	195 (13.2)	**683 (12.7)[B,L]**	**20 (12.6)[L]**	41 (15.9)	939 (12.9)
Mean Mat. Age at Men	12.9 (1.5)	12.8 (1.5)	**12.6 (1.5)[L]**	13.0 (1.4)	12.8 (1.5)
Maternal Parenting Score (18M)					
Mean (SD)	41.4 (4.2)	**40.6 (4.6)[B]**	40.7 (4.6)	41.0 (4.5)	40.7 (4.5)
Maternal Interaction Score, 42 M					
Mean (SD)	29.1 (4.6)	**28.5 (4.9)[B]**	28.3 (4.6)	28.7 (4.5)	28.6 (4.8)

Table A.2 continued

Partner Interaction Score, 42 M					
Mean (SD)	20.6 (7.4)	20.1 (8.2)	21.6 (7.5)	21.4 (7.2)	20.3 (8.0)
Day Care (15M)					
Yes (%)	93 (6.0)	322 (5.7)	**16 (9.5)^{F}**	23 (8.4)	454 (5.9)
Day Care (42 M)					
Yes (%)	515 (35.6)	1902 (73.4)	68 (43.3)	**106 (44.0)^{B,F}**	2591 (36.8)
Older Sibling					
Yes (%)	990 (63.9)	**2915 (51.6)^{B}**	**103 (60.2)^{F}**	**155 (58.5)^{F}**	4163 (54.5)
Older Brother					
Yes (%)	613 (40.0)	**1688 (29.9)^{B}**	**66 (38.6)^{F}**	**100 (37.7)^{F}**	2467 (32.3)
Older Sister					
Yes (%)	572 (36.9)	**1691 (29.9)^{B}**	55 (32.1)	**81 (30.5)^{B}**	2399 (31.4)

*Assessment performed before exclusions using 1 month data were applied, so sample sizes are slightly larger than that used in final analyses.

^{B,F,L,E} Indicate that a particular value is different from B (breast), F(formula), L (late soy), E (early soy) at the $p<0.05$ level.

Table A.3. Girls' BMI Z-scores for age, by feeding group

Feeding Group	n	Mean z-score (SD)	Breast Referent β (95% CI)	Early Formula Referent β (95% CI)
Age 7				
Primarily Breast	550	0.06 (0.97)	0.	-0.14 (-0.23, -0.05)
Early Formula	1733	0.21 (0.96)	0.14 (0.05, 0.23)	0.
Early Soy	47	0.18 (1.24)	0.11 (-0.17, 0.40)	-0.03 (-0.31, 0.25)
Late Soy	83	-0.06 (0.88)	-0.13 (-0.35, 0.09)	-0.27 (-0.48, -0.06)
Age 8				
Primarily Breast	460	0.20 (0.94)	0.	-0.16 (-0.25, -0.06)
Early Formula	1515	0.35 (0.90)	0.16 (0.06, 0.25)	0.
Early Soy	37	0.49 (0.86)	0.29 (-0.01, 0.60)	0.14 (-0.16, 0.43)
Late Soy	71	0.10 (0.82)	-0.10 (-0.33, 0.13)	-0.26 (-0.47, 0.04)
Age 9				
Primarily Breast	522	0.06 (1.03)	0.	-0.20 (-0.29, -0.10)
Early Formula	1746	0.25 (0.97)	0.20 (0.10, 0.29)	0.
Early Soy	45	0.31 (0.91)	0.26 (-0.04, 0.56)	0.06 (-0.23, 0.35)
Late Soy	79	-0.03 (1.01)	-0.09 (-0.31, 0.15)	-0.28 (-0.50, -0.05)

Table A.4: Distribution of covariate characteristics for lost to follow up subjects, across feeding group categories, and stratum-specific median follow up times.

