

# Early-life soy exposure and age at menarche

Margaret A. Adgent<sup>a,c</sup>, Julie L. Daniels<sup>a</sup>, Walter J. Rogan<sup>c</sup>, Linda Adair<sup>a</sup>, Lloyd J. Edwards<sup>b</sup>, Daniel Westreich<sup>a,d</sup>, Mildred Maisonet<sup>e</sup> and Michele Marcus<sup>e</sup>

<sup>a</sup>Departments of Epidemiology and <sup>b</sup>Biostatistics, Gillings School of Global Public Health, UNC-Chapel Hill, Chapel Hill, <sup>c</sup>National Institute of Environmental Health Sciences, NIH, Research Triangle Park, <sup>d</sup>Department of Obstetrics and Gynecology and Duke Global Health Institute, Duke University, Durham, NC, and <sup>e</sup>Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA

## Summary

### Correspondence:

Dr Margaret Adgent,  
Epidemiology Branch,  
National Institute of  
Environmental Health  
Sciences, PO Box 12233,  
Mail Drop A3-05,  
Research Triangle Park,  
NC 27709-2233, USA.  
E-mail:  
adgentma@mail.nih.gov

Adgent MA, Daniels JL, Rogan WJ, Adair L, Edwards LJ, Westreich D, Maisonet M, Marcus M. Early-life soy exposure and age at menarche. *Paediatric and Perinatal Epidemiology* 2012; **26**: 163–175.

This study examines the timing of menarche in relation to infant-feeding methods, specifically addressing the potential effects of soy isoflavone exposure through soy-based infant feeding. Subjects were participants in the Avon Longitudinal Study of Parents and Children (ALSPAC). Mothers were enrolled during pregnancy and their children have been followed prospectively. Early-life feeding regimes, categorised as *primarily breast*, *early formula*, *early soy* and *late soy*, were defined using infant-feeding questionnaires administered during infancy. For this analysis, age at menarche was assessed using questionnaires administered approximately annually between ages 8 and 14.5. Eligible subjects were limited to term, singleton, White females. We used Kaplan–Meier survival curves and Cox proportional hazards models to assess age at menarche and risk of menarche over the study period.

The present analysis included 2920 girls. Approximately 2% of mothers reported that soy products were introduced into the infant diet at or before 4 months of age (early soy). The median age at menarche [interquartile range (IQR)] in the study sample was 153 months [144–163], approximately 12.8 years. The median age at menarche among early soy-fed girls was 149 months (12.4 years) [IQR, 140–159]. Compared with girls fed non-soy-based infant formula or milk (early formula), early soy-fed girls were at 25% higher risk of menarche throughout the course of follow-up (hazard ratio 1.25 [95% confidence interval 0.92, 1.71]). Our results also suggest that girls fed soy products in early infancy may have an increased risk of menarche specifically in early adolescence. These findings may be the observable manifestation of mild endocrine-disrupting effects of soy isoflavone exposure. However, our study is limited by few soy-exposed subjects and is not designed to assess biological mechanisms. Because soy formula use is common in some populations, this subtle association with menarche warrants more in-depth evaluation in future studies.

**Keywords:** ALSPAC, menarche, infant diet, soy milk, endocrine disruptor.

## Introduction

Some of the soy isoflavones, specifically genistein and daidzein, are weak oestrogenic compounds that are present in soy protein and various products derived from soybeans.<sup>1–3</sup> Demonstrating structural and functional similarity to 17 $\beta$ -oestradiol, soy isoflavones can bind to oestrogen receptors and can act as either oestro-

gen agonists or antagonists.<sup>4–6</sup> The biological activity of soy isoflavones has been demonstrated widely *in vitro* and in animal models,<sup>7–11</sup> as well as in adult humans.<sup>12,13</sup> Few studies, however, have addressed the effects of early-life soy protein exposure on long-term outcomes such as reproductive development and function.

The timing of puberty, or the transition from childhood to the adult reproductive stage, is a complex

process hypothesised to be determined, in part, by early-life endocrine system function.<sup>14</sup> Accordingly, exposure to soy isoflavones in early infancy may have lasting effects on later reproductive development. Animal studies have shown that pubertal markers, such as the age at vaginal opening, occur earlier in female rodents fed genistein or a soy-based diet during various time points in early development.<sup>15–18</sup> Early pubertal onset has also been observed in girls exposed to other endocrine-disrupting compounds pre- and postnatally.<sup>19,20</sup> Overall, however, epidemiological studies on the effects of early-life endocrine disruptor exposure on the timing of puberty have been equivocal.<sup>21–24</sup>

Infants exposed to soy-based products, such as soy-based infant formula (SBF), have not been studied well with respect to potentially endocrine-disrupting effects of early-life soy exposure. SBF is commonly used in the US, accounting for approximately 13–20% of the infant formula sold.<sup>25,26</sup> The urinary concentration of total isoflavones among infants exclusively fed SBF is approximately 500 times the concentration of those fed cows' milk formula,<sup>27</sup> and plasma isoflavone concentrations per bodyweight are an order of magnitude higher in SBF-fed infants than in adults consuming diets containing soy protein.<sup>28</sup> 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 early-life endocrine disruption.

## Methods

### Study sample

Avon Longitudinal Study of Parents and Children is an ongoing, prospective, longitudinal study that enrolled pregnant women residing in the Avon region of the UK, who were expected to deliver between 1 April 1991 and 31 December 1992. A total of 14 062 livebirths were recruited into the study during pregnancy. Of these, 13 978 (52% boys and 48% girls) were twins or singletons alive at 1 year. Mothers provided consent for participation. Ethical approval for the study was obtained from the ALSPAC Law and Ethics Committee and the local NHS Research Ethics Committees and the present analysis was approved by the Institutional

Review Board of the University of North Carolina at Chapel Hill.

