

A matched case-control study of breastfeeding and other factors in pediatric Crohn's disease

Theresa A. Mikhailov^{1*}, Melissa A. Christensen², Sylvia E. Furner³

1 Department of Pediatrics, Division of Critical Care, Medical College of Wisconsin, Milwaukee, WI, USA **2** Department of Surgery Children's Hospital of Wisconsin, Milwaukee, WI, USA **3** Department of Epidemiology and Biostatistics, School of Public Health, University of Illinois Chicago, Chicago, IL, USA

Abstract

Background: The incidence of inflammatory bowel disease (IBD) in children and adolescents has increased over the last 40 years. This increase is due primarily to the rise in Crohn's disease (CD). The pathogenesis of CD remains uncertain but is thought to involve both environmental and genetic factors. We conducted this matched case-control of children diagnosed with CD to determine if CD in children is associated with breastfeeding in infancy.

Methods: We matched sibling controls to unrelated cases by age and sex. We obtained demographic and clinical data from the medical records and questionnaires. We analyzed data by McNemar's test, t-test, the non-zero correlation test, and conditional logistic regression analysis.

Results: The association between breastfeeding and CD was protective but not significant [$\psi = 0.63$ (0.31–1.30) $n=152$, McNemar's $\chi^2 = 1.58$, $p=0.21$]. There was no significant trend in development of CD based on duration of breastfeeding. Breastfeeding [dichotomous $\psi = 0.61$ (0.27–1.38) $m=76$ pairs] and [ordinal $\psi = 0.80$ (0.27–2.41), 0.40 (0.11–1.43), 0.62 (0.24–1.58) for < 3 months, 3–6 months, and > 6 months, respectively, vs. none] was protective in the development of CD controlling for family history of CD and of ulcerative colitis (UC), diarrhea, hospitalization, and any illness other than diarrhea in infancy, maternal age, and any smoke exposure. Any smoke exposure was a statistically significant risk factor for CD with an odds ratio greater than 2.

Conclusion: We found no association between breastfeeding in infancy and CD in children but a consistent association between any smoke exposure and CD in children.

Citation: Mikhailov TA, Christensen A, Furner SE (2017) A matched case-control study of breastfeeding and other factors in pediatric Crohn's disease. *Adv Pediatr Res* 4:4. doi:10.12715/apr.2017.4.4

Received: November 17, 2016; **Accepted:** March 1, 2017; **Published:** April 1, 2017

Copyright: © 2017 Mikhailov et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Competing interests: The authors have declared that no competing interests exist.

Sources of funding: This study was supported in part by the Medical College of Wisconsin's Digestive Disease Center, Grant Number 3303359.

* Email: tmikhail@mcw.edu

Introduction

Inflammatory bowel disease (IBD), which includes Crohn's disease (CD) and ulcerative colitis (UC), is distributed unevenly within populations and throughout the world. Onset can occur at any age but incidence is highest in adolescence and early

adulthood [1, 2]. Several studies have suggested that the incidence of IBD in children and adolescents has significantly increased over the last 40 years and that the increase is due primarily to an increase in the incidence of CD [3-6]. The first prospective, population-based North American study of the incidence of IBD in childhood was conducted in

Wisconsin from 2000 to 2001. The reported annual incidence of IBD in children younger than 18 years was 7.05 per 100,000 person-years, the highest reported incidence in children to date [7]. In this study, the annual incidence of CD was 4.56 per 100,000 person-years for children younger than 18 years, and of UC was 2.14 per 100,000 person-years for children younger than 18 years.

Despite intense research, the pathogenesis of IBD has not been defined. Numerous studies have identified genes which may increase the risk of having IBD [8, 9]; however, studies in twins and families with IBD suggest that genes alone cannot account for the risk of developing IBD [10-14]. Moreover, studies in urban versus rural, immigrant, and northern versus southern geographic populations demonstrate changes in risk for IBD that implicate environmental influences [15, 16]. Recent reviews describe the development of IBD as a complex combination of genetic and environmental factors but emphasize the importance of neonatal and early childhood environmental factors in altering the gut microbiome, thereby altering gastrointestinal immune development [17, 18].