	Primarily Breast	Early Formula	Early Soy	Late Soy	Median Follow Up Time (IQR)
Study Sample	124	513	8	13	140 (118 – 157)
Birth Weight					
≤ 2500g	2 (1.7)	8 (1.6)	0	0	140 (140 - 141)
>2500g	119 (98.4)	498 (98.4)	8 (100)	13 (100)	140 (117 – 157)
Missing	3	7	0	0	
Breast Feeding Duration					
None	0	152 (29.6)	2 (25)	1 (7.7)	129 (116 – 141)
<3 mo	0	222 (43.3)	2 (25)	2 (15.4)	140 (119 – 157)
3-<6 mo	0	95 (18.5)	3 (37.5)	0	140 (118 – 157)
6+ mo	124 (100)	44 (8.6)	1 (12.5)	10 (76.9)	140 (117 – 157)
Missing	0	0	0	0	
Mean (SD) (mo)	10.3 (3.0)	1.9 (2.7)	3.6 (5.3)	7.4 (5.0)	
Pre-pregnancy BMI					
≥ 25	18 (15.7)	107 (23.1)	1 (14.3)	2 (15.4)	132 (124 – 141)
< 25	97 (84.4)	356 (76.9)	6 (85.7)	11 (84.6)	140 (117 - 157)
Missing	9	50	1	0	
Mean (SD)	22.6 (3.1)	23.2 (4.1)	24.3 (6.5)	23.1 (8.1)	
Maternal Age at Menarche					
8-11	20 (17.5)	75 (17.0)	1 (16.7)	1 (9.1)	140 (118 – 141)
12-14	77 (67.5)	295 (66.7)	3 (50.0)	9 (91.8)	140 (118 – 157)
15+	17 (14.9)	72 (16.3)	2 (33.3)	1 (9.1)	140 (115 – 157)
Missing	10	71	2	2	
Mean (SD)(y)	12.9 (1.6)	13.0 (1.5)	13.0 (1.7)	13.1 (1.4)	
Prenatal Smoking					
Yes	13 (11.0)	108 (22.2)	3 (37.5)	4 (30.8)	137 (117 – 141)
No	105 (89.0)	379 (77.8)	5 (62.5)	9 (69.2)	140 (118 – 157)
Missing	6	26	0	0	
BMI for age, age 7-9					
>85th Percentile	19 (17.9)	119 (26.9)	1 (14.3)	3 (27.3)	128 (115 – 141)
≤ 85th Percentile	87 (82.1)	324 (73.1)	6 (85.7)	8 (72.7)	140 (128 – 157)
Missing	18	70	1	2	
Infant Health at 6 months					
Very Health/ Minor Problem	114 (98.3)	477 (96.8)	5 (71.4)	13 (100)	140 (117 – 157)
Sometimes ill/ Mostly unwell	2 (1.7)	16 (3.3)	2 (28.6)	0	134 (117 – 141)
Missing	8	20	1	0	
Maternal Age					
Mean (y)	29.3 (4.8)	27.7 (4.6)	30.9 (4.7)	29.5 (5.7)	

IQR: Inter-quartile Range

Discussion:

Table A.1 describes the distribution of ages at which early soy users initiated soy product use. The distribution across these ages suggests a large amount of heterogeneity in the age of introduction to soy. This may have implications with respect to conclusions that can be drawn from this exposure group.

Table A.2 assesses the association between covariates that are associated with various developmental outcomes measured in this analysis, and infant feeding group. The primary interest, given the aims of this study, are covariates that are associated both with the outcomes of interest and also with a differential distribution between early soy and early formula use. However, since all feeding groups were included in models, a covariate with a differential proportion in any of the feeding groups will be considered as a potential confounder of the whole model. Differences across feeding groups were assessed via pairwise chi-squared tests for categorical variables, and ANOVA with tukey adjustment for multiple comparisons for continuous variables.

Breastfeeding duration was variable across all feeding groups. All "formula" groups (early formula, early soy, late soy) had a higher prevalence of prenatal smoking and "single" marital status than the primarily breastfed.

Early Soy differed from other feeding groups in the proportion of male births, low birth weight, girls with BMI above the 85th percentile at age 8 (vs. late soy only), most infant health variables, infant crying, maternal perception of infant health, childhood soy consumption, maternal prenatal vegetarian status (vs. early formula only), day care attendance at 15 months (vs. formula only), and presence of any sibling or an older brother (vs. formula only). Mothers in the early soy group also had an earlier age at menarche, compared to the late soy group.

Early formula feeders had a lower level of maternal education and lower maternal age than all other feeding groups. They also had lower maternal parenting and interaction scores (vs. primarily breast), higher pre-pregnancy BMI (vs. primarily breast and late soy), lower birth weight (vs. primarily breast and late soy), and a higher proportion of girls with BMI above the 85th percentile, as

compared to the primarily breast fed. . They had higher proportions of childhood illnesses reported than the primarily breast fed group, but less than the early soy group.

The late soy group had a higher proportion of non-white subjects compared to all other feeding groups. Compared to the primarily breast and early formula groups, the late soy group had a higher proportion of vegetarian mothers. They also had a lower proportion of subjects in day care at 42 months, compared to the early soy group, and a higher proportion for the same variable, compared to the primarily breast fed group. Partner interaction scores at 42 months of age was not associated with any feeding group.

In Table A.3, it is show that there is no statistical difference in BMI z-score at ages 7, 8 and 9 between the early formula and early soy exposure groups among girls. Early formula fed girls do appear to be heavier than breast fed girls, in general.

In Table A.4, the stratum-specific distributions of lost to follow up subjects across feeding groups is broadly similar to the distributions of the sample population, as shown in Table A.2. These relationships were not evaluated for statistical difference, and it is noted that the number of lost to follow up subjects are small in the early and late soy groups, so it is difficult to draw inferences from their distributions as shown here.

The median follow up time contributed by lost to follow up subjects is 140 months (IQR: 118 – 157). The distribution of follow up times within strata of each covariate is similar in most cases. Exceptions include shorter follow up time for subjects that did not breastfeed, who had high pre-pregnancy BMI, who smoked prenatally, had high BMI for age, and were ill as infants.