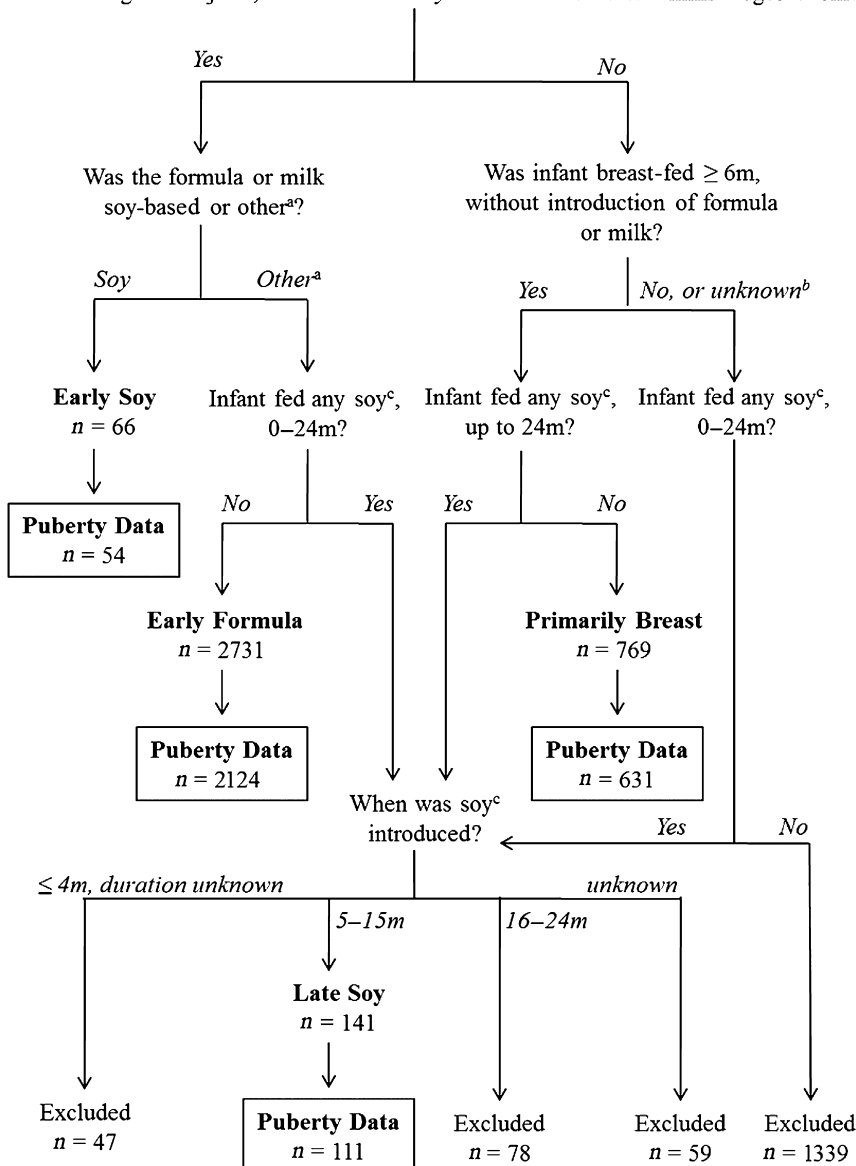
The present investigation was restricted to singleton, term, White females ( $n = 5230$ ), then further restricted to those for whom sufficient infant-feeding data were available and for whom at least one puberty questionnaire was completed between the ages of approximately 8 and 14.5 years of age ( $n = 2920$ ). We restricted the analysis to White girls because over 90% of the study sample was White, which limited our ability to control for confounding by race.

### Exposure assessment

Exposure assessment methods have been described previously.<sup>29</sup> In brief, mothers completed infant-feeding questionnaires at 1, 6, 15 and 24 months postpartum, in which they reported current breast-feeding habits, the age at which other milks or formulas were introduced into the child's diet (including formula/baby milk, soy milk, soy formula, follow-on milk, goats' milk, hypo-allergenic formula and cows' 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, as this questionnaire most thoroughly and proximally characterised feeding behaviours in early infancy. Responses from the 15-month questionnaire were used to characterise these early exposures if the 6-month questionnaire was missing or incomplete, as well as to characterise exposures later in infancy. Subjects were categorised into four mutually exclusive feeding groups: primarily breast-fed, early formula, early soy and late soy (Figure 1). *Primarily breast-fed infants* were breast fed until  $\geq 6$  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 had no reported 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 sustained use at 6 months of age. *Late soy-fed infants* were introduced to soy milk or soy formula any time between 5 months of age and 15 months of age. For the early formula and early soy exposure groups, the minimum exposure window of 4–6 months of age was defined as such in order to capture a period in

Of 5230 eligible subjects, was infant fed *any* formula or non-breast milk at ages 4–6m?



**Figure 1.** Study sample and exposure characterisation. Subjects were classified into four mutually exclusive feeding groups (*primarily breast, early formula, early soy, late soy*) based on age and duration of formula, milk and breastmilk feeding. Boxes contain the final study sample size by exposure group, for subjects with complete exposure and outcome data. <sup>a</sup>Other: formula/baby milk, goats' milk, hypo-allergenic formula, follow-on milk and cows' milk. <sup>b</sup>Unknown: data were insufficient to determine breast-feeding duration or milk/formula use up to 6 months. <sup>c</sup>Soy: soy milk or soy formula. m, months.

infancy where the diet was predominantly milk-based, and also to allow for sufficient sample size to be included in each exposure group. Subjects were excluded if feeding profiles were not sufficiently complete to estimate duration of a particular feeding method, or if the feeding profile did not fit in a previously described exposure category. Subjects who only reported soy use between 16 and 24 months were also excluded because it was assumed that exposure would be low compared with the earlier time periods when milk or formula comprised most of the diet. 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 and 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.<sup>30</sup> Questionnaires were completed by a caregiving

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. If multiple questionnaires contained discordant age at menarche responses for the same individual ( $n = 528$ , 18%), the response that was provided first in the series of Growing and Changing questionnaires was used as the girl's age at menarche. ALSPAC subjects were also invited to attend regular research clinic visits, where age at menarche, among other developmental characteristics, was assessed. For 138 (4.7%) subjects in the present analysis, missing questionnaire data on age at menarche were estimated from this clinic data.

Some subjects reported a menarche event, but did not report an age at menarche ( $n = 99$ ). For 60 of these subjects, age at menarche was estimated as the mid-point 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 one or more questionnaires were skipped between a negative and positive menarche response, an estimated age was not derived and this subject was not included in any analysis ( $n = 39$ ). In an alternative analysis, age at menarche was imputed for these 60 subjects as part of a larger multiple imputation (MI) exercise (described in the following section). Ages derived using the mid-point approach were included in the complete case analysis, while imputed values were used in MI models only.