A number of environmental factors have been considered in the development of IBD. These include smoking, oral contraceptive pill (OC) use, diet, preceding infections, antibiotic use, and breastfeeding [18-21]. Previous studies have failed to clearly define the relationship between breastfeeding and the development of CD or UC [20, 22, 23]. This may be due to study design issues such as poor definition of cases, poor selection of controls, inadequate power, and systematic bias [24]. Breastfeeding has been investigated as a factor influencing the development of many chronic diseases in children [25]. Human breast milk contains many substances that may influence the growth and development as well as the function of the gastrointestinal tract [26]. Some of these factors may have age-dependent effects. Since it is posited that CD pathogenesis is a complex interaction between genetic susceptibility factors, priming by enteric microflora, and immune-mediated tissue injury, it is quite plausible that breastfeeding in infancy may have a protective effect on the development of CD, at least in childhood, and possibly throughout life [27]. We hypothesized that breastfeeding has a protective effect on the development of CD in children. We conducted a

matched case-control study to determine if children who were breastfed in infancy were less likely to develop CD than children who were not breastfed and to determine if other environmental factors altered the risk of developing CD in childhood.

Methods

Subject selection

This study was approved by the Institutional Review Board of the Children's Hospital of Wisconsin. We conducted this as a matched case-control study. We identified prevalent cases from the population of individuals followed for CD at the Children's Hospital of Wisconsin. Individuals were considered eligible for participation if they were diagnosed with CD on or after January 1, 2000 at less than 18 years of age based on standard diagnostic criteria (clinical, radiological, endoscopic, and histological). Individuals were not eligible to participate if they had UC or an indeterminate form of IBD. To ensure controls with a similar risk of disease to that of cases, we recruited unaffected siblings of cases to participate as control subjects. For the purposes of this study, siblings had to have the same mother and father as the case sibling to be eligible for inclusion as controls. This was determined by simple questioning and was not confirmed by laboratory testing. Siblings were not eligible to participate as control subjects if they had CD. Only one affected sibling from any given family was eligible to participate as a case. If there were two or more members of the same family that met the standard diagnostic criteria for CD, we included the first eligible sibling who agreed to participate in the study. We obtained informed consent from the subject if 18 years of age or older and from the parent (or legal guardian) of all other subjects, and assent from all subjects over 7 years of age.

We obtained demographic and clinical data from the medical records. We obtained additional demographic data as well as early childhood history by questionnaire (completed by subject's parent or other individual familiar with the subject's early childhood). Specific questions pertained to duration of breastfeeding, age at introduction to formula, age at introduction to solid foods, and history of

infectious illnesses in infancy (and prior to onset of CD). Other pertinent data collected included maternal age, maternal smoking and passive smoking exposure (for the child and for the mother during pregnancy), and maternal use of OC before and during pregnancy with the subject. Additional data collected included family history of CD or UC, number of siblings, and history of antibiotic use in the years prior to diagnosis. We asked adolescents and young adults to complete a separate questionnaire regarding smoking history, antibiotic use, and for female subjects, OC use and gynecological history.

We derived a variable (any illness) that was determined by an affirmative answer to any of the infectious illnesses in infancy. We also derived the smoking variables – any smoke exposure, any passive smoke exposure, and any passive *in utero* smoke exposure. We defined any smoke exposure by an affirmative answer to any of the six smoking-related variables. We defined any passive smoke exposure by an affirmative answer to any of the five smoking-related variables [Mother smoked during this pregnancy? Was child's mother exposed to smoke during pregnancy? Did child live with someone who smoked? (parent's survey) Does the child live with someone who smokes? (child's survey) Does the child work with someone who smokes? (child's survey)]. We defined any passive smoke exposure *in utero* by an affirmative answer to the question "Mother smoked during this pregnancy?" or "Was child's mother exposed to smoke during pregnancy?"

Sample size

Based on standard formulae for sample size for a pair-matched study design [28], a total of 146 pairs of cases and controls was calculated to determine a risk effect of not breastfeeding in infancy on the development of CD consistent with that reported in the literature [$\psi=1.9$ with a two-sided level of statistical significance $\alpha=0.05$ and a power of $(1-\beta)=0.8$]. This calculation was based on an assumed proportion of individuals who were not initially breastfed during infancy of 40%. This sample size was calculated to evaluate the hypothesis that individuals who were not breastfed in infancy are more likely to develop CD than are individuals who were breastfed in infancy.