### Analysis

All analyses were completed using SAS 9.1.3 and 9.2 (SAS Institute Inc., Cary, NC, USA). Chi-squared tests and  $t$ -tests ( $\alpha = 0.05$ ) were used to compare the distribution of covariates between included and excluded subjects, and between exposure groups. Time-to-event analyses were used to assess differences in age at menarche between exposure groups. 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* (LTF) (dropped out before the 14.5-year assessment), or *administratively censored* (completed the 14.5-year assessment). Crude Kaplan–Meier survival curves,

median age at menarche and interquartile range [IQR: 25th–75th percentile] estimates for the main exposure and all covariates were obtained using life table analysis (SAS PROC LIFETEST). The difference in Kaplan–Meier curves was evaluated using log-rank  $P$ -values ( $\alpha = 0.05$ ). We obtained adjusted Kaplan–Meier curves and median age at menarche by exposure group by calculating inverse probability-of-exposure weights,<sup>31,32</sup> using polytomous logistic regression, and then performing a weighted life table analysis.

Hazard ratios (HRs) for menarche were estimated using Cox proportional hazards modelling. The early formula group was used as the reference group in all models. Relative precision of estimated HRs was compared using confidence limit ratios, calculated as the upper 95% confidence limit divided by the lower 95% confidence limit. The proportional hazards assumption was assessed for all variables modelled, and continuous-time interaction terms were included for variables in violation of this assumption [pre-pregnancy body mass index (BMI) and maternal age at menarche]. Potential outliers were identified using deviance residuals ( $> \pm 3$ ) and  $dfbeta$  influence statistics ( $> 2/\sqrt{n}$ ), and were excluded only in explicitly identified sensitivity analyses. The HRs for our main exposure did not violate the proportional hazards assumption according to the formal SAS diagnostics. They were not, however, identical. To illustrate changes in the HRs over time, a series of average HRs for different periods of follow-up [11 years, 12 years, 13 years and complete duration of follow-up (~14.5 years)] were estimated.<sup>33</sup> Cox models were carried out using both complete case analysis and MI analysis. For the complete case analysis, only subjects with complete data on necessary covariates were modelled, which resulted in a loss of 18% of subjects ( $n = 537$ ). For the analyses including imputed data, values for missing outcome ( $n = 60$ ) or necessary covariates were estimated from available data on all adjustment variables, as well as birthweight, maternal education, maternal age at delivery, breast-feeding duration, marital status, child's BMI and completion history for the menarche question on the 'Growing and Changing' questionnaires (based on which questionnaires contained negative responses, which contained missing responses and which contained positive responses) using SAS PROC MI (five iterations). Cox models of imputed data were run and summarised in SAS PROC MIANALYZE.

Covariates were included as confounders if they were associated with infant-feeding method and age at

menarche or censoring in these data (via univariable association), or in relevant literature, but not hypothesised to be along the causal pathway between infant feeding and age at menarche. Variables examined included child's birthweight, breast-feeding duration, maternal report of milk allergy at 6 months, vegetarian diet/soy consumption in childhood, maternal perception of infant health, prenatal vegetarian diet, maternal age at menarche, pre-pregnancy BMI, prenatal smoking and maternal age and education at delivery. The final models were adjusted for pre-pregnancy BMI, maternal report of smoking in the last 2 months of pregnancy (yes/no) and maternal age at menarche (continuous years).

To investigate the hypothesis that childhood BMI may be along the causal pathway between infant-feeding method and age at menarche, we estimated the association between feeding group and BMI age-adjusted z-scores<sup>34</sup> using linear regression. Height and weight measurements were obtained through either clinic visit or self-report between the ages of 7 and 9 years. The proportion of overweight girls in each feeding group was also compared using chi-squared tests ( $\alpha = 0.05$ ), where overweight was defined as having a BMI greater than the 85th percentile for age at any time between ages 7 and 9 years.

To complement the analysis described above, log-binomial regression was used to estimate the risk of achieving early menarche, defined as <12 years of age (11.9 years corresponds to the 25th percentile for timing of menarche in non-Hispanic White females in the National Health and Nutrition Examination Survey, 1988–1994<sup>35</sup>). For censored subjects, those censored after age 12 were included in the reference group, while those censored before age 12 were excluded ( $n = 473$ , 16%).

### Sensitivity analyses

A set of sensitivity analyses were performed to assess whether proportional hazards model results were affected by informative censoring.<sup>36</sup> First, we assumed that all LTF subjects were at low risk of 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 LTF subjects were at high risk of reaching menarche in the study period, and follow-up times were modelled as events occurring at the time of dropout ('high risk'). Cox

models were repeated for each hypothetical scenario, with the resulting estimates representing extreme upper and lower bounds of the possible influence of informative censoring on our results, were it to occur non-differentially across exposure groups. In order to characterise potential differences in censoring across feeding groups, chi-squared tests, Fisher's exact tests and *t*-tests ( $\alpha = 0.05$ ) were used to compare characteristics of LTF subjects with subjects who reported events or were administratively censored.

### Results

Among 5230 eligible ALSPAC subjects, 2920 (56%) had sufficient infant-feeding and puberty data to be included in this analysis. Of the 2310 excluded girls, 1523 had incomplete infant-feeding profiles ( $n = 1137$ , 49%) or did not fit into a defined exposure category (e.g. introduced to soy between 16 and 24 months, introduced to formula at 5 months) ( $n = 386$ , 17%), 748 (32%) did not complete any 'Growing and Changing' questionnaire, and 39 (2%) did not report an age at menarche, as previously described. On average, excluded girls had lower birthweight, were more likely to be ill as infants, were exposed to prenatal tobacco smoke and born to younger, heavier mothers than those that remained in the final study sample (Table 1). Excluded girls were also more likely to have missing data on nearly all covariates examined. There were no differences 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, had shorter breast-feeding duration and were more likely to be classified in the early formula group, as compared with the included study sample.

Early soy-exposed girls were similar to early formula-exposed girls with respect to birthweight and childhood BMI, as well as maternal pre-pregnancy BMI, age at menarche, prenatal smoking status, marital status and prenatal vegetarian diet (Table 2). Early soy-exposed girls did experience longer breast-feeding duration than early formula-fed girls and were more likely to continue a vegetarian diet/soy consumption in childhood. Additionally, their mothers were more likely to be older at delivery, have higher education and have reported their infant to have had poor health or milk allergy. Late soy and primarily breast-fed girls both differed from early formula-fed girls with respect to breast-feeding duration, childhood BMI,

**Table 1.** Characteristics of eligible<sup>a</sup> Avon Longitudinal Study of Parents and Children (ALSPAC) study sample ( $n = 5230$ ), distinguished as those included in the present analysis ( $n = 2920$ ) and those excluded because of missing exposure or outcome data ( $n = 2310$ )

	Study sample included in present analysis $n$ (%)	Median age at menarche (months) [IQR <sup>b</sup> ]	Lost to follow up <sup>c</sup> $n$ (%)	Excluded $n$ (%)
Study sample	2920	153 [144–163]	666	2310
Feeding group				
Early formula	2124 (72.7)	153 [144–163]	513 (77.0)	607 (77.1)*
Early soy	54 (1.9)	149 [140–159]	8 (1.2)	12 (1.5)
Late soy	111 (3.8)	153 [147–159]	21 (3.2)	30 (3.8)
Primarily breast	631 (21.6)	154 [145–165]	124 (18.6)	138 (17.5)*
Missing	–	–	–	1523
Birthweight (g)				
$\leq 2500$	36 (1.2)	151 [146–160]	10 (1.5)	52 (2.3)*
$> 2500$	2843 (98.8)	153 [144–163]	646 (98.5)	2231 (97.7)*
Missing	41	153 [144–161]	10	27
Mean $\pm$ SD	3439 $\pm$ 436	–	3464 $\pm$ 460	3408 $\pm$ 464*
Breast-feeding duration (months)				
Mean $\pm$ SD	4.3 $\pm$ 4.6	–	3.7 $\pm$ 4.4	3.7 $\pm$ 4.4*
Missing	0	–	0	142
Pre-pregnancy body mass index				
$\geq 25$	538 (20.0)	150 [141–160]	128 (21.1)	433 (21.8)
$< 25$	2156 (80.0)	154 [145–164]	478 (78.9)	1552 (78.2)
Missing	226	154 [144–162]	60	325
Mean $\pm$ SD	22.8 $\pm$ 3.7	–	23.0 $\pm$ 4.0	23.1 $\pm$ 4.0*
Maternal age at menarche (years)				
8–11	493 (19.1)	147 [138–155]	98 (16.9)	376 (19.6)
12–14	1771 (68.6)	154 [145–163]	389 (67.1)	1276 (66.6)
15+	317 (12.3)	161 [151–171]	93 (16.0)	263 (13.7)
Missing	339	154 [145–165]	86	395
Mean $\pm$ SD	12.8 $\pm$ 1.5	–	12.8 $\pm$ 1.5	12.8 $\pm$ 1.6
Prenatal smoking				
Yes	423 (14.8)	152 [142–163]	128 (20.2)	419 (21.2)*
No	2433 (85.2)	154 [144–163]	506 (79.8)	1559 (78.8)*
Missing	64	153 [147–160]	32	332
Body mass index, age 7–9				
Ever $> 85$ th percentile	613 (24.8)	148 [137–155]	92 (19.1)	270 (24.4)
Never $> 85$ th percentile	1856 (75.2)	156 [147–165]	389 (80.9)	835 (75.6)
Missing	451	154 [143–162]	185	1205
Infant health at 6 months				
Healthy/minor problems	2781 (98.0)	153 [144–163]	617 (96.9)	1777 (96.4)*
Sometimes ill/unwell	56 (2.0)	153 [148–170]	20 (3.1)	66 (3.6)*
Missing	83	154 [144–160]	29	467
Maternal age at delivery (years)				
$\leq 30$	1942 (66.5)	154 [144–163]	477 (71.6)	1674 (72.5)*
$> 30$	978 (33.5)	153 [144–164]	189 (28.4)	636 (27.5)*
Missing	0	–	0	0
Mean $\pm$ SD	28.8 $\pm$ 4.5	–	28.1 $\pm$ 4.7	27.6 $\pm$ 5.0*

\* $P < 0.05$ , comparing proportion or mean of excluded with proportion or mean of included.

<sup>a</sup>Term, singleton, White females born in the ALSPAC cohort.

<sup>b</sup>Interquartile range: 25th–75th percentile.

<sup>c</sup>Subjects who completed at least one ‘Growing and Changing’ questionnaire, but did not report menarche and did not complete the ‘Growing and Changing’ questionnaire administered at 14.5 years; a subset of study sample ( $n = 2920$ ).

	Early formula	Early soy	Late soy	Primarily breast
<i>n</i>	2124	54	111	631
Mean ± SD				
Birthweight (g)	3426 ± 436	3479 ± 528	3527 ± 435*	3462 ± 422
Missing, <i>n</i>	30	1	1	9
Breast-feeding duration (months)	2.4 ± 3.0	4.1 ± 4.3*	7.4 ± 4.8*	10.5 ± 3.0*
Missing, <i>n</i>	0	0	0	0
Pre-pregnancy BMI	23.1 ± 3.9	22.9 ± 4.1	22.1 ± 3.8*	22.3 ± 3.0*
Missing, <i>n</i>	166	3	7	50
Maternal age at menarche (years)	12.8 ± 1.5	12.5 ± 1.3	13.1 ± 1.6	13.0 ± 1.5*
Missing, <i>n</i>	258	6	13	62
Maternal age at delivery (years)	28.3 ± 4.4	30.6 ± 5.0*	29.7 ± 4.4*	30.2 ± 4.4*
Missing, <i>n</i>	0	0	0	0
Age-adjusted BMI z-score, age 7	0.21 ± 0.96	0.24 ± 1.23	-0.02 ± 0.86*	0.04 ± 0.96*
Missing, <i>n</i>	513	9	19	106
Age-adjusted BMI z-score, age 8	0.35 ± 0.90	0.51 ± 0.86	0.13 ± 0.78*	0.17 ± 0.93*
Missing, <i>n</i>	638	18	32	157
Age-adjusted BMI z-score, age 9	0.24 ± 0.96	0.47 ± 0.92	0.01 ± 0.96	-0.01 ± 1.02*
Missing, <i>n</i>	919	19	48	220
<i>n</i> (%)				
Prenatal smoking				
Yes	361 (17)	9 (17)	10 (9)*	43 (7)*
No	1712 (83)	45 (83)	99 (90)	577 (93)
Missing, <i>n</i>	51	0	2	11
Perceived infant health				
Healthy/minor problems	2038 (98)	43 (86)*	105 (99)	595 (99)
Sometimes ill/unwell	39 (2)	7 (14)	1 (1)	9 (1)
Missing, <i>n</i>	47	4	5	27
Marital status				
Single	348 (17)	10 (20)	13 (12)	71 (12)*
Married	1705 (83)	41 (80)	96 (88)	540 (88)
Missing, <i>n</i>	71	3	2	20
Maternal education <sup>a</sup>				
O-Level, Voc, CSE, None	1300 (64)	23 (43)*	40 (36)*	264 (43)*
A-Level, Degree	744 (36)	30 (57)	70 (64)	349 (57)
Missing, <i>n</i>	80	1	1	18
BMI, age 7-9				
Ever >85th percentile	465 (26)	17 (35)	13 (13)*	118 (21)*
Never >85th percentile	1303 (74)	32 (65)	86 (87)	435 (79)
Missing, <i>n</i>	356	5	12	78
Vegetarian/soy diet in childhood				
Yes	44 (3)	5 (11)*	14 (14)*	33 (6)*
No	1718 (97)	39 (89)	85 (86)	519 (94)
Missing, <i>n</i>	362	10	12	79
Milk allergy at 6 months				
Yes	21 (1)	26 (50)*	16 (15)*	6 (1)
No	2063 (99)	26 (50)	90 (85)	601 (99)
Missing, <i>n</i>	40	2	5	24
Prenatal vegetarian diet				
Yes	97 (5)	5 (9)	16 (15)*	55 (9)*
No	2005 (95)	49 (91)	93 (85)	564 (91)
Missing, <i>n</i>	22	0	2	12