Matching process

We began recruitment of subjects in June 2004 and continued through to June 2010. Altogether, we enrolled 258 cases and 202 unaffected sibling controls in the study. Of these, 216 cases and 194 controls returned completed surveys. After excluding ineligible subjects, subjects who withdrew after enrollment, duplicate controls from within an individual family, and controls who were too young or too old to match any cases, 185 cases and 137 controls were available for matching. We identified optimal case-control matches by age and sex and confirmed that no cases were matched to their own sibling. To address the considerable missing data for most specific illnesses during infancy, we conducted sensitivity analyses. Based on these analyses, we eliminated all subjects who were missing answers for diarrhea ($n=34$). Also, we eliminated all matched pairs ($m=8$ pairs) whose age difference did not meet the study criteria of 12 months. This resulted in a final sample of 152 subjects or 76 pairs (43 male pairs, 33 female pairs) for further analysis.

Statistical analysis

We compared data for cases and their matched controls by chi-squared test or McNemar's test for categorical variables and by independent samples t-test for continuous variables with unequal sample sizes. We calculated matched odds ratios and 95% confidence intervals to estimate the strength and precision of the association between breastfeeding and the risk of developing CD. We assessed confounding and interaction initially through analysis comparing disease status to breastfeeding status stratified by each of the potential confounding factors. We repeated the analysis of cases and controls comparing disease status and breastfeeding status stratified by each of the potential confounding factors on the final sample of 152 subjects. We assessed interaction effect by the Breslow-Day test for homogeneity of the odds ratio. Using the non-zero correlation test, we assessed for a dose response in the relationship between duration of breastfeeding and CD, stratified by each of the potential confounding factors. We used conditional logistic regression (conditioned on age- and sex-matched pairs) to generate crude and adjusted matched odds

ratios and 95% confidence intervals of the association between breastfeeding and CD. We assessed for interaction between the exposure and other potential risk factors by entering appropriate interaction terms into the model and assessing effect on model fit. To address the primary hypothesis, the exposure variable was a dichotomous variable or an ordinal variable. We performed all statistical analyses using SAS version 9.1.

Results

Univariate analysis

As expected based on matching criteria, cases and controls were of similar age at enrollment (13.7 ± 3.1 years for cases vs. 13.8 ± 3.6 years for controls, $t=0.23$, $df=150$, $p=0.82$) (see Table 1) and cases were of similar age at time of diagnosis (11.6 ± 3.1 years for males vs. 12.7 ± 2.8 years for females, $p=0.15$). Of the 152 subjects, 142 were white, 8 were African-American, 1 was another race, and 1 did not specify race. Mother's age at child's birth did not differ between cases and controls (29.2 ± 4.0 years for cases vs. 29.9 ± 4.2 years for controls, $p=0.25$). Many subjects reported a family history of CD ($n=42$) or UC ($n=16$) but even more reported a family history of other gastrointestinal conditions ($n=57$). None of the family history reports differed significantly between cases and controls.

Overall, 105 (69.1%) subjects were breastfed, and of those, duration of breastfeeding was missing for only 1 subject (see Table 1). The proportion of subjects who were breastfed did not differ significantly between cases and controls [$\psi=0.65$ (0.32–1.30), $\chi^2=1.51$, $p=0.22$]. The duration of breastfeeding was almost 2 months shorter in cases than controls but the difference was not significant (6.8 ± 6.4 months vs. 8.7 ± 8.4 months, $p=0.21$). Cases were introduced to formula at a slightly younger age than controls (1.8 ± 2.7 months vs. 2.4 ± 3.1 months, $p=0.18$) and were given solid foods at about the same age as controls (4.6 ± 2.1 months vs. 4.9 ± 2.5 months, $p=0.40$) but neither difference was statistically significant. Cases and controls did not differ significantly with regard to the types of formula or the first solid food they were fed.

Ear infections and fever were the most common illnesses reported in infancy, but the frequency of these did not differ between cases and controls. Vomiting and diarrhea during infancy were less commonly reported, but there was no difference in frequency of these illnesses between cases and controls. Cases were more likely to have had other illnesses ($p=0.04$), respiratory distress ($p=0.05$), or hospitalization ($p=0.11$) in infancy than were controls. Cases and controls did not differ with respect to the derived variable "any illness." The number of antibiotic courses taken by subjects in the three-year period prior to diagnosis did not differ significantly between cases and controls.