**Table 2.** Distribution of characteristics [mean ± SD for continuous variables, *n* (%) for categorical variables] by feeding group

\* $P < 0.05$ , compared with early formula feeding group.

<sup>a</sup>Levels of education, ranked lowest to highest: None, CSE (Certification of Secondary Education), Voc (Vocational), O-Level (Ordinary Level), A-Level (Advanced Level) and Degree (University Degree). BMI, body mass index.

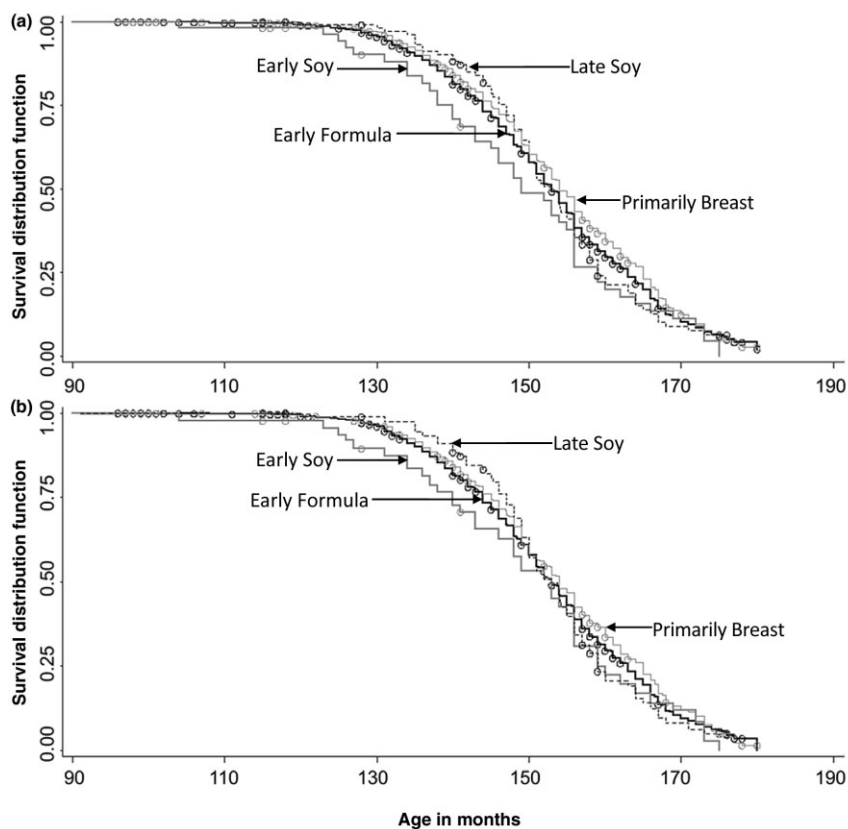
childhood vegetarian diet/soy consumption, and maternal pre-pregnancy BMI, age at delivery, education, prenatal smoking status and prenatal vegetarian diet.

The median age at menarche for the study sample was 153 months (12.8 years) [IQR, 144–163]. The median follow-up time contributed by those LTF ( $n = 666$ , 22.8%) was 140 months (11.7 years) [IQR, 117–157]. Nearly 2% of girls in the final study sample ( $n = 54$ ) 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. Three per cent of all girls were administratively censored, suggesting that they experienced menarche at some point after age 14.5 years. None of the administratively censored were early soy-fed.

In crude analyses, there was no difference in age at menarche between strata of birthweight, maternal perception of infant health, or maternal age at delivery (Table 1), or of marital status, maternal education, prenatal vegetarian diet, childhood vegetarian diet/soy consumption, or maternal report of milk allergy at 6 months (not shown). Earlier age at menarche was

observed among girls with high BMI and among girls whose mothers had a high pre-pregnancy BMI or a young age at menarche. Early censoring also occurred among girls with high BMI, and among those whose mothers had high pre-pregnancy BMI (respective median [IQR] follow-up times among LTF subjects: 128 [115–141] and 132 [124–141]). Age at menarche was slightly earlier among girls exposed to prenatal tobacco smoke (log-rank  $P = 0.06$ ), and was later among girls with mothers who had an older age at menarche (Table 1). These results are similar to previously published associations in this cohort.<sup>37</sup>

Across infant-feeding groups, the crude median age at 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]). The unadjusted Kaplan–Meier curves were not statistically different over time for any feeding group, compared with the early formula reference (log-rank  $P_{\text{early soy}} = 0.12$ ,  $P_{\text{late soy}} = 0.82$ ,  $P_{\text{primarily breast}} = 0.07$ ), and following adjustment, the median age for all groups shifted to 153 months, with the exception of primarily breast feeding, which remained at 154 months. However, prior to



**Figure 2.** Unadjusted (a) and confounding-adjusted (b) Kaplan–Meier survival curves for each feeding group are shown. Censored observations are indicated by  $\circ$ . Survival for each feeding group is indicated as follows: — early formula; — early soy; .... late soy; — primarily breast. Confounding-adjusted curves are adjusted for pre-pregnancy body mass index, maternal age at menarche and prenatal smoking using inverse probability of treatment weights estimated with polytomous logistic regression.