Maternal OC use before and during pregnancy did not differ significantly between cases and controls. Only 18 (11.8 %) mothers smoked during pregnancy but 37 (24.3 %) lived or worked closely with a smoker during pregnancy. Of the children, 42 (27.6 %) lived with a smoker during their childhood (by parental report) but only 8 (5.3 %) smoked (by self-report) prior to enrollment in the study. Of these 8 children, 4 were cases and 4 were controls. Only one of these cases reported smoking prior to diagnosis with CD. There were no significant differences between cases and controls for any of the specific smoking factors listed above. However, using the derived smoking variables, cases were more likely than controls to report any smoke exposure ($p=0.03$), any passive smoke exposure ($p=0.04$), and any passive *in utero* smoke exposure ($p=0.05$).

Of the female subjects, 43 (65.2%) were post-menarche, which was not significantly different between cases and controls, although 6 subjects were too young to complete this portion of the questionnaire. For those who reported an age at menarche, there was no difference between cases and controls (13.0 ± 2.1 years vs. 12.7 ± 1.0 years, $n=43$, $p=0.52$). Of the female subjects, 9 (13.6 %) reported taking OC, 6 cases and 3 controls, and 2 reported having been pregnant, 1 case and 1 control. The case who had been pregnant did not indicate whether this was before or after her diagnosis with CD. Neither of these differed significantly between cases and controls although controls tended to be younger than cases at the age they began taking OC ($p=0.07$) and had taken OC for a longer period of time than cases ($p=0.04$).

Table 1. Characteristics of the study population

| | Cases (n, %) | Controls (n, %) |
|--|--------------|-----------------|
| Female | 33 (43.4) | 33 (43.4) |
| *Age at enrollment (years), mean +/- SD | 13.7 +/- 3.1 | 13.8 +/- 3.6 |
| *Maternal age at child's birth (years), mean +/- SD | 29.2 +/- 4.0 | 29.9 +/- 4.2 |
| Breastfed | 49 (64.5) | 56 (73.7) |
| *Breastfeeding duration (months), mean +/- SD | 6.8 +/- 6.4 | 8.7 +/- 8.4 |
| *Age at introduction of formula (months), mean +/- SD | 1.8 +/- 2.7 | 2.4 +/- 3.1 |
| *Age at introduction of first food (months), mean +/- SD | 4.6 +/- 2.1 | 4.9 +/- 2.5 |
| Family history of CD | 18 (23.7) | 24 (31.6) |
| Family history of UC | 8 (10.5) | 8 (10.5) |
| Family history of other GI conditions ¹ | 33 (43.4) | 24 (31.6) |
| Vomiting | 14 (18.4) | 8 (10.5) |
| Diarrhea | 15 (19.7) | 11 (14.5) |
| Ear infections | 34 (44.7) | 33 (43.4) |
| Fever | 29 (38.2) | 23 (30.3) |
| ††Respiratory distress | 8 (10.5) | 2 (2.6) |
| †Other illness | 10 (13.2) | 3 (3.9) |
| Hospitalization | 11 (14.5) | 5 (6.6) |
| Any illness | 47 (61.8) | 43 (56.6) |
| Mother smoked during the pregnancy | 10 (13.2) | 8 (10.5) |
| Child's mother was exposed to smoke during pregnancy | 21 (27.6) | 16 (21.1) |
| Child lived with someone who smoked (parent survey) | 24 (31.6) | 18 (23.7) |
| Child ever smoked (child survey) | 4 (5.3) | 4 (5.3) |
| Child lives with someone who smokes (child survey) | 14 (18.4) | 10 (13.2) |
| Child works with someone who smokes (child survey) | 6 (7.9) | 4 (5.3) |
| *Age at menarche (years), mean +/- SD | 13.0 +/- 2.1 | 12.7 +/- 1.0 |
| †Child was exposed to any smoking | 34 (44.7) | 21 (27.6) |
| †Child was passively exposed to smoking | 33 (43.4) | 21 (27.6) |
| †Child was passively exposed in utero to smoking | 27 (35.5) | 16 (21.1) |
| Child was passively exposed to any smoking after birth | 30 (39.5) | 22 (28.9) |

Characteristics of the study population of unmatched cases and controls were compared by chi-squared test of proportions, except as otherwise indicated.

- *comparison by independent samples t-test.
- †p-value < 0.05 for comparison between cases and controls (or for age at diagnosis between males and females)
- ††p-value < 0.10 for comparison between cases and controls

¹ Family history of other gastrointestinal conditions includes gastroesophageal reflux disease (2 cases, 4 controls), cancer (7 cases, 6 controls), diverticular disease (3 cases, 5 controls), irritable bowel syndrome (14 cases, 9 controls), and motility issues (2 cases, 1 control).

CD: Crohn's disease; UC: ulcerative colitis

Assessment for confounding and interaction

The association between breastfeeding and development of CD tended to be protective but was not statistically significant [$\psi = 0.63$ (0.31-1.30) $n=152$, McNemar’s $\chi^2 = 1.58$, $p=0.21$]. Stratification by sex or by age did not demonstrate confounding in this association as expected since these were the matching criteria. Stratified analysis of the association between breastfeeding and development of CD failed to demonstrate confounding by or interaction with any of the available covariates, including diarrhea during infancy and all of the smoking exposures.

Assessment for dose response

To assess for a dose response of breastfeeding duration on the development of CD, a trend test was conducted. As evidenced by the non-zero correlation test, there was no statistically significant trend in development of CD based on duration of breastfeeding (none, less than or equal to 3 months, greater than 3 months but less than or equal to 6 months, and greater than 6 months) assessed by matched pairs. As evidenced by the non-zero correlation test, there was no significant trend in development of CD with increased duration of breastfeeding for pair matched analysis when stratified by any covariates.

Multivariable analysis

Based on the analysis to this point, multivariable analysis was conducted using breastfeeding as a dichotomous variable or breastfeeding duration as an ordinal variable (with dummy variables using no breastfeeding as the referent). No interaction was demonstrated by conditional logistic regression analysis and, therefore, no interaction terms were included in further model development. Family history of CD, family history of UC, diarrhea in infancy, hospitalization, and any smoke exposure were included as covariates based both on support from previous literature and on the results of the preceding analysis. Any illness was redefined as any illness in infancy other than diarrhea or hospitalization and was included as a covariate in the

model. Maternal age was included as a covariate due to support from the previous literature. Maternal age was entered into models as a continuous variable but all other covariates were entered into the models as dichotomous variables.

In the final model with breastfeeding as a dichotomous exposure, breastfeeding tended to be a protective factor for the development of CD [$\psi = 0.61$ (0.27–1.38) $n=152$] controlling for family history of CD and of UC, diarrhea, hospitalization, and any illness other than diarrhea in infancy, maternal age, and any smoke exposure (see Table 2).

Table 2. Logistic regression models for breastfeeding as a dichotomous variable

| Factor | Initial model with breastfeeding only | Final model with breastfeeding and all covariates |
|---------------------------------|---------------------------------------|---|
| | Point estimate (95% CI) | Point estimate (95% CI) |
| Breastfeeding (dichotomous) | 0.63 (0.31–1.30) | 0.61 (0.27–1.38) |
| Family history of CD | | 0.64 (0.26–1.60) |
| Family history of UC | | 1.19 (0.31–4.59) |
| Diarrhea | | 1.17 (0.38–3.61) |
| Hospitalization | | 1.96 (0.56–6.82) |
| Any illness other than diarrhea | | 1.26 (0.56–2.87) |
| Maternal age | | 0.96 (0.88–1.05) |
| Any smoke exposure | | 2.29 (1.05–5.03) |

CD: Crohn’s disease; CI: confidence interval; UC: ulcerative colitis

In the final model with breastfeeding as an ordinal exposure, breastfeeding tended to be protective in the development of CD controlling for the same factors listed above. The greatest protective effect in both models was for intermediate duration of breastfeeding rather than the longest duration of breastfeeding (see Table 3). Any smoke exposure was a statistically significant risk factor for the development of CD with an odds ratio of greater than 2 in both final models (see Tables 2 and 3).

Table 3. Logistic regression models for breastfeeding using dummy variables for ordinal exposure

| Factor | Initial model with breastfeeding only | Final model with breastfeeding and all covariates |
|--|---------------------------------------|---|
| | Point estimate (95% CI) | Point estimate (95% CI) |
| Breastfeeding duration (≤ 3 months vs. none) | 0.89 (0.34–2.34) | 0.80 (0.27–2.41) |
| Breastfeeding duration (> 3 but ≤ 6 months vs. none) | 0.53 (0.18–1.62) | 0.40 (0.11–1.43) |
| Breastfeeding duration (> 6 months vs. none) | 0.57 (0.25–1.27) | 0.62 (0.24–1.58) |
| Family history of CD | | 0.67 (0.26–1.68) |
| Family history of UC | | 1.32 (0.34–5.15) |
| Diarrhea | | 1.14 (0.35–3.72) |
| Hospitalization | | 2.22 (0.61–8.14) |
| Any illness other than diarrhea | | 1.23 (0.53–2.88) |
| Maternal age | | 0.96 (0.87–1.05) |
| Any smoke exposure | | 2.35 (1.06–5.18) |