Length of follow-up	Feeding group	Events	HR <sup>a</sup>	95% CI	CLR
11 years	Early formula	112	1.00	Reference	–
	Early soy	6	2.06	[0.91, 4.69]	5.45
	Late soy	2	0.41	[0.10, 1.65]	16.4
	Primarily breast	27	0.92	[0.60, 1.41]	2.35
12 years	Early formula	416	1.00	Reference	–
	Early soy	16	1.57	[0.95, 2.59]	2.72
	Late soy	12	0.58	[0.33, 1.03]	3.12
	Primarily breast	115	0.97	[0.78, 1.19]	1.52
13 years	Early formula	905	1.00	Reference	–
	Early soy	30	1.33	[0.92, 1.91]	2.07
	Late soy	49	1.01	[0.76, 1.35]	1.78
	Primarily breast	265	0.95	[0.82, 1.09]	1.32
Complete follow-up	Early formula	1278	1.00	Reference	–
	Early soy	41	1.25	[0.92, 1.71]	1.86
	Late soy	69	1.01	[0.79, 1.29]	1.63
	Primarily breast	408	0.95	[0.84, 1.06]	1.26

<sup>a</sup>Adjusted for pre-pregnancy body mass index, prenatal smoking and maternal age at menarche.

CI, confidence interval; CLR, confidence limit ratio.

approximately 150 months, unadjusted and adjusted Kaplan–Meier curves (Figure 2) suggest more menarche events in early adolescence in the early soy group, as compared with all other feeding groups, despite nearly equivalent median ages.

Averaged across the full length of follow-up, the hazard of reaching menarche among girls in the early soy feeding group was approximately 1.25 times [95% confidence interval (CI) 0.92, 1.71] that among girls in the early formula feeding group, adjusted for maternal pre-pregnancy BMI, prenatal smoking and maternal age at menarche (Table 3). A modest reduction in hazard was associated with breast feeding (HR: 0.95 [95% CI 0.84, 1.06]), while no association was observed for late soy feeding, over the full length of follow-up. Associations were similar in crude analysis of all eligible subjects, and in crude complete case analysis, but were slightly attenuated for early soy exposures when identified outliers were removed ( $n = 14$ ) (HR: 1.20 [95% CI 0.87, 1.64]). In addition, results were similar for all exposure groups when MI was implemented to account for missing covariates (MI HR<sub>early soy</sub>: 1.22 [95% CI 0.88, 1.70]; MI HR<sub>late soy</sub>: 1.06 [95% CI 0.85, 1.32]; MI HR<sub>primarily breast</sub>: 0.97 [95% CI 0.87, 1.08]). As childhood BMI

z-scores and overweight status were not significantly different between early soy and early formula feeding groups, the association between early soy feeding and menarche was not likely to be mediated by childhood BMI in this analysis.

To further characterise the variability of hazards over time, a series of average HRs for increasingly longer periods of follow-up are also shown in Table 3. For early soy exposure, these serial HR estimates suggest an elevated risk of menarche in the earlier stages of adolescence that decreased with increasing length of follow-up. Late soy exposure was associated with a decreased risk of menarche up to approximately 12 years, and subsequently reached the null with increasing length of follow-up. A small decrease in risk was associated with primarily breast feeding at each follow-up interval. These findings are similarly conveyed in adjusted Kaplan–Meier curves (Figure 2b).

Loss to follow up in the early soy group tended to occur earlier in the study than in the other feeding groups (median LTF for early soy = 123 months [IQR: 99–141]), yet the overall proportion of LTF in the early soy group ( $n = 8$ , 15%) was less than other feeding

**Table 3.** Hazard ratios (HRs) [95% CI] for varying lengths of follow-up

groups (e.g.  $n = 513$ , 24% in early formula). The possible effect of LTF subjects on reported effect estimates, were they to be non-differentially informative as opposed to random, was assessed in sensitivity analyses. If LTF subjects were all 'high risk' for menarche, the HR estimate associated with early soy exposure for the full length of follow-up decreased to 1.07 [95% CI 0.80, 1.44], whereas if these subjects were all 'low risk,' the HR increased to 1.50 [95% CI 1.08, 2.00]. However, while LTF subjects (Table 1), particularly those in the early formula feeding group (Table S1), do appear to have older maternal age at menarche and a higher proportion of prenatal smoking than other subjects, these characteristics would be likely to induce bias in opposing directions, suggesting that LTF subjects are neither high nor low risk. Of note, early formula LTF subjects also demonstrated a lower proportion of childhood overweight than other early formula subjects. However, the proportion missing for this variable was also substantially greater for LTF subjects ( $n = 141$ , 27% vs.  $n = 215$ , 13%), thus limiting accurate characterisation of the differences.

Early menarche (<12 years) was reported for 599 (20%) of 2920 subjects. In a complete case analysis ( $n = 2028$ ), the adjusted risk ratios for early menarche [95% CIs] were 1.53 [1.07, 2.18], 0.64 [0.37, 1.11] and 0.98 [0.81, 1.17] for early soy, late soy and primarily breast feeding exposures, respectively, in comparison with the early formula reference group.

## Discussion

In this study, over the full course of follow-up, early-life exposure to soy products was associated with a small, imprecisely estimated increased risk of menarche, while the median age at menarche did not differ substantially between any of the feeding groups. However, variations in risk over time were evident in multiple approaches to analysis, including survival curves, serial HRs, and the estimation of early menarche risk, that suggest that the risk of menarche associated with early soy exposure was greatest in the early stages of adolescence. This change in risk over time is likely to reflect both statistical and biological phenomena. Regardless of exposure status, nearly all girls reach menarche at later stages of follow-up. Accordingly, the risk ratio of menarche for time points later in adolescence is expected to approach 1.0, and therefore the HR averaged over the full course of follow-up is lower than it is for shorter periods of

follow-up. In addition, it is possible that biological susceptibility to soy among a sensitive subset of exposed girls would result in elevated hazards in early, but not late adolescence, as we observed here.