CD: Crohn’s disease; CI: confidence interval; UC: ulcerative colitis

Discussion

We did not find a statistically significant association between breastfeeding in infancy and development of CD in children. We found a tendency toward a protective effect of breastfeeding in infancy on the development of CD in children [$\psi=0.61$ (0.27–1.38)] controlling for family history of CD and of UC, diarrhea, hospitalization, and any illness other than diarrhea in infancy, maternal age, and any smoke exposure. We also found that longer duration of breastfeeding tended to be more protective on the development of CD [$\psi=0.80$ (0.27–2.41), $\psi=0.40$ (0.11–1.43), and $\psi=0.62$ (0.24–1.58) for ≤ 3 months, > 3 but ≤ 6 months, and > 6 months respectively, vs. none] controlling for family history of CD and of UC, diarrhea, hospitalization, and any illness other than diarrhea in infancy, maternal age, and any smoke exposure.

We could not assess the effect of sex or age on these associations since these were matching criteria for the study. Furthermore, we could not assess the effect of race or ethnicity on this association since there was limited variability in these characteristics within the study population. We could not adequately assess the effect of birth order or number of siblings on the association since, by design of the study, cases could be only children but the controls could not. We were able to assess the individual effect of family history, history of specific illnesses in infancy, and history of smoking exposure on the association controlling for the other factors listed. Of these, family history of CD tended to have a protective effect and family history of UC tended to be a risk factor for the development of CD. Many of the specific illnesses in infancy appeared to confound the association between breastfeeding and CD but this may have been influenced by the large proportion of missing data and the small number of affirmative answers for many of these illnesses. Many of the smoking exposures appeared to be risk factors for the development of CD. Some of these may have been influenced by the proportion of missing data. However, the derived variable, any smoke exposure, was consistently a significant risk factor with odds ratios greater than 2.0 for the development of CD.

Even though this study did not demonstrate statistically significant findings, its findings support those of the preponderance of the literature, which shows a protective effect of breastfeeding on the development of CD [20]. Furthermore, this study demonstrated a protective effect of a similar magnitude to that reported in previous studies [29–31]. Additionally, this study supported a beneficial effect to longer duration of breastfeeding on the development of CD, as did previous studies [32, 33].

There are several important strengths of the design of this study. First, the inclusion criteria for cases were quite rigorous, ensuring that only individuals with definitive CD were included as cases, and further ensuring that the cases were pediatric cases and residents of Wisconsin at diagnosis thereby creating a pediatric, population-based sample of cases. Second, since the study was population-based rather than clinic-based, it is less likely to suffer from selection bias. However, controls were selected from a pool of eligible unaffected siblings of children with CD and

matched by age and sex to unrelated cases thus demonstrating the protective effects of breastfeeding in a sample of children with a potential risk of acquiring CD comparable to that of their affected siblings. Selection of controls from the general population would be unlikely to provide controls with a comparable genetic susceptibility to CD since the prevalence of the known genetic susceptibility variants is lower in the general population than it would be expected to be in siblings of cases. Furthermore, this is a study of pediatric-onset CD. Although pediatric-onset CD and adult-onset CD are likely the same condition, one epidemiologic feature is strikingly different: In pediatric-onset CD, males are more commonly affected, whereas in adult-onset CD, females are more commonly affected. Consistent with previous studies, we found more affected males than females in our study population.

We hypothesized that breastfeeding would be a statistically significant protective factor for CD, but this was not the case. This was most likely related to inadequate sample size. Our sample size calculations suggested that we would need 146 pairs of subjects to achieve statistical significance. We were only able to obtain 114 pairs and some of these had to be eliminated due to missing data on a key confounder (i.e., diarrhea during infancy) or inappropriate matching (i.e., age difference between case and control greater than 12 months). Thus, our final analysis was conducted on only 76 pairs. Furthermore, the sample size calculations were based on the assumption that 60% of subjects would have been breastfed in infancy. However, the rate of breastfeeding for our study sample was 73.7% for controls. Based on this rate, our study had power of only 23.6% to demonstrate the expected protective effect of breastfeeding. One other important factor that was not addressed by this study was exclusivity of breastfeeding. Some investigators have found a potentially greater protective effect of exclusive breastfeeding than non-exclusive breastfeeding [34]. We were unable to determine exclusivity of breastfeeding in our study.

The design of the questionnaire may have been confusing, causing some participants to skip particular questions or to answer questions incorrectly. Controls were asked if they had a family history of CD (other than the sibling who was

enrolled in the study) and, if so, to identify the affected relative. Some controls indicated that their only affected relative was their sibling so these were considered negative for family history; however, some controls did not indicate who the affected relative was, so they could have been misclassified as positive for family history of CD if the only affected relative was their sibling. The use of prevalent cases in the study may have allowed misclassification related to the timing of an exposure relative to onset of disease. The only factors likely to have been misclassified were antibiotic use and the passive smoke exposures obtained from the adolescent surveys since the timing of these exposures was not as easy to confirm. With regard to the controls, there might have been greater risk of misclassification since the reference date was not always clear at the time the survey was completed.

While it is likely that breastfeeding has a protective effect on the development of CD, this study cannot definitively demonstrate this. The retrospective design of this study, and of most published studies, leaves some uncertainty. Even in a study of pediatric-onset CD, the interval from infancy to the development of disease is typically more than ten years. This interval is adequate for recall of major factors such as whether a child was breastfed or not but perhaps not as much so for specific details such as the duration of breastfeeding, the first solid food, or occurrence of specific illnesses in infancy. Ideally, these specific details would be better obtained prospectively. Future studies should be conducted in a manner that allows for the prospective collection of data regarding breastfeeding, including duration and exclusivity, and other factors in infancy and early childhood. Finally, future studies should be designed to consider genetic susceptibility to CD so that all factors involved in the etiopathogenesis of CD are included.

References

1. Calkins BM, Lilienfeld AM, Garland CF, Mendeloff AI. Trends in incidence rates of ulcerative colitis and Crohn's disease. *Dig Dis Sci*. 1984 Oct;29(10):913-20.
2. Bernstein CN, Blanchard JF, Rawsthorne P, Wajda A. Epidemiology of Crohn's disease and ulcerative colitis