An association between early soy exposure and early menarche is biologically plausible. 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 axis regulates hormone signalling 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.<sup>38</sup> Oestrogen receptors are present in the hypothalamus,<sup>39,40</sup> suggesting that this region in particular may be susceptible to isoflavone binding. Accordingly, rodent models have demonstrated that neonatal exposure to oestrogenic compounds has altered hypothalamic characteristics and function.<sup>17</sup> In addition, recent evidence has suggested that manganese, which is found in marginally higher concentrations in soy formula than in cows' milk formula,<sup>41</sup> may be associated with hypothalamic function and early pubertal onset.<sup>42,43</sup> Although this mechanism was not considered in our original study hypothesis, it may warrant further exploration.

In general, we observed that early soy-fed infants and their mothers were not readily distinguishable from those in the early formula reference group with respect to demographic and other descriptive characteristics. Importantly, the few characteristics by which these feeding groups differed, including maternal age, education, breast-feeding duration, child health and diet, were not shown to be associated with age at menarche and therefore should not bias our findings. For the remaining characteristics evaluated, early formula and early soy exposure groups were similar.

A mild and particularly unstable decrease in hazard of menarche in early adolescence within the late soy feeding group was also observed. That early soy and late soy feeding were not similarly associated with age at menarche suggests that dose and timing of soy exposure may be important for the induction of developmental effects. The exact amount of soy intake in our sample is unknown, so a true dose–response relationship could not be assessed. However, as the diet becomes increasingly diverse with increasing age, exposure to soy formula is presumably higher during

early infancy as compared with later, possibly influencing the observed associations. In addition, as early formula and late soy feeders were found to consistently differ on most demographic and life style characteristics, there is increased potential for residual confounding in early formula and late soy comparisons, relative to early formula and early soy comparisons. Thus, late soy exposure findings should be interpreted with particular caution.

Strom *et al.*<sup>24</sup> reported on various endocrine-sensitive outcomes among adults who had participated in cows' milk formula and soy formula clinical trials in infancy, and observed no difference in recalled 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 survival analysis, we did observe that early soy-exposed girls were at higher risk of menarche in early adolescence. This finding reflects a subtle characteristic of the association between soy infant feeding and menarche that was not captured in previous studies.

The longitudinal nature of our study was advantageous in that it allowed for age at menarche to be repeatedly assessed and estimated within approximately 1 year of the menarche event until age 14.5. There is potential, however, for selection bias in our final study sample, particularly if there is a characteristic that was associated with infant-feeding method, age at menarche and inclusion in our final study sample. As stated previously, excluded girls had lower birthweight, shorter breast-feeding duration, were more likely to be ill as infants, exposed to prenatal tobacco smoke and born to younger, heavier mothers. However, none of these characteristics were differentially associated with early soy and early formula feeding, and with age at menarche, so the associations observed among early soy-fed girls are not likely to be biased due to selection in this way. However, it is possible that other unmeasured factors may have influenced participation in the study and may have allowed for some selection bias to persist.

Given the distribution of predictive characteristics such as maternal age at menarche, pre-pregnancy BMI and prenatal smoking among the LTF subjects (Table 1, Table S1), there is no clear indication that informative censoring is present in this analysis that would result in biased estimates. For early formula-fed girls specifically, the characteristics for which differences were

observed between LTF and other subjects would not be likely to induce bias in a consistent direction (i.e. the net bias may be minimal given that estimates may be both over- and underestimated because of a higher proportion of prenatal smoking and older maternal age at menarche, respectively, in the LTF). In addition, we adjusted for maternal age at menarche and prenatal smoking to control for bias related to LTF. Still, sensitivity analyses demonstrated that our results do have the potential to be biased by informative censoring, were this type of censoring to be present, non-differentially by exposure status. Of particular concern, the 'high risk' model yielded estimates for early soy exposure that were nearly null. While this model is an unrealistic treatment of our data,<sup>36</sup> it is important to consider that our estimates may be overestimating the true association between early soy feeding and age at menarche due to censoring, and replication of our findings is imperative before meaningful conclusions can be drawn.

This analysis was conducted using a large, longitudinal cohort study that is generalisable to the UK.<sup>44</sup> 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, although ALSPAC is large, we had only 54 girls in the early soy group, which limited the power and precision of our analysis. Our analysis was also restricted to White girls, and our final study sample differed from excluded ALSPAC participants on multiple characteristics, such as prenatal tobacco smoke exposure and maternal age at delivery. Lastly, it is unclear whether the infant formula products used in the UK at this time were substantially different from those used, for example, in the US in the present day. Therefore, the generalisability of these results may be limited.

In summary, our study yielded preliminary findings that exposure to soy products in early infancy may contribute to a small increase in risk of menarche in early adolescence. This study contributes to a growing literature on the potential for endocrine disruptors to affect pubertal onset, making use of a uniquely wide exposure contrast between those who were and were not fed soy products in infancy. However, additional research is needed given the limitations of our study, such as low prevalence of exposure and large loss to follow up in our study sample. Future studies in populations with higher soy formula prevalence, such as in the US, will be informative.

## Acknowledgements

We are extremely grateful to all the families who took part in this 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 currently provide core support for ALSPAC. This publication is the work of the authors and they will serve as guarantors for the contents of this paper. The research was specifically funded by NICHD/NIH (T32HD052468-01A2), 2008–2013, the Intramural Research Program of the NIH, and Centers for Disease Control and Prevention. The authors declare they have no actual or potential competing financial interests.