- in a central Canadian province: A population-based study. *Am J Epidemiol.* 1999 May 15;149(10):916-24.
3. Barton JR, Gillon S, Ferguson A. Incidence of inflammatory bowel disease in Scottish children between 1968 and 1983; marginal fall in ulcerative colitis, three-fold rise in Crohn's disease. *Gut.* 1989 May;30(5):618-22.
 4. Cosgrove M, Al-Atia RF, Jenkins HR. The epidemiology of paediatric inflammatory bowel disease. *Arch Dis Child.* 1996 May;74(5):460-1.
 5. Armitage E, Drummond H, Ghosh S, Ferguson A. Incidence of juvenile-onset Crohn's disease in Scotland. *Lancet.* 1999 May 1;353:1496-7.
 6. Sawczenko A, Sandhu BK, Logan RF, Jenkins H, Taylor CJ, Mian S, Lynn R. Prospective survey of childhood inflammatory bowel disease in the British Isles. *Lancet.* 2001 Apr 7;357(9262):1093-4.
 7. Kugathasan S, Judd RH, Hoffmann RG, Heikenen J, Telega G, Khan F, et al. Epidemiologic and clinical characteristics of children with newly diagnosed inflammatory bowel disease in Wisconsin: A statewide population-based study. *J Pediatr.* 2003 Oct;143(4):525-31.
 8. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature.* 2001 May 31;411:599-603.
 9. Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature.* 2001 May 31;411:603-6.
 10. Monsen U, Brostrom O, Nordenvall B, Sorstad J, Hellers G. Prevalence of inflammatory bowel disease among relatives of patients with ulcerative colitis. *Scand J Gastroenterol.* 1987 Mar;22(2):214-8.
 11. Monsen U, Bernell O, Johansson C, Hellers G. Prevalence of inflammatory bowel disease among relatives of patients with Crohn's disease. *Scand J Gastroenterol.* 1991 Mar;26(3):302-6.
 12. Tysk C, Lindberg E, Jarnerot G, Floderus-Myrhed B. Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. *Gut.* 1988 Jul;29(7):990-6.
 13. Thompson NP, Driscoll R, Pounder RE, Wakefield AJ. Genetics versus environment in inflammatory bowel disease: Results of a British twin study. *BMJ.* 1996 Jan 13;312:95-6.
 14. Subhani J, Montgomery S, Pounder R, Wakefield A. Concordance rates of twins and siblings in inflammatory bowel disease. *Gut.* 1998;42:A40.
 15. Gilat T, Langman M, Rozen P. Environmental factors in inflammatory bowel disease. *Front Gastrointest Res.* 1986;11:158-76.
 16. Sonnenberg A, McCarty DJ, Jacobsen SJ. Geographic variation of inflammatory bowel disease within the United States. *Gastroenterology.* 1991 Jan;100(1):143-9.
 17. Mezoff EA, Aly H. The winding road to understanding the neonatal origins of inflammatory gastrointestinal disorders. *J Pediatr Gastroenterol Nutr.* 2013 Nov;57(5):543-9.
 18. Ponder A, Long MD. A clinical review of recent findings in the epidemiology of inflammatory bowel disease. *Clin Epidemiol.* 2013 Jul 25;5:237-47.
 19. Abegunde AT, Muhammad BH, Bhatti O, and Ali T. Environmental risk factors for inflammatory bowel diseases: Evidence based literature review. *World J Gastroenterol* 2016 Jul 21;22(27):6296-317.
 20. Klement E, Cohen RV, Boxman J, Joseph A, Reif S. Breastfeeding and risk of inflammatory bowel disease: A systematic review with meta-analysis. *Am J Clin Nutr.* 2004 Nov;80(5):1342-52.
 21. Ng SC, Bernstein CN, Vatn MH, Lakatos PL, Loftus EV, Jr, Tysk C, et al. Geographical variability and environmental risk factors in inflammatory bowel disease. *Gut* 2013;62:630-49.
 22. Ekbohm A, Adami HO, Helmick CG, Jonzon A, Zack MM. Perinatal risk factors for inflammatory bowel disease: A case-control study. *Am J Epidemiol.* 1990;132(6): 1111-9.
 23. Thompson NP, Montgomery SM, Wadsworth ME, Pounder RE, Wakefield AJ. Early determinants of inflammatory bowel disease: Use of two national longitudinal birth cohorts. *Eur J Gastroenterol Hepatol.* 2000;12(1): 25-30.
 24. Mikhailov TA, Furner SE. Breastfeeding and genetic factors in the etiology of inflammatory bowel disease in children. *World J Gastroenterol* 2009;15:270-9.
 25. Davis MK. Breastfeeding and chronic disease in childhood and adolescence. *Pediatr Clin North Am.* 2001 Feb;48(1):125-41, ix.
 26. Carver JD, Barness LA. Trophic factors for the gastrointestinal tract. *Clin Perinatol.* 1996 Jun;23(2):265-85.
 27. Shanahan F. Crohn's disease. *Lancet.* 2002 Jan 5;359:62-9.
 28. Schlesselman, JJ, Stolley, PD. Case-control studies: design, conduct, analysis. New York: Oxford University Press; 1982.
 29. Corrao G, Tragnone A, Caprilli R, Trallori G, Papi C, Andreoli A, et al. Risk of inflammatory bowel disease attributable to smoking, oral contraception and breastfeeding in Italy: A nationwide case-control study. Cooperative Investigators of the Italian Group for the Study of the Colon and the Rectum (GISC). *Int J Epidemiol.* 1998 Jun;27(3):397-404.
 30. Klement E, Reif S. Breastfeeding and risk of inflammatory bowel disease. *Am J Clin Nutr.* 2005 Aug;82(2):486.

31. Gruber M, Marshall JR, Zielezny M, Lance P. A case-control study to examine the influence of maternal perinatal behaviors on the incidence of Crohn's disease. *Gastroenterol Nurs*. 1996 Mar-Apr;19(2):53-9.
32. Bergstrand O, Hellers G. Breast-feeding during infancy in patients who later develop Crohn's disease. *Scand J Gastroenterol*. 1983 Oct;18(7):903-6.
33. Rigas A, Rigas B, Glassman M, Yen YY, Lan SJ, Petridou E, et al. Breast-feeding and maternal smoking in the etiology of Crohn's disease and ulcerative colitis in childhood. *Ann Epidemiol*. 1993 Jul;3(4):387-92.
34. Koletzko S, Sherman P, Corey M, Griffiths A, Smith C. Role of infant feeding practices in development of Crohn's disease in childhood. *BMJ (Clin Res Ed)* 1989;298:1617-8.