## References

- Rostagno MA, Palma M, Barroso CG. Pressurized liquid extraction of isoflavones from soybeans. *Analytica Chimica Acta* 2004; **522**:169–177.
- Rostagno MA, Palma M, Barroso CG. Fast analysis of soy isoflavones by high-performance liquid chromatography with monolithic columns. *Analytica Chimica Acta* 2007; **582**:243–249.
- Kuiper GG, Lemmen JG, Carlsson B, Corton JC, Safe SH, van der Saag PT, et al. Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinology* 1998; **139**:4252–4263.
- 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. *The Journal of Steroid Biochemistry and Molecular Biology* 2006; **101**:246–253.
- 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. *The Journal of Nutrition* 1997; **127**:263–269.
- Dixon RA, Ferreira D. Genistein. *Phytochemistry* 2002; **60**:205–211.
- Dang ZC. Dose-dependent effects of soy phyto-oestrogen genistein on adipocytes: mechanisms of action. *Obesity Reviews* 2009; **10**:342–349.
- Akiyama T, Ishida J, Nakagawa S, Ogawara H, Watanabe S, Itoh N, et al. Genistein, a specific inhibitor or tyrosine-specific protein kinases. *The Journal of Biological Chemistry* 1987; **262**:5592–5595.
- 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. *The Journal of Biological Chemistry* 2003; **278**:962–967.
- 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: Clinical and Experimental* 2008; **57**:S24–S31.
- 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. *Nutrition Research* 2008; **28**:368–376.
- 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. *Human Reproduction Update* 2009; **15**:423–440.
- Cassidy A, Bingham S, Setchell KD. Biological effects of a diet of soy protein rich in isoflavones on the menstrual cycle of premenopausal women. *The American Journal of Clinical Nutrition* 1994; **60**:333–340.
- Mouritsen A, Aksglaede L, Sorensen K, Mogensen SS, Leffers H, Main KM, et al. Hypothesis: exposure to endocrine-disrupting chemicals may interfere with timing of puberty. *International Journal of Andrology* 2010; **33**:346–359.
- 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. *Toxicological Sciences* 1999; **51**:236–244.
- 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. *Toxicological Sciences* 2003; **71**:74–83.
- 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; **29**:988–997.
- Gallo D, Cantelmo F, Distefano M, Ferlini C, Zannoni GF, Riva A, et al. Reproductive effects of dietary soy in female Wistar rats. *Food and Chemical Toxicology* 1999; **37**:493–502.
- 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; **11**:641–647.
- Vasiliu O, Muttineni J, Karmaus W. In utero exposure to organochlorines and age at menarche. *Human Reproduction* 2004; **19**:1506–1512.
- Gladden BC, Ragan NB, Rogan WJ. Pubertal growth and development and prenatal and lactational exposure to polychlorinated biphenyls and dichlorodiphenyl dichloroethene. *Journal of Pediatrics* 2000; **136**:490–496.
- Warner M, Samuels S, Mocarrelli P, Gerthoux PM, Needham L, Patterson DG, et al. Serum dioxin concentrations and age at menarche. *Environmental Health Perspectives* 2004; **112**:1289–1292.
- Leijds MM, Koppe JG, Olie K, van Aalderen WM, Voogt P, Vulsma T, et al. Delayed initiation of breast development in

- girls with higher prenatal dioxin exposure; a longitudinal cohort study. *Chemosphere* 2008; **73**:999–1004.
- 24 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; **286**:807–814.
- 25 National Toxicology Program. *NTP Brief of Soy Infant Formula*. Washington, D.C.: National Institutes of Health, US Department of Health and Human Services, 2010. [http://ntp.niehs.nih.gov/ntp/ohat/genistein-soy/SoyFormulaUpdt/FinalNTPBriefSoyFormula\\_9\\_20\\_2010.pdf](http://ntp.niehs.nih.gov/ntp/ohat/genistein-soy/SoyFormulaUpdt/FinalNTPBriefSoyFormula_9_20_2010.pdf) [last accessed November 2011].
- 26 Bhatia J, Greer F, American Academy of Pediatrics Committee on Nutrition. Use of soy protein-based formulas in infant feeding. *Pediatrics* 2008; **121**:1062–1068.
- 27 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. *Journal of Exposure Science & Environmental Epidemiology* 2009; **19**:223–234.
- 28 Setchell KD, Zimmer-Nechemias L, Cai J, Heubi JE. Exposure of infants to phyto-oestrogens from soy-based infant formula. *Lancet* 1997; **350**:23–27.
- 29 Adgent M, Daniels JL, Edwards L, Siega-Riz AM, Rogan WJ. Early life soy exposure and gender-role play behavior in children. *Environmental Health Perspectives* 2011; DOI: 10.1289/ehp.1103579.
- 30 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. *Paediatric and Perinatal Epidemiology* 2009; **23**:492–504.
- 31 Hernan MA, Robins JM. Estimating causal effects from epidemiological data. *Journal of Epidemiology and Community Health* 2006; **60**:578–586.
- 32 Cole SR, Hernan MA. Adjusted survival curves with inverse probability weights. *Computer Methods and Programs in Biomedicine* 2004; **75**:45–49.
- 33 Hernan MA. The hazards of hazard ratios. *Epidemiology* 2010; **21**:13–15.
- 34 Centers for Disease Control and Prevention. A SAS Program for the CDC Growth Charts. Updated 2009. <http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm> [last accessed April 2011].
- 35 Chumlea WC, Schubert CM, Roche AF, Kulin HE, Lee PA, Himes JH, *et al.* Age at menarche and racial comparisons in US girls. *Pediatrics* 2003; **111**:110–113.
- 36 Allison PD. Heterogeneity, repeated events and other topics. In: *Survival Analysis Using SAS®: A Practical Guide, Second Edition*. Editor: McDaniel G. Cary, NC: SAS Institute Inc., 2010; pp. 283–287.
- 37 Maisonet M, Christensen KY, Rubin C, Holmes A, Flanders WD, Heron J, *et al.* Role of prenatal characteristics and early growth on pubertal attainment of British girls. *Pediatrics* 2010; **126**:E591–E600.
- 38 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; **121** (Suppl. 3):S192–S207.
- 39 Shughrue PJ, Lane MV, Merchenthaler I. Comparative distribution of estrogen receptor-alpha and -beta mRNA in the rat central nervous system. *The Journal of Comparative Neurology* 1997; **388**:507–525.
- 40 Perez SE, Chen EY, Mufson EJ. Distribution of estrogen receptor alpha and beta immunoreactive profiles in the postnatal rat brain. *Brain Research. Developmental Brain Research* 2003; **145**:117–139.
- 41 Cockell KA, Bonacci G, Belonje B. Manganese content of soy or rice beverages is high in comparison to infant formulas. *Journal of the American College of Nutrition* 2004; **23**:124–130.
- 42 Lee B, Hiney JK, Pine MD, Srivastava VK, Dees WL. Manganese stimulates luteinizing hormone releasing hormone secretion in prepubertal female rats: hypothalamic site and mechanism of action. *Journal of Physiology* 2007; **578**:765–772.
- 43 Hiney JK, Srivastava VK, Dees WL. Manganese induces IGF-1 and cyclooxygenase-2 gene expressions in the basal hypothalamus during prepubertal female development. *Toxicological Sciences* 2011; **121**:389–396.
- 44 Golding J, Pembrey M, Jones R, Team AS. ALSPAC – the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatric and Perinatal Epidemiology* 2001; **15**:74–87.

## Supporting information

Additional Supporting Information may be found in the online version of this article:

**Table S1.** Comparison of characteristics of lost to follow up (LTF) subjects and subjects reporting events or administrative censoring (AC), within feeding group.